

Are Currency Crises Predictable? A Test

ANDREW BERG and CATHERINE PATTILLO*

This paper evaluates three models for predicting currency crises that were proposed before 1997. The idea is to answer the question: if we had been using these models in late 1996, how well armed would we have been to predict the Asian crisis? The results are mixed. Two of the models fail to provide useful forecasts. One model provides forecasts that are somewhat informative though still not reliable. Plausible modifications to this model improve its performance, providing some hope that future models may do better. This exercise suggests, though, that while forecasting models may help indicate vulnerability to crisis, the predictive power of even the best of them may be limited. [JEL F31, F47]

In recent years, a number of researchers have claimed success in systematically predicting which countries are more likely to suffer currency crises. The Asian crisis has stimulated further work in this area, with several papers already claiming to be able to “predict” the incidence of this crisis using pre-crisis data.¹

It may seem unlikely that it should be possible to systematically predict currency crises. It is reasonable to doubt that sharp and predictable movements in the exchange rate are consistent with the actions of forward-looking speculators. Early theoretical models of currency crises suggested, however, that crises may

*Andrew Berg and Catherine Pattillo are Economists in the Research Department. They would like to thank, without implication, Graciela Kaminsky, Andy Rose, and Aaron Tornell for help reproducing and interpreting their results, Brooks Dana Calvo, Maria Costa, Manzoor Gill, and Nada Mora for superb research assistance, and Eduardo Borensztein, Robert Flood, Steve Kamin, an anonymous referee, and many IMF colleagues for useful comments.

¹IMF (1998), Kaminsky (1998a and 1998b), Radelet and Sachs (1998b), Corsetti, Pesenti, and Roubini (1998a), and Tornell (1998), among others.

be predictable even with fully rational speculators.² In “second-generation” models, a country may be in a situation in which an attack, while not inevitable, might succeed if it were to take place; the exact timing of crises would be essentially unpredictable. Even here, though, it may be possible to identify whether a country is in a zone of vulnerability—that is, whether fundamentals are sufficiently weak that a shift in expectations could cause a crisis. In this case, the relative vulnerability of different countries might predict the relative probabilities of crises in response to a shock such as a global downturn in confidence in emerging markets.³

It is one thing to say that currency crises may be predictable in general, however, and another that econometric models estimated using historical data on a panel or cross section of countries can foretell crises with any degree of accuracy. It is an open question whether crises are sufficiently similar across countries and over time to allow generalizations from past experience. For example, models estimated over countries without capital mobility may not work in a world of capital mobility.⁴ Moreover, many factors that may indicate a higher probability of crisis, such as inadequate banking supervision or a vulnerable political situation, are not easily quantified.

The possible endogeneity of policy to the risk of crisis may also limit the predictability of crises. For example, authorities within a country, or their creditors, might react to signals so as to avoid crises.⁵ Policymakers are often fighting the previous battle, so they are likely to respond to the most obvious indicators from a previous crisis. On the other hand, a focus by market participants on a particular variable could result in its precipitating a crisis where one might not otherwise have occurred.

The flurry of work between the 1994 and 1997 crises and the large number of crises observed in 1997 provides an excellent opportunity to test existing state-of-the-art “early warning systems” out of sample. The 1997 Asian crises that we look at here present special challenges, however, on two grounds. First, many analysts have argued that the causes of the Asian crises lie not in the traditional macroeconomic fundamentals but rather in structural and microeconomic problems such as weak banking supervision, poor corporate governance, and even corruption.⁶ Data on these are hard to come by, and the emphasis on these issues is somewhat new, so the available empirical models focus rather on the typical macroeconomic variables. This bodes ill for the predictability of the Asian crises with these models. A contrasting line of thought, but also with pessimistic

²Krugman (1979). In this model, though, the exchange rate does not jump and indeed there are no capital gains or losses of any sort at the point of crisis, so the relevance to the type of crises most people have in mind may be limited.

³See Flood and Marion (1998) for a survey of this literature.

⁴Flood and Marion (1994) discuss and present some evidence on the predictability of currency crises in capital-controlled developing economies.

⁵Initially successful early warning systems might thus cease to work following publication. This is a version of the Lucas critique.

⁶Radelet and Sachs (1998a) emphasize the inability of fundamentals to explain the crises, while Corsetti, Pesenti, and Roubini (1998b and 1998c) focus more on the structural and microeconomic explanations. See Lane and others (1999) and Berg (1999) and references therein for overviews.

implications for us, is that the Asian crises were largely “bank run” phenomena—panic attacks against otherwise viable exchange rate regimes. This distinguishes these crises from those emphasized in most of the empirical models, and suggests that, at best, only a few variables that measure exposure to panicky capital outflows would be helpful predictors of crisis.⁷ When a crisis will strike would be difficult or impossible to foretell.

On the other hand, the 1994 Mexico crisis, which was the immediate inspiration for much of the recent work on crises, does not in many respects look that different from Thailand’s. Sachs (1997) argues that Thailand’s 1997 crisis “has the same hallmarks [as the 1994 crisis]: overvaluation of the real exchange rate, coupled with booming bank lending, heavily directed at real estate.” In any case, each set of new crises always presents some new features, so the existence of some novelty in the Asian crises does not invalidate them as tests of the models we consider.

Ultimately, the question of whether crises are predictable can only be settled in practice. The recent work claiming success in predicting crises has focused almost exclusively on in-sample prediction—that is, on formulating and estimating a model using data on a set of crises, and then judging success by the plausibility of the estimated parameters and the size of the prediction errors for this set of crises.⁸ The key test is not, however, the ability to fit a set of observations after the fact, but the prediction of future crises. Given the relatively small number of crises in the historical data, the danger is acute that specification searches through the large number of potential predictive variables may yield spurious success in “explaining” crises within the sample. The possibility that the determinants of crises may vary importantly through time also suggests the importance of testing the models out of sample.

This paper evaluates three different models proposed before 1997 for predicting currency crises. The idea is to try to answer the question: if we had been using these models in late 1996, how well armed would we have been to predict the Asian crisis? For each of the three models, we duplicate the original results as closely as possible. We then reestimate the models using data through 1996, as would have a researcher who at the end of 1996 aimed to predict crises the following year. We use two samples of countries: the same as the original paper, and another common sample for purposes of comparing the three methods. We then use the models to forecast events in 1997. We generate a ranking of countries according to predicted probability or severity of crisis in 1997 for each model, and then compare the predicted and actual rankings.

We chose the following three approaches based on their promise as early warning systems, their potential applicability to the 1997 crises, and their success within sample:

⁷None of the precrisis models used a measure of short-term external debt relative to reserves, a variable much emphasized by many advocates of the “bank run” interpretation of these crises, such as Radelet and Sachs (1998b).

⁸Exceptions are Tornell (1998), discussed below, and Kaminsky (1998a), which, while it presents out-of-sample estimates of the probability of currency crisis, does not provide tests of whether these forecasts are better than, for example, guesswork. In addition, Furman and Stiglitz (1998) carry out an exercise similar to ours. Their conclusions are largely consistent with our own, with some differences as noted below.

- Kaminsky, Lizondo, and Reinhart (1998) (hereafter KLR) monitor a large set of monthly indicators that signal a crisis whenever they cross a certain threshold. This approach has the potential attraction that it produces thresholds beyond which a crisis is more likely. This accords with the common practice of establishing certain warning zones, such as current account deficits beyond 5 percent of GDP or reserves less than three months of imports. The authors claim some success in developing a set of indicators that reliably predict the likelihood of crisis. Moreover, Kaminsky (1998a and 1998b) and Goldstein (1998) have asserted that this method can be applied successfully to the 1997 crises.
- Frankel and Rose (1996) (FR) develop a probit model of currency crashes in a large sample of developing countries. Their use of annual data permits them to look at variables, such as the composition of external debt, that are available only at that frequency.
- Sachs, Tornell, and Velasco (1996) (STV) restrict their attention to a cross section of countries in 1995, analyzing the incidence of the “tequila effect” following the Mexico crisis. They concentrate on a more structured hypothesis about the cause of this particular episode, emphasizing interactions among weak banking systems, overvalued real exchange rates, and low reserves. They claim to explain most of the cross-country pattern of currency crisis in emerging markets in 1994–95. Their approach has also been applied to analyzing the Asian crisis.⁹

I. Three Methods for Predicting Crises

Kaminsky, Lizondo, and Reinhart (1998) Signals Approach

The Model

For KLR, a currency crisis occurs when a weighted average of monthly percentage depreciations in the exchange rate and monthly percentage declines in reserves exceeds its mean by more than three standard deviations.¹⁰ KLR propose the monitoring of several indicators that may tend to exhibit unusual behavior during a 24-month window prior to a crisis. They choose 15 candidate indicator variables based on theoretical priors and on the availability of monthly data.¹¹ An indicator issues a signal whenever it moves beyond a given threshold level.

⁹Tornell (1998), Radelet and Sachs (1998b), Corsetti, Pesenti, and Roubini (1998a), and IMF (1998) estimate variants of STV for 1997, all with some success.

¹⁰Weights are calculated so that the variance of the two components of the index are equal. See Berg and Pattillo (1998) as well as KLR for further details regarding the methodology.

¹¹Indicators are (1) international reserves in U.S. dollars; (2) imports in U.S. dollars; (3) exports in U.S. dollars; (4) terms of trade; (5) deviations of the real exchange rate from a deterministic time trend (in percentage terms); (6) the differential between foreign and domestic real interest rates on deposits; (7) “excess” real M1 balances, where excess is defined as the residuals from a regression of real M1 balances on real GDP, inflation, and a deterministic time trend; (8) the money multiplier of M2; (9) the ratio of domestic credit to GDP; (10) the real interest rate on deposits; (11) the ratio of (nominal) lending to deposit rates; (12) the stock of commercial bank deposits; (13) the ratio of broad money to gross international reserves; (14) an index of output; and (15) an index of equity prices measured in U.S. dollars. The indicator is defined as the annual percentage change in the level of the variable (except for the deviation of the real exchange rate from trend, “excess” real M1 balances, and the three interest rate variables).

We can consider the performance of each indicator in terms of the matrix at right. Cell A represents the number of months in which the indicator issued a good signal, B is the number of months in which the indicator issued a bad signal or “noise,” C is the number of months in which the indicator failed to issue a signal that would have been a good signal, and D is the number of months in which the indicator did not issue a signal that would have been a bad signal. For each indicator, KLR find the “optimal” threshold, defined as that threshold that minimizes the noise-to-signal ratio B/A .¹²

	Crisis within 24 months	No crisis within 24 months
Signal was issued	A	B
No signal was issued	C	D

The thresholds are calculated in terms of the percentiles of each country’s distribution for the variable in question. An optimal threshold for a given predictor, such as domestic credit growth, might be 80, for example, meaning that a signal is considered to be issued whenever domestic credit growth in a given country is in the highest 20 percent of observations for that country. The optimal threshold is constrained to be the same across countries. Thus, minimizing the noise-to-signal ratio for the sample of countries yields an optimal threshold percentile for each indicator that is the same for all countries. The corresponding country-specific threshold value of the underlying variable associated with that percentile will differ across countries, however.

The KLR approach is bivariate, in that each indicator is analyzed, and optimal thresholds calculated, separately. Kaminsky (1998a) calculates a single composite indicator of crisis as a weighted sum of the indicators, where each indicator is weighted by the inverse of its noise-to-signal ratio. She then calculates a probability of crisis for each value of the aggregate index by observing how often within the sample a given value of the aggregate index is followed by a crisis within 24 months.

Table 1 presents an analog of a regression output for the KLR model, as estimated in the in-sample period of 1970 to April 1995.¹³ The first column shows the noise-to-signal ratio estimated for each indicator (defined as the number of bad signals as a share of possible bad signals $(B/(B+D))$ divided by the number of good signals as a share of possible good signals $(A/(A+C))$). Column 2 shows how much higher is the probability of a crisis within 24 months when the indicator emits a signal than when it does not (within sample). When the noise-to-signal ratio is less than 1, this number is positive, implying that crises are more likely when the indicator signals than when it does not. Indicators with noise-to-signal ratios equal to or above unity are not useful in anticipating crises.

¹²If the absence of a crisis within 24 months is considered the null hypothesis, then observations of type B are Type I errors, while observations of type C are Type II errors. The procedure can be thought of as minimizing the ratio of Type I errors, as a share of tranquil periods $(B/(B+D))$ to 1 – Type II errors as a share of crisis periods $(A/(A+C))$.

¹³The in-sample period for the KLR model stops in April 1995 because of the 24-month prediction window. A person implementing the KLR model in April 1997 (right before the Thai crisis) would estimate the thresholds based on the performance of predictive variables measured only through April 1995, since after that month it would be impossible to know (yet) whether a crisis was to occur within 24 months.

Table 1. Performance of Indicators—In-Sample

Indicator	23-Country Sample, 1970–April 1995		
	Noise/signal (adjusted) ^a (1)	$P(\text{crisis/signal})$ $-P(\text{crisis})^b$ (2)	Number of crises with data (3)
Real exchange rate ^c	0.25	29	70
M2/reserves growth rate	0.39	17	68
Export growth rate	0.60	9	68
International reserves growth rate	0.44	15	68
Excess M1 balances ^d	0.60	9	67
Domestic credit/GDP growth rate	0.78	4	65
Real interest rate	0.76	4	38
M2 multiplier growth rate	1.14	-2	68
Import growth rate	1.16	-2	68
Industrial production growth rate	1.14	-2	54
Terms of trade growth rate	0.93	1	55
Lending rate/deposit rate	1.04	-1	29
Bank deposit growth rate	1.63	-6	68
Stock price index growth rate	1.59	-6	44
Real interest differential	1.34	-4	38
Current account/GDP	0.42	16	70
M2/reserves (level)	0.45	14	51

^aRatio of false signals (measured as a proportion of months in which false signals could have been issued $[B/(B+D)]$) to good signals (measured as a proportion of months in which good signals could have been issued $[A/(A+C)]$).

^b $P(\text{crisis/signal})$ is the percentage of the signals issued by the indicator that were followed by at least one crisis within the subsequent 24 months $[A/(A+C)]$ in terms of the matrix in the text). $P(\text{crisis})$ is the unconditional probability of a crisis, $(A+C)/(A+B+C+D)$.

^cDeviation from deterministic trend.

^dResidual from regression of real M1 on real GDP, inflation, and a deterministic trend.

We find eight indicators to be informative: deviations of the real exchange rate from trend, the growth in M2 as a fraction of reserves, export growth, change in international reserves, “excess” M1 balances, growth in domestic credit as a share of GDP, the real interest rate, and the growth in the terms of trade.¹⁴

Predicting 1997

We have already calculated the optimal thresholds and resulting noise-to-signal ratios for the different indicators. To forecast for the post-April 1995 period, we

¹⁴These indicators are also all informative in the KLR analysis. These results are quite similar to those obtained by KLR with a different sample of countries and time period, though they found a further four indicators to be informative. See the Appendix for more detail, as well as a full analysis of in-sample performance.

apply these thresholds to the values of the predictive variables after this date, determining whether they are issuing signals or not. The first column of Table 2 shows the performance of the Kaminsky (1998a) composite measures of the probability of crisis based on the weighted sum of indicators signaling.

A natural question is whether the estimated probability of crisis is above 50 percent prior to actual crises. The summary statistics rows show that only 4 percent of the time was the predicted probability of crisis above 50 percent in cases when there was a crisis within the next 24 months, during the period May 1995 to December 1997. If we are more interested in predicting crises than predicting tranquil periods and are not so worried about calling too many crises, we may want to consider an alarm to be issued when the estimated probability of crisis is above 25 percent. Table 2 shows that the estimated probabilities are above 25 percent in 25 percent of the precrisis observations. Sixty-three percent of alarms, however, are false at the 25 percent cutoff.

This is not very good performance: most crises are missed and most alarms are false. These forecasts are, nonetheless, better than random guesses, both economically and statistically. The actual out-of-sample frequency of crisis following an alarm (defined as an estimated probability above 25 percent) is 37 percent. The frequency of crisis following periods without such alarms is 24 percent. And a χ^2 test of the goodness of fit results rejects at the 5 percent level of significance the hypothesis that the number of successfully called crises is no higher than if the warnings were uninformative.¹⁵

So far we have examined the ability of the model to predict the approximate *timing* of crises for each country.¹⁶ We can also evaluate the cross-sectional success of the models' predictions in identifying which countries are vulnerable in a period of global financial turmoil such as 1997. The question here is whether the models assign higher predicted probabilities of crisis to those countries that had the biggest crises. We can then evaluate forecast performance by comparing rankings of countries based on the predicted and actual crisis indices. As we will see, this also allows us to compare forecasts across models with different definitions of crisis. Table 3 shows countries' actual crisis index and predicted probability of crisis in 1997 for the various different forecasting methods.¹⁷ The table also shows the Spearman correlation between the actual and predicted rankings and its associated *p*-value, as well as the R^2 from a bivariate regression of the actual rankings on the predictions.

The KLR-based forecasts are somewhat successful at ranking countries by severity of crisis. The forecasted probabilities are significantly correlated with the actual rankings of countries in 1997 by their crisis index. They explain 28 percent of the variance.

To get a richer sense of how useful this general approach would have been, we now examine more closely the predictions of the KLR-based model for four Asian

¹⁵This is true for both the 50 percent and the 25 percent cutoff.

¹⁶We say approximate because the model only attempts to place the crisis within a 24-month window.

¹⁷The predicted crisis probability is the average of the probabilities during January to December 1996, using the out-of-sample estimates. The actual crisis index used to rank the countries for 1997 is the maximum value of the monthly crisis index for each country during 1997.

Table 2. Goodness-of-Fit of KLR Model—Out of Sample

Cutoff of 50 Percent
Goodness-of-Fit Table^a

Predicted	Original specification			Predicted	Augmented with current account and M2/reserves		
	Actual				Actual		
	Tranquil	Crash	Total		Tranquil	Crash	Total
Tranquil	337	117	454	Tranquil	338	122	460
Crash	1	5	6	Crash	0	0	0
Total	338	122	460	Total	338	122	460

Summary Statistics

	Original	Augmented
<i>p</i> -value for χ^2 test of independence	0.002	No crisis called
Percent of observations correctly called	74	73
Percent of crises correctly called ^b	4	0
Percent of tranquil periods correctly called ^c	100	100
False alarms as a percent of total alarms ^d	17	No crisis called
Probability of crisis given:		
an alarm ^e	83	No crisis called
no alarm ^f	26	27

Cutoff of 25 Percent
Goodness-of-Fit Table^a

Predicted	Original specification			Predicted	Augmented with current account and M2/reserves		
	Actual				Actual		
	Tranquil	Crash	Total		Tranquil	Crash	Total
Tranquil	286	91	377	Tranquil	279	83	362
Crash	52	31	83	Crash	59	39	98
Total	338	122	460	Total	338	122	460

Summary Statistics

	Original	Augmented
<i>p</i> -value for χ^2 test of independence	0.014	0.001
Percent of observations correctly called	69	69
Percent of crises correctly called ^b	25	32
Percent of tranquil periods correctly called ^c	85	83
False alarms as a percent of total alarms ^d	63	60
Probability of crisis given:		
an alarm ^e	37	40
no alarm ^f	24	23

^aTable shows number of observations.

^bA precrisis period is correctly called when the estimated probability of crisis is above the cutoff probability and a crisis ensues within 24 months.

^cA tranquil period is correctly called when the estimated probability of crisis is below the cutoff probability and no crisis ensues within 24 months.

^dA false alarm is an observation with an estimated probability of crisis above the cutoff (an alarm) not followed by a crisis within 24 months.

^eThis is the number of precrisis periods correctly called as a share of total predicted precrisis periods.

^fThis is the number of periods where tranquility is predicted and a crisis actually ensues as a share of total predicted tranquil periods (observations for which the predicted probability of crisis is below the cutoff).

Table 3. Correlation of Actual and Predicted Rankings Based on KLR, FR, and STV

	KLR		FR		STV				
	Predicted probabilities of crisis in 1997		Predicted probabilities of crisis in 1997 ^d		Predicted probabilities of crisis in 1997				
	Actual 1997 crises index	Noise-to-signal weighted sum of indicators ^a Original ^b Augmented ^c	Actual 1997 crisis index	Table 4 Model 1	Table 4 Model 2	Table 5 Model 3	Table 5 Model 4		
Thailand	1	16	7	3	7	11	2	7	5
Korea	2	4	5	2	7	7	3	12	11
Indonesia	3	18	11	2	7	7	1	14	9
Malaysia	4	8	13				4	6	6
Zimbabwe	5	3	3				5	23	12
Taiwan Province of China	6	5	4				9	11	22
Colombia	7	9	12	8	8	6	8	18	4
Philippines	8	1	1	7	8	8	6	1	1
Brazil	9	2	2	10	6	5	14	4	2
Turkey	10	7	10	1	3	2	7	9	21
Venezuela	11	14	16	5	10	12	21	22	13
Pakistan	12	10	9	6	11	9	10	17	20
South Africa	13	6	8				12	15	16
Jordan	14	15	18				17	20	15
India	15	20	21	14	13		13	5	19
Sri Lanka	16	17	19	11	14	13	15	16	17
Chile	17	18	20	15	9	10	16	19	14
Bolivia	18	20	21	13	12	12	22	13	10
Argentina	19	12	17	16	5	3	23	2	7
Mexico	20	13	14	12	4		18	21	18
Peru	21	11	6	9	1	4	20	8	23
Uruguay	22	20	21	4	2	1	11	3	3
Israel	23	20	15				19	10	8
Correlation ^e		0.54	0.60		0.33	0.12		0.11	0.23
p-value		0.007	0.003		0.253	0.694		0.612	0.295
R ²		0.28	0.36		0.11	0.02		0.01	0.05

^aBased on average of weighted sample conditional probabilities during 1996, using out-of-sample estimates.

^bOriginal KLR variables.

^cAddition of current account and M2/reserves in levels to original variables.

^dAverage predicted probabilities for 1996, where model was estimated up to April 1995.

^eSpearman Rank Correlation of the fitted values and the actual crisis index and its p-value. The R² is from a regression of fitted values on actual values.

crisis countries (where crisis is identified according to the KLR definition): Korea, Indonesia, Malaysia, and Thailand, and one Asian and three Latin American non-crisis countries: Philippines, Argentina, Brazil, and Mexico.¹⁸ Figure 1 presents the KLR composite measure of estimated probability of crisis, with vertical lines at crisis dates.

The KLR probability forecasts do not paint a clear picture of substantial risks in crisis compared to noncrisis countries. Two (then) noncrisis countries, Brazil and the Philippines, consistently present risks of crisis above 30 percent during 1996. One crisis country, Korea, also presents risks above 30 percent, though only in the first half of the year, while Malaysia is generally above 20 percent. Estimated crisis risks remain below 17 percent in 1996 for the crisis and noncrisis countries Argentina, Mexico, Indonesia, and Thailand.

In sum, the KLR is a mixed success. The fitted probabilities from the weighted sum of indicators are statistically significant predictors of crisis probability in 1997. Still, the overall explanatory power is fairly low, as demonstrated by the low R^2 statistic in the regression of the actual on the predicted crisis rankings and the overall goodness of fit for the out-of-sample predictions.

Frankel and Rose (1996) Probit Model

The Model

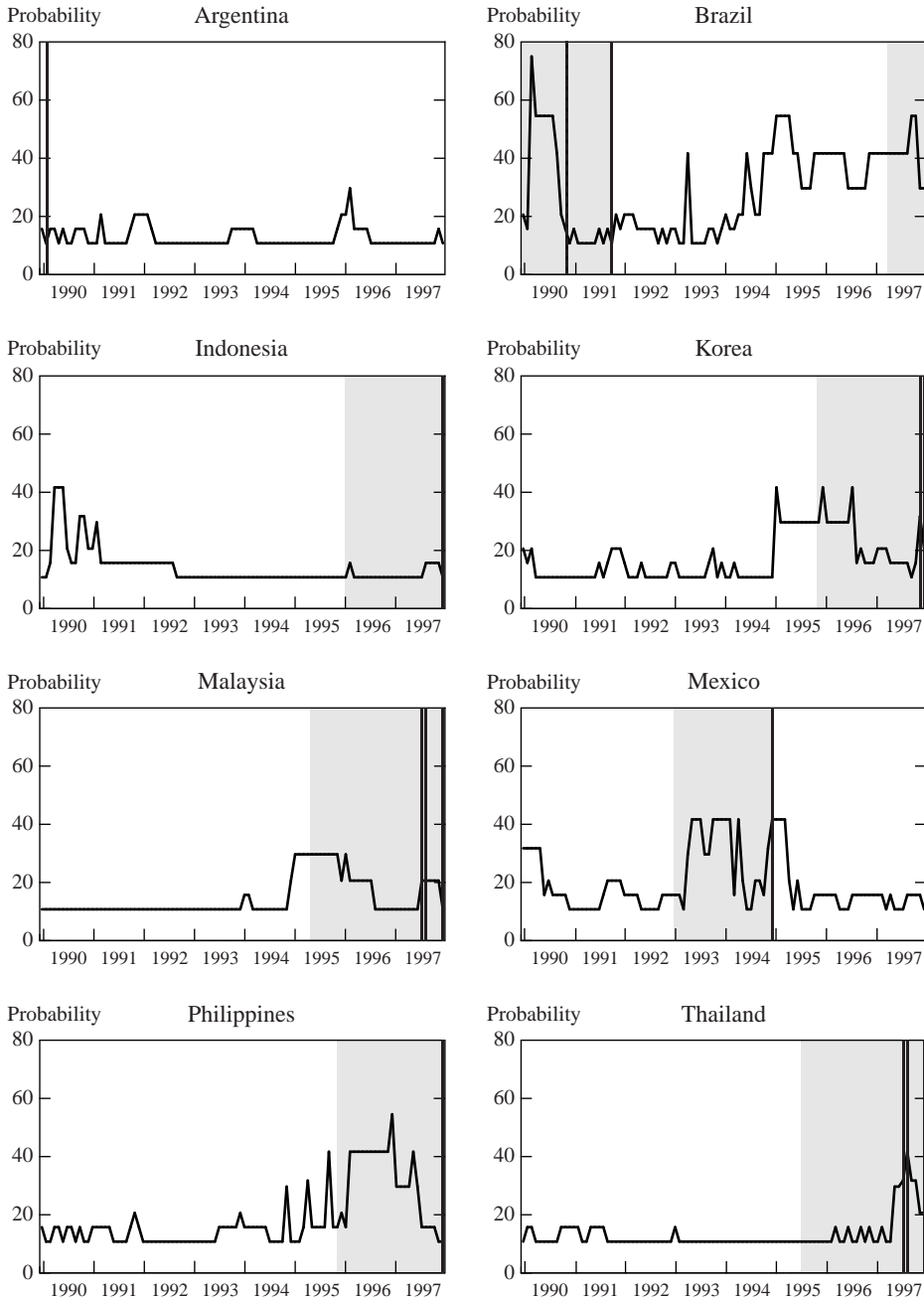
FR estimate the probability of a currency crash using annual data for more than 100 developing countries from 1971–92, a much broader sample of countries than the other two papers. The use of annual data may restrict the applicability of the approach as an early warning system, but it permits the analysis of variables such as the composition of external debt for which higher frequency data are rarely available. FR test the hypothesis that certain characteristics of capital inflows are positively associated with the occurrence of currency crashes: low shares of FDI; low shares of concessional debt or debt from multilateral development banks; and high shares of public-sector, variable-rate, short-term, and commercial bank debt.¹⁹

¹⁸These countries are an interesting but nonrandom subsample. We use them only to illustrate the conclusions from the broader sample.

¹⁹The complete list of variables is as follows. Domestic macroeconomic variables: (1) the rate of growth of domestic credit, (2) the government budget as percent of GDP, (3) and the growth rate of real GNP. Measures of vulnerability to external shocks include: (1) the ratio of total debt to GNP, (2) the ratio of reserves to imports, (3) the current account as a percentage of GDP, and (4) the degree of overvaluation, defined as the deviation from the average bilateral real exchange over the period. Foreign variables are represented by (1) the percentage growth rate of real OECD output (in U.S. dollars at 1990 exchange rates and prices), and (2) a “foreign interest rate” constructed as the weighted average of short-term interest rates for the United States, Germany, Japan, France, the United Kingdom, and Switzerland, with weights proportional to the fractions of debt denominated in the relevant currencies. Characteristics of the composition of capital inflows are expressed as a percentage of the total stock of external debt and include (1) amount of debt lent by commercial banks, (2) amount that is concessional, (3) amount that is variable rate, (4) amount that is public sector, (5) amount that is short-term, (6) amount lent by multilateral development banks (includes the World Bank and regional development banks but not the International Monetary Fund), and (7) the flow of FDI as a percentage of the debt stock.

ARE CURRENCY CRISES PREDICTABLE?

Figure 1. KLR Crisis Probabilities for Selected Countries



Note: The solid vertical lines represent crisis dates. The areas with dashed lines denote the 24 months prior to crises.

FR define a currency crash as a nominal exchange rate depreciation of at least 25 percent that also exceeds the previous year's change in the exchange rate by at least 10 percent. Thus, the type of currency crisis considered does not include speculative attacks successfully warded off by the authorities through reserve sales or interest rate increases. FR argue that it is more difficult to identify successful defenses, since reserve movements are noisy measures of exchange market intervention and interest rates were controlled for long periods in most of the countries in the sample.

Table 4 (column 1) presents the FR benchmark probit regression, estimated from 1970 through 1996 for purposes of forecasting 1997. The coefficients reflect the effect of one-unit changes in regressors on the probability of a currency crash (expressed in percentage points) evaluated at the mean of the data.²⁰ We can conclude that the probability of a crisis increases when foreign interest rates are high, domestic credit growth is high, the real exchange rate is overvalued relative to the average level for the country, the current account deficit and the fiscal surplus are large as a share of GDP, external concessional debt is small, and FDI is small relative to the total stock of external debt.²¹ As noted in the Appendix, the in-sample goodness of fit of the FR model is reasonably high.

Predicting 1997

The FR model estimated through 1996 can easily generate out-of-sample predictions for 1997. We cannot directly analyze goodness of fit for this model, as there were no crisis countries in 1997 according to the FR definition.²² Instead, we can compare the predicted probabilities of crisis and actual values of nominal exchange rate depreciation for 1997 for predictions based on model 1 of Table 4 (Table 3). Overall, the forecasts are not successful, with a correlation of 33 percent. The fraction of the variance of the rankings accounted for (measured by the R^2) is 11 percent, and the prediction is not significant.²³ In sum, the FR model fails to provide much useful guidance on crisis probabilities in 1997.

²⁰Thus, an increase in the degree of exchange rate overvaluation by 1 percentage point would increase the estimated probability of crisis by 0.172 percentage points.

²¹This contrasts somewhat from the published FR results, particularly in the significance of the current account and the real exchange rate and the insignificance of reserves/imports. These changes result from several differences in specification. In addition to the inclusion of more recent years, the most important changes were that we exclude countries with a population below 1 million or annual per capita GDP below \$1,000 and that we have fixed an error that resulted in a miscalculated real exchange rate measure. See the Appendix for details.

²²This reflects the fact that the use of annual frequency does not work well here; because the devaluations happened toward the end of the year, none of the Asian countries are identified as crisis countries in 1997.

²³This correlation is based on the 13 countries for which data are available that are part of the 23-country common sample. Based on the full sample where data are available (25 out of the 41 countries included in model 3A of Appendix Table A3), the forecasts are even less successful.

ARE CURRENCY CRISES PREDICTABLE?

Table 4. Frankel and Rose: Probit Estimates of Probability of a Currency Crash, 1970–96

	Model 1 FR specification		Model 2 Modified	
	dF/dx	$ z ^a$	dF/dx	$ z ^a$
Commercial bank share of total debt	0.022	0.1	0.121	1.0
Concessional share	-0.296	-2.3 **	-0.305	-3.3 ***
Variable rate share	0.020	0.1	-0.089	-0.6
Short-term share	0.106	0.6	0.161	1.1
FDI/debt	-0.795	-2.5 **	-0.576	-1.8 *
Public sector share	0.212	1.7 *	0.247	2.4 **
Multilateral share	0.021	0.1	0.045	0.4
Debt/GNP	-0.025	-0.5	0.021	0.6
Reserves/imports	-0.007	-1.3		
Reserves/M2			-0.206	-3.5 ***
Current account/GDP	-0.697	-2.4 **	-0.679	-2.8 ***
Overvaluation ^b	0.172	2.9 ***	0.107	2.4 **
Government budget surplus/GDP	0.767	2.6 ***	0.595	2.6 **
Domestic credit growth	0.182	4.4 ***	0.119	3.1 ***
GDP growth rate	-0.058	-0.9	-0.017	-0.3
Foreign interest rate	1.007	2.0 **	0.909	2.3 **
Northern (OECD) growth	0.414	0.4	0.033	0.0
Open			-0.239	-4.1 ***
Sample size	464		448	
Pseudo R^2	0.20		0.32	

Goodness of Fit

	Model 1			Model 2 Actual		
	Tranquil	Crash	Total	Tranquil	Crash	Total
Cutoff probability of 50 percent ^c						
Predicted tranquility	398	50	448	381	40	421
Predicted crash	6	10	16	7	20	27
Total	404	60	464	388	60	448
Cutoff probability of 25 percent ^d						
Predicted tranquility	373	34	369	342	22	369
Predicted crash	31	26	57	41	38	79
Total	404	60	426	383	60	448

^aOne, two, and three asterisks denote significance at the 10, 5, and 1 percent levels, respectively.

^bDefined as the deviation from the average real exchange rate over the period.

^cA crisis is correctly called when the estimated probability of crisis is above 50 percent if a crisis ensues within 24 months. A tranquil period is correctly called when the estimated probability of crisis is below 50 percent and there is no crisis within 24 months.

^dA crisis is correctly called when the estimated probability of crisis is above 25 percent if a crisis ensues within 24 months. A tranquil period is correctly called when the estimated probability of crisis is below 25 percent and there is no crisis within 24 months.

Sachs, Tornell, and Velasco (1996) Cross-Country Regressions

The Model

STV analyze the impact of Mexico's financial crisis of December 1994 on other emerging markets in 1995. They examine the determinants of the magnitude of the currency crisis in a cross section of 20 countries in 1995. This approach cannot hope to shed light on the timing of crises. Rather, it may answer the question of which countries are most likely to suffer serious attacks in the event of a change in the global environment. This approach is potentially attractive, even for our purposes, for a number of reasons. First, the timing may be much harder to predict than the incidence of a crisis across countries. Moreover, the determinants of crisis episodes may have varied importantly over time. STV can impose more economic structure on their analysis by focusing on a particular set of crises (those occurring at one time). STV argue that a key feature of the 1995 crises was that the attacks hit hard only at already vulnerable countries. In a rational panic, investors identify a country as being likely to suffer from a large devaluation in the face of an outflow, and validate their own concerns by fleeing the country. Thus, countries with overvalued exchange rates and weak banking systems were subject to more severe attacks, but only if they had low reserves relative to monetary liabilities (so that they could not easily accommodate the capital outflow) and weak fundamentals (so that fighting the attack with higher interest rates would be too costly).

The original STV model was not designed to predict future crises but rather to explain events in 1995. For our purposes, it is important for the crises that affected mostly Asian countries in 1997 to have been broadly similar to the 1995 crises. And in fact a number of researchers have argued since 1997 that the two sets of crises share many characteristics. Radelet and Sachs (1998a) argue that the 1997 and 1995 crises shared important characteristics, though their interpretation of post-Thailand Asian crises relies more heavily on contagion effects. The IMF (1998) argues that the STV results apply to the Asian crisis and constructs a composite indicator of crises on that basis. Radelet and Sachs (1998b), Tornell (1998), and Corsetti, Pesenti, and Roubini (1998a) also apply models in the STV spirit to both sets of crises.

Tequila Crisis Models

STV define a crisis index (*IND*) as the weighted sum of the percent decrease in reserves and the percent depreciation of the exchange rate, from November 1994 to April 1995. They argue that countries had more severe attacks when their banking systems were weak (proxied by a lending boom variable (*LB*) measuring growth in credit to the private sector from 1990 through 1994) and when the exchange rate was overvalued (measured as the degree of depreciation from 1986–89 to 1990–94 (*RER*)). Moreover, they find that these factors only matter for countries with low reserves (*DLR*), measured as having a reserves/M2 ratio in the lowest quartile, and “weak fundamentals” (*DWF*), which means having *RER* in the lowest three quartiles or *LB* in the highest three quartiles.

Thus, they estimate across the i countries in their sample an equation of the form:

$$IND_i = \beta_1 + \beta_2 RER_i + \beta_3 LB_i + \beta_4 RER_i \cdot DLR_i + \beta_5 LB_i \cdot DLR_i + \beta_6 RER_i \cdot DWF_i + \beta_7 LB_i \cdot DWF_i + \varepsilon_i.$$

Regression 1 of Table 5 reproduces the original STV benchmark regression, using their data.²⁴ The results emphasized by STV are, first, that the effect of *RER* is significantly negative for countries with low reserves and weak fundamentals (the sum of estimates of $\beta_2 + \beta_4 + \beta_6$ is negative), and the effect of *LB* is significantly positive for these same countries (the sum of estimates of $\beta_3 + \beta_5 + \beta_7$ is positive). They take the high R^2 of the regression (0.69) to indicate that the model explains the pattern of contagion well.

To apply this model to the 1997 crises, we run the model over the original STV sample (row 2 of Table 5) as well as the same sample of 23 countries to which we apply the KLR approach (row 3). The regression coefficients change substantially. The STV hypotheses now receive only mixed support. For example, when revised data are used (row 2), the effect of *RER* with low reserves and weak fundamentals ($\beta_2 + \beta_4 + \beta_6$) is now insignificantly different from zero, while the coefficient on *LB* with low reserves ($\beta_3 + \beta_5$) increases significantly.

The fragility of the STV results with respect to the data revisions that have taken place since their estimations and to the addition of three countries to the sample casts some doubt on the usefulness of this specification for the Asian crises. We nonetheless generate predictions for 1997 based on these estimates drawn from the Tequila crisis.

Predicting 1997

To implement the STV model for 1997, we mechanically update the STV variables and apply the coefficients from the STV regressions for the Tequila crisis to obtain predicted values for the 1997 crises. For the dependent variable that measures the severity of the crisis, we measure percent depreciation of the nominal exchange rate from April 1997 through December 1997. For the explanatory variables, we move all the definitions forward two years. We then calculate forecasts of devaluation using the coefficient estimates from the STV benchmark specification estimated for the Tequila crisis.

Column 7 of Table 3 shows the country rankings based on the actual value of the crisis index for 1997, defined, analogously to STV, as the change in the nominal exchange rate between April and December 1997. Column 8 presents country rankings based on applying the coefficients from the STV regression estimated over the 23-country sample to the updated *LB* and *RER* variables and associated dummy variables.

STV themselves try many variants of their benchmark regression, in their case to demonstrate robustness. For example, the STV definition in terms of the average level of the real exchange rate in the 1990 through 1994 period divided by the

²⁴Regression 1 differs slightly from the published benchmark regression, as discussed in the Appendix.

Table 5. STV: 1994/95 Regressions

Regression Number	Regression	Number of Countries	R^2	\bar{R}^2	Results ^{a,b}						
					STV hypotheses: $\beta_2 = 0$	$\beta_{2+\beta_4} = 0$	$\beta_{2+\beta_4+\beta_6} < 0$	$\beta_3 = 0$	$\beta_{2+\beta_5} = 0$	$\beta_{2+\beta_5+\beta_7} > 0$	
1	STV benchmark, rerun fixing Taiwan	20	0.69	0.55	6.61 (2.311)	0.28 (0.539)	-2.16 (0.994)	1.86 (0.810)	-5.49 (2.648)	3.88 (1.463)	
2	STV with revised data	20	0.66	0.50	6.46 (1.860)	-0.06 (0.228)	-1.59 (1.133)	1.74 (0.654)	-16.87 (7.292)	3.75 (1.405)	
3	23-country sample	23	0.44	0.22	4.18 (1.811)	-0.34 (0.205)	-3.48 (1.615)	0.94 (0.617)	-8.11 (6.578)	1.89 (1.354)	
4	Alternate <i>RER</i> definition (1) <i>RER</i> 1994/90	23	0.41	0.19	-4.18 (2.329)	0.78 (0.875)	-4.29 (1.784)	-0.42 (0.157)	0.22 (0.771)	2.15 (1.582)	

^aCoefficients in bold are significant at the 5-percent level. Bolded coefficients are significantly inconsistent with the STV hypothesis. Figures in parentheses are standard errors.

^bThe β s are coefficients from the regression $IND = \beta_2 RER + \beta_3 LB + \beta_4 RER \cdot DLR + \beta_5 RER \cdot DWF + \beta_6 RER \cdot DWF + \beta_7 LB \cdot DWF$, where *RER* is the degree of real depreciation, *LB* is a measure of the lending boom, *DLR* is a dummy variable for countries with low reserves, and *DWF* is a dummy for countries with weak fundamentals (see text for explanations).

average level during 1986 through 1989 clearly has an arbitrary element, and they also try other measures, such as the percent change in the real exchange rate from 1990 to 1994.

None of these forecasts performs well. The most successful specification, based on Table 5, regression 4, employs one of the alternative definitions of *RER*. Its forecast rankings of crisis severity are insignificant predictors of the actual rankings and explain only 5 percent of the variance of the actual country rankings.²⁵

A recent paper (Tornell, 1998) may seem to contradict the results in this paper. Tornell estimates a model very similar to STV, stacking observations from the 1994/95 crisis and the 1997 crisis. He finds that his new model: (1) fits fairly well, with significant coefficients plausibly signed; (2) has coefficients that appear stable between the two sets of crises; and (3) when fitted with the 1994 observations only and forecasting for 1997, produces good predictions, much better than the STV forecasts examined here and comparable to the KLR-weighted sum of indicators-based probabilities.

Rather than providing a counterexample to the results presented here, this effort illustrates the importance of testing models out of the sample used to formulate them, as we do here. A variety of apparently small modifications characterizes the difference between the specification in STV and Tornell (1998), and yet these respecifications apparently make the difference between success and failure in predicting the incidence of the 1997 crises “out of sample.”²⁶

This suggests that specification uncertainty can be as important as parameter uncertainty across crisis episodes, at least for techniques such as STV that rely on a small number of observations and relatively complex models. Only the application of models to episodes that postdate the design of the model provides an appropriately tough test. Unfortunately for our purposes, the apparent need for a separate specification search for the new set of crises casts some doubt on the usefulness of this sort of approach for predicting future crises.

II. Do Additional Variables Help?

We have seen that even the most successful of the models under consideration (KLR) has fairly low explanatory power. None of these papers was meant to be the last word on forecasting, however, so it is reasonable to ask whether it would have been possible to do better with some relatively minor modifications. We have already corrected some errors in the previous versions, as would anyone

²⁵In light of this predictive failure, we have also considered a much less ambitious test of the STV model, justified by the idea that we may reasonably expect some constancy of the general model of crisis episodes even if parameter constancy fails to hold. It turns out, however, that even when reestimated using 1996 and 1997 data to explain the 1997 results, the STV model applied to the 1997 crisis meets with little success. The results vary strongly depending on the exact specification, but the fit is always poor. Compared with its application to the 1994 crisis, the coefficients are economically and statistically different, and the explanatory power of the regressions is much lower. Naturally, the in-sample results for 1997 are superior to the out-of-sample predictions we have already analyzed. It is remarkable, though, that the STV regression reestimated with 1997 data performs somewhat worse than the KLR out-of-sample forecasts.

²⁶Bussière and Mulder (1999) confirm this conclusion. They find that the Tornell (1998) model performs poorly at predicting 1998 crises.

implementing them in early 1997. We have also looked at robustness to alternate samples and, in the case of STV, to changes in the definition of some of the explanatory variables. Here, though, we go one step further and ask whether the addition of some plausible right-hand-side variables would have greatly improved the performance of the models. To some extent we are, then, deviating here from the approach of testing “pure” out-of-sample forecasts.

KLR omitted several variables that even prior to 1997 were clearly identified in the literature as important potential determinants of crisis, most notably the level of the ratio of M2 to reserves and the ratio of the current account to GDP. KLR used the rate of growth of M2/reserves, but most discussions of crisis vulnerability even then focused on the level of this variable. KLR did not use the current account. We find that in the KLR framework both the level of M2/reserves and the ratio of the current deficit to GDP are highly informative over the in-sample period, as Table 1 shows.²⁷ As shown in the second column of Table 2, the KLR model augmented with these two additional variables performs noticeably better out-of-sample than the original model. For example, 32 percent of the precrisis observations are called correctly at the 25 percent cut-off, compared with 25 in the original model. In the rank correlation test, the augmented model’s predictions are more highly correlated with the actual ranking of crises, with a correlation coefficient of 0.60 compared with 0.54 for the original model (columns 2 and 3 of Table 3).

For the FR model, we also tried alternative explanatory variables, all estimated using data through 1996. We saw in the original FR specification that the ratio of reserves to imports does not seem to matter. Measuring reserves as a ratio to short-term external debt and to broad money (M2) have both been suggested as alternative ways of measuring the adequacy of reserves.²⁸ We find that both the ratio of reserves to short-term external debt and that of reserves to M2 are separately significant predictors of crisis. When all three reserve ratios are included, the ratio of reserves to M2 is significant at the 1 percent level, while the ratio of reserves to short-term external debt is significant at the 10 percent level. The ratio of reserves to imports is insignificant and wrongly signed. The degree of openness of the economy may indicate the flexibility of the adjustment mechanism in the country and hence the probability of crisis. We find that more open economies, as measured by the share of exports and imports in GDP, were significantly less likely to suffer a crisis.²⁹ Changes in the terms of trade had no apparent impact on the likelihood of crisis, while measuring the debt composition variables as a share of GDP rather than total debt also had no effect. Interacting short-term external debt with credit growth, in the spirit of STV, also did not help predict crises.

²⁷The current account is measured as a moving average of the previous four quarters. We use our interpolated monthly GDP series to form the ratio of the current account to the moving average of GDP over the same period.

²⁸See Calvo and Mendoza (1996) on Mexico for an emphasis on the ratio of M2 to reserves and Radelet and Sachs (1998a) on the Asian crises for a focus on short-term external debt/reserves. The inclusion of the ratio of reserves to short-term external debt is particularly in violation of the out-of-sample spirit of this paper, as most of the interest in this variable postdates the Asian crises.

²⁹Milesi-Ferretti and Razin (1998) make this argument and include this variable in a similar regression with some success.

As a result of this specification search, regression 2 of Table 4 includes the ratio of reserves to M2 and the degree of openness of the economy. These additions do not help performance in 1997, as shown in column 6 of Table 3, which shows that the correlation of predicted and actual rankings of crises in 1997 is still small and insignificantly different from zero.

We did not attempt to add variables to the STV model, partly because the small sample size renders the exercise particularly prone to data mining and also because STV themselves consider and reject the main alternative candidate explanatory variables. We noted above that we have investigated a variety of different specifications suggested by STV themselves, without success.

III. Is It Fair to Compare Such Different Models?

We have judged these models based on their forecasting performance. Only the KLR model was designed explicitly with this objective in mind, and so it is perhaps not surprising that it is the most successful. However, FR is also a panel-based approach, and it is a reasonable test of the model to ask how well it fits in more recent years. And the value of the STV model depends in part on its applicability to crises in general, not just to those over which it was estimated.³⁰

We have analyzed and compared results from three models that differ in critical ways. Most fundamentally, they are models with different crisis definitions—that is, dependent variables—and different samples. Since each model is forecasting something different, the comparison of typical statistics such as the R^2 is not helpful. We have therefore relied on goodness of fit, where applicable, and more generally on the rank correlation of predicted probabilities and actual incidence of crisis in 1997 in assessing the models.³¹

It is nonetheless important to keep in mind that success has different meanings for each of the models. For STV, it would imply that the relative severity of crisis was predictable, given the time period during which attacks might be expected to occur. KLR (and even more so FR, because of the shorter forecast interval) attempt as well the more ambitious task of predicting the timing of crises. It is perhaps surprising that KLR achieves some success at both ranking (as measured by the correlations of predicted and actual for 1997) and timing, as measured by the goodness-of-fit statistics.

The three models embody different definitions of crisis. STV and KLR agree on looking at a crisis index that combines information on reserve losses and exchange rate depreciations, on the grounds that they are trying to measure pressure on the exchange regime, whether it results in a devaluation or not. FR measure only the exchange rate, though largely on the practical grounds that data on reserve changes are noisy. FR and KLR choose to look for discrete crises defined as extreme values of the underlying index. This approach may be justified on the

³⁰Many have tried to apply the model to other crises, as mentioned in footnote 9.

³¹For the same reason, it is also not helpful to directly compare probabilities of crisis across models. Where the crisis events are more common, the unconditional probabilities, and hence the mean forecast probabilities in an unbiased model, are higher.

grounds that crises represent a structural break in the behavior of the exchange rate and reserves compared to other times; the models are attempting, then, to predict the breaks, not the behavior in between. STV do not predict crises as discrete events; rather, they try to predict the severity of crises as measured by the percent change in a crisis index over a particular period.

Different crisis definitions yield different results, and all operational definitions of crisis contain measurement error in that they only imperfectly capture whatever we have in mind by currency crises.³² This may worsen the performance of the models, though it may mean that they “really” work better than reported, in that some of the false alarms or missed crises may have been due to measurement error of the dependent variable. We have not explored this issue here.

The models in their original forms were estimated over quite different samples: FR used the broadest possible sample of developing countries over 22 years; STV estimated over only a cross section of “emerging markets”—that is, countries in the IFC database—at a particular time characterized by contagion and crisis; and KLR included an eclectic mix of developing and developed countries, the latter in particular chosen partly because they had crises, over 25 years. We have to some extent tested whether these differences in sample were important, by reestimating the models over the original and over a common sample. We have found that the KLR results were fairly robust to this change, though we find fewer indicators to be informative. The FR specification changed in some important ways with the restriction of the sample.³³ The STV results turned out to be most fragile to changes in sample, both over the original time period and also with respect to future crisis episodes. It turns out, though, that this variation in performance of the STV and FR models across samples did not matter along one important dimension: in no case did the out-of-sample forecasts predict crises well.

The models forecast over different time horizons. FR and STV forecast roughly one year out, while KLR considers an alarm to be correct if a crisis happens any time within a 24-month window. This difference is not responsible for the superior performance of the KLR model, as it performs about as well when attempting to forecast crises 12 months ahead rather than 24.

Furman and Stiglitz (1998) apply the KLR methodology to predicting the Asian crisis and, while they do not systematically evaluate the results, conclude that it does not work well, noting some success but many false positives. They dismiss what success they do observe largely on the argument that the method of measuring predictive variables in terms of percentiles is biased in favor of predicting crises in countries that have previously had little volatility in predictive variables. For example, even relatively small real exchange rate appreciation results in a large percentile deviation in historically tranquil countries, such as the Asian crisis countries. We find this argument unconvincing. There are many reasons why measures that compare variables to their own history may pick up important trends efficiently.³⁴

³²See Milesi-Ferretti and Razin (1998) on sensitivity to alternative crisis definitions in the FR model.

³³See the Appendix for details.

³⁴A doctor may well ask whether a patient has lost weight, not how his weight compares to the standard charts, when looking for signs of sickness. (We thank Joseph Stiglitz for this analogy.)

Ultimately, the question is empirical. In fact, the KLR model does not tend to systematically overpredict crises in-sample in relatively tranquil countries.

The models analyzed in this paper are, with the partial exception of STV, reduced form and nonstructural. An alternative approach is to estimate a well-defined structural model. Blanco and Garber (1986) estimate a model of currency crisis probability for Mexico that achieves some success. The results of this sort of model are hard to compare with those we consider here. First, their results are essentially a special case, in that they fit a specific structural model. The first-generation model they estimate, with excess domestic credit creation driving a crisis, is more plausibly applied to the specific crises they consider (Mexico's in the 1970s and 1980s) than in many other cases. Their estimation depends on using the interest rate differential as a measure of expected devaluation. The empirical relevance of this assumption is doubtful, despite its plausibility.³⁵ Moreover, they estimate only one period ahead, a horizon that may be of limited use for policymakers.³⁶

IV. Conclusion

We have examined the extent to which models formulated and estimated prior to 1997 would have helped predict the 1997 currency crises. The exercise is thus "out of sample" both in the sense that we estimate the models using data only through 1996 and, equally important, in that the models themselves were specified prior to 1997. The results of this unusually tough test are generally though not unambiguously negative. Two of the three models (STV and FR) provide forecasts that are no better than guesswork. Ex ante plausible variations in sample and specification did not change this result.

The KLR model, in contrast, achieved a measure of success. The probabilities of crisis it generated during the period May 1995 to December 1996 were statistically significant predictors of actual crisis incidence over the subsequent 24 months. Moreover, its forecasted cross-country ranking of severity of crisis is a significant predictor of the actual ranking. This success should not be exaggerated. The model does not explain a large part of the actual variation in outcomes. When this model issued an alarm during the May 1995 to December 1996 period, a crisis would actually have followed in 1997 37 percent of the time.³⁷ This compares with a 27 percent unconditional probability of crisis in 1997. And the model explains only 28 percent of the variation in actual crisis rankings.

We also tried adding various explanatory variables to the models. Plausible modifications to the STV and FR models did not yield useful forecasts, even some, such as the inclusion of short-term external debt, actually inspired by events in 1997. The addition of two variables to the KLR model that were widely

³⁵The interest differential did not signal an expected exchange rate change in advance of Mexico's 1994 crisis, as Werner (1996) discusses.

³⁶Their estimated probabilities of crisis are generally somewhat lower than those of KLR largely because they are trying to predict a much rarer event than KLR (a crisis next month, as opposed to a crisis sometime within the next 24 months).

³⁷An alarm here is defined as a predicted probability above 25 percent.

considered good indicators prior to 1997—the level of the current account balance and M2/reserves—improves performance somewhat.

The answer to the question posed in the title of this paper is thus “yes, but not very well.” The answer is “yes” since the KLR forecasts, and even more so the modified model, are clearly better than a naive benchmark of pure guesswork. We say “not very well” because even the KLR model issues more false alarms than accurate warnings, while it misses most crises.

We have judged the forecasts of these models against a naive alternative of pure guesswork, and the statistically significant results do not imply that the KLR model does better than the analysis of informed observers. Systematic comparisons against alternative benchmarks would be interesting. It is not easy to find more challenging comparators, however. First, ratings agencies such as Moody’s did not warn markets against the East Asian crises of 1997.³⁸ Goldfajn and Valdés (1998) show that exchange rate expectations of currency traders do not help predict crises. And there is little evidence that interest differentials systematically predict crises.³⁹

The out-of-sample comparison of different approaches provides some insight into important issues in the empirical modeling of currency crises. We have found that reestimating the panel-based KLR and FR models over different samples of countries and longer time periods has preserved most of the economically important results. The STV model has proved largely unstable. More recent efforts to apply STV-like models to the Asian crises have met with some success. While this may help explain the crisis, it seems that the approach of carefully fitting a small set of crises is not promising as a way to predict the next round. To put it another way, specification uncertainty appears to be as important as parameter uncertainty for STV-type approaches, which represent a more complex specification fitted to many fewer observations.

We have also shed some light on the styled facts about crises. All three approaches demonstrate that the probability of a currency crisis increases when domestic credit growth is high, the real exchange rate is overvalued relative to trend, and the ratio of M2 to reserves is high. Both FR and KLR also suggest that a large current account deficit is an important risk factor.⁴⁰ These conclusions imply that elements of both first- and second-generation models are relevant: M2/reserves would seem to play a more important role in second-generation models of crisis that emphasize multiple equilibria, while the other variables are more suggestive of traditional first-generation models.

Where do we go from here? In this paper we have seen that the addition of some plausible variables improves performance of the KLR model somewhat. In

³⁸After years of stable or increasing ratings, the first downgrade in the Asian crisis countries was a negative outlook in Thailand in February 1997 (Moody’s). The rest were not downgraded until mid- to late 1997. See Adams and others (1998).

³⁹The nominal interest differential alone does not predict crises well in our sample of countries. In a bivariate probit regression (not shown), the nominal interest differential is statistically significant, but the goodness of fit is much worse than for the KLR model we have considered, and the out-of-sample forecasts are not helpful. The real interest differential does worse still.

⁴⁰The real exchange rate and the current account are not significant in the original FR specification, as discussed in the Appendix.

a related paper, we depart from the entire “indicators” methodology that looks for discrete thresholds and calculates signal-to-noise ratios.⁴¹ Instead, we apply a probit regression technique to the same data and crisis definition as in KLR. In the process we test some of the basic assumptions of the KLR approach. Specifically, we embed the KLR approach in a multivariate probit framework in which the independent variable takes the value of one if there is a crisis in the subsequent 24 months and zero otherwise. These probit models provide generally better forecasts than the KLR models. In the process, we find also that the data do not generally support one of the basic ideas of the KLR indicator approach: that it is useful to interpret predictive variables in terms of discrete thresholds, the crossing of which is particularly significant for signaling a crisis.

A variety of specification issues appear worth exploring, particularly in the context of probit-based models estimated on panel data. We can be confident that future papers will predict past crises. This exercise suggests, though, that while crisis forecasting models may help indicate vulnerability, the predictive power of even the best of them may be limited.

APPENDIX

Issues in Reestimation and In-Sample Results

In the text we present the KLR, FR, and STV models estimated with a common sample, and analyze the success of the out-of-sample predictions for 1997. This appendix fills in some of the steps. First, we discuss issues involved in the reestimation of the models, including the effects of updating the estimation period, changing the sample, fixing any errors in the original estimates, and using more recently available and hence revised data. Second, we evaluate the in-sample performance of the models.

Kaminsky, Lizondo, and Reinhart (1998) Signals Approach

We first reproduce the KLR results using the same 20-country, 1970–95 sample they use.⁴² Our results are broadly similar to those of KLR, though column 1 of Table A1 shows slightly weaker performance than reported by KLR for most of the indicators. Differences are starker for four indicators, for which KLR find a noise-to-signal ratio substantially below unity while we find a ratio above unity. Thus, although KLR find 12 informative indicators—that is, those with noise-to-signal ratios below unity—we find only 8 of these to be informative.⁴³

Next, we modify the sample in two ways. First, we estimate only through April 1995. This reflects the information available to the analyst just before the Thai crisis of July 1997,

⁴¹See Berg and Pattillo (1998 and 1999).

⁴²Argentina, Bolivia, Brazil, Chile, Colombia, Denmark, Finland, Indonesia, Israel, Malaysia, Mexico, Norway, Peru, the Philippines, Spain, Sweden, Thailand, Turkey, Uruguay, and Venezuela.

⁴³There are a number of possible reasons for the differences in results. We have found that our implementation of the KLR definition of crisis results in a set of crisis dates that do not fully match the KLR crisis dates as reported in Kaminsky and Reinhart (1996). Specifically, we fail to match 14 out of 76 KLR crises. Some of this discrepancy may come from differences in the raw data. We have found that seemingly small differences due to revisions in International Financial Statistics (IFS) data can strongly influence the results, and furthermore they and we separately “cleaned” the data of errors.

Table A1. Performance of Indicators

Indicator	KLR Sample Rerun			Original KLR Estimates		
	Noise/signal (adjusted) ^a (1)	$P(\text{crisis}/\text{signal})$ $-P(\text{crisis})^b$ (2)	Number of crises with data (3)	Noise/signal (adjusted) ^a (4)	$P(\text{crisis}/\text{signal})$ $-P(\text{crisis})^b$ (5)	Number of crises with data (6)
Real exchange rate ^c	0.24	32	70	0.19	39	72
M2/reserves growth rate	0.46	16	69	0.48	17	70
Export growth rate	0.49	15	69	0.42	20	70
International reserves growth rate	0.53	13	69	0.55	16	72
Excess M1 balances ^d	0.69	7	68	0.52	15	66
Domestic credit/GDP growth rate	0.71	6	66	0.62	11	62
Real interest rate	0.74	5	42	0.77	6	44
M2 multiplier growth rate	0.80	4	69	0.61	11	70
Import growth rate	1.19	-3	69	1.16	-3	71
Industrial production growth rate	1.23	-3	58	0.52	16	57
Terms of trade growth rate	1.42	-5	63	0.77	6	58
Lending rate/deposit rate	1.44	-5	34	1.69	-9	33
Bank deposit growth rate	1.53	-6	69	1.20	-4	69
Stock price index growth rate	1.81	-9	47	0.47	18	53
Real interest differential	1.97	-9	42	0.77	6	44
Current account/GDP						
M2/reserves (level)						

^aRatio of false signals (measured as a proportion of months in which false signals could have been issued $[B/(B+D)]$) to good signals (measured as a proportion of months in which good signals could have been issued $[A/(A+C)]$).

^b $P(\text{crisis}/\text{signal})$ is the percentage of the signals issued by the indicator that were followed by at least one crisis within the subsequent 24 months ($[A/(A+C)]$ in terms of the matrix in the text). $P(\text{crisis})$ is the unconditional probability of a crisis, $(A+C)/(A+B+C+D)$.

^cDeviation from deterministic trend.

^dResidual from regression of real M1 on real GDP, inflation, and a deterministic trend.

since the evaluation of an observation requires knowing whether there will be a crisis within 24 months. Second, we change the sample of countries: we omit the five European countries from the sample and add other emerging market economies. This sample is more appropriate for our concern with crises in “emerging markets” and also serves as an informal test of robustness of the KLR approach.⁴⁴ Table 1 in the text shows that indicator performance over the larger sample is broadly similar to results using the KLR sample. The average noise-to-signal ratio falls a little for the informative indicators in the 23-country sample (as well as for the entire set of indicators).

So far we have looked at each indicator separately. Following Kaminsky (1998a), we next calculate the weighted-sum-based probabilities of crisis.⁴⁵ This produces a series of estimated probabilities of crisis for each country. These should be interpreted as the predicted probability of crisis within the next 24 months, based on the (weighted) number of indicators signaling in a given month.⁴⁶

How good are these in-sample forecasts in predicting crises during January 1970 to April 1995? For zero/one dependent variables, it is natural to ask what fraction of the observations are correctly called. A cutoff level for the predicted probability of crisis is defined such that a crisis is predicted if the estimated probability is above this threshold. The resulting goodness-of-fit data are shown in the first two columns of Table A2 for two cutoffs: 50 percent and 25 percent.⁴⁷

What can we conclude? The first column of Table A2 displays the goodness-of-fit measures for the KLR weighted-sum-based probabilities, using the original specification and our new sample. The model correctly calls most observations at the 50 percent cutoff, almost entirely through correct prediction of tranquil periods (i.e., those that are not followed by crises within 24 months). Almost all (91 percent) of the crisis months (i.e., observations followed by a crisis within 24 months) are missed. Even with so few crisis observations correctly called, 44 percent of alarms (i.e., observations where the predicted probability of crisis is above 50 percent) are false, in that no crisis in fact ensues within 24 months. Next, we add the two new variables, current account and M2/reserves in levels. As the second column of Table A2 shows, the addition of these variables only modestly improves the performance of the KLR-based probabilities. A χ^2 test rejects the null that the forecasts and actual outcomes are independent at the 1 percent level.

With a lower cutoff of 25 percent, 41 percent of crisis observations are correctly called by the original KLR model. The probability of a crisis within 24 months is now 37 percent if there is an alarm, much higher than the unconditional probability of crisis of 16 percent in this sample. Now, however, 63 percent of alarms are false. A χ^2 test also rejects the null that the forecasts and actual outcomes are independent at the 1 percent level here.

Our analysis of the in-sample success of the KLR-type models suggests that the approach can indeed be useful and the model does significantly better than guesses based on the unconditional probability of crisis. Nonetheless, most crises are still missed and most alarms are false.

⁴⁴We add the following to the 15 KLR emerging market economies: India, Jordan, Korea, Pakistan, South Africa, Sri Lanka, Taiwan Province of China, and Zimbabwe.

⁴⁵Two issues regarding the treatment of missing data in the KLR framework deserve mention. A key variable is *c24*, which is defined to equal one if there is a crisis in the next 24 months. This variable is defined as long as one observation is available (either a crisis or noncrisis month) in the relevant 24-month period. Secondly, the weighted sum of indicators signaling is calculated provided that data on at least one of the indicators is available. The weighted-sum-based probabilities are calculated using the same principle.

⁴⁶Unlike Kaminsky (1998a), we use only the good indicators, that is, those with noise-to-signal ratio less than one.

⁴⁷See Table 2 footnotes for precise definitions of “correctly called” and related terms.

Table A2. Goodness of Fit of KLR Model—In Sample

Cutoff of 50 Percent
Goodness-of-Fit Table^a

Predicted	Original Specification			Predicted	Augmented with Current Account and M2/reserves		
	Actual				Actual		
	Tranquil	Crash	Total		Tranquil	Crash	Total
Tranquil	5,541	1,114	6,655	Tranquil	5,581	1,115	6,696
Crash	90	115	205	Crash	50	114	164
Total	5,631	1,229	6,860	Total	5,631	1,229	6,860

Summary Statistics

	Original	Augmented
Percent of observations correctly called	82	83
Percent of crises correctly called ^b	9	9
Percent of tranquil periods correctly called ^c	98	99
False alarms as a percent of total alarms ^d	44	30
Probability of crisis given:		
an alarm ^e	56	70
no alarm ^f	17	17

Cutoff of 25 Percent
Goodness-of-Fit Table^a

Predicted	Original Specification			Predicted	Augmented with Current Account and M2/reserves		
	Actual				Actual		
	Tranquil	Crash	Total		Tranquil	Crash	Total
Tranquil	4,790	728	5,518	Tranquil	4,568	658	5,227
Crash	841	501	1,342	Crash	1,063	571	1,634
Total	5,631	1,229	6,860	Total	5,631	1,229	6,860

Summary Statistics

	Original	Augmented
Percent of observations correctly called	77	75
Percent of crises correctly called ^b	41	46
Percent of tranquil periods correctly called ^c	85	81
False alarms as a percent of total alarms ^d	63	65
Probability of crisis given:		
an alarm ^e	37	35
no alarm ^f	13	13

^aTable shows number of observations.

^bA precrisis period is correctly called when the estimated probability of crisis is above the cutoff probability and a crisis ensues within 24 months.

^cA tranquil period is correctly called when the estimated probability of crisis is below the cutoff probability and no crisis ensues within 24 months.

^dA false alarm is an observation with an estimated probability of crisis above the cutoff (an alarm) not followed by a crisis within 24 months.

^eThis is the number of precrisis periods correctly called as a share of total predicted precrisis periods.

^fThis is the number of periods where tranquility is predicted and a crisis actually ensues as a share of total predicted tranquil periods (observations for which the predicted probability of crisis is below the cutoff).

Frankel and Rose (1996) Probit Model Using Multi-Country Sample

Table A3 (column 1) presents our reproduction of the FR benchmark probit regression, using the same sample of annual data for over 100 developing countries for 1970–92. FR conclude from this and a variety of similar regressions that the probability of a crisis increases when output growth is low, domestic credit growth is high, foreign interest rates are high, and FDI as a proportion of total debt is low. They also found support for the prediction that crashes tend to occur when reserves are low and the real exchange rate is overvalued.⁴⁸

We made several revisions to the FR benchmark regression before updating it to 1996. As with the other papers, we used currently available, and hence revised, data from the same World Bank source as FR.⁴⁹ In addition, we corrected an error in the original FR calculation of the overvaluation variable.⁵⁰

The net effect of all these changes is shown in the second regression of Table A3. Overall, the model performs somewhat better than the original FR regression. The corrected overvaluation variable now has a much stronger and more significant effect. Higher northern (OECD) growth now significantly decreases the risk of crisis, and the effect of foreign interest rates is smaller and insignificant.⁵¹

We now estimate the model through 1996 for purposes of generating predictions for 1997. As the third regression in Table A3 shows, the results are similar to the 1970 to 1992 regressions. A large share of debt which is concessional now reduces the risk of crisis.⁵²

Next, we change the sample. The sample of countries used in the original FR regressions is substantially different from those in the KLR and STV regressions. In particular, a large number of least-developed countries (such as the countries of the Council for Mutual Economic Assistance) and small island economies (for example, São Tomé, Cape Verde, and Vanuatu) are included. Because of concerns that crises in these countries may have different determinants and to maximize comparability with the other papers, we have rerun the FR regression over a smaller sample of 41 countries made up of all developing countries with per capita incomes above \$1,000 and population above 1 million for which there are data.⁵³

⁴⁸Although the authors highlight the importance of low reserves and overvaluation in their conclusion, their results show significant effects were not robust and were found in fewer than half of the specifications they tested. The result that faster domestic growth reduces the probability of crisis is also not robust, as illustrated by the benchmark regression itself.

⁴⁹This changed not only some of the data but also the sample, because some of the data that had previously been available, largely from the early 1970s, are now considered to be of unacceptable quality, while other formerly unavailable observations now had data. The net effect is to increase the number of observations from 780 in FR to 881, though the overlap of common data points is only 729 observations.

⁵⁰We also made two other technical modifications. First, we used percent changes instead of log differences in comparing the devaluations with the 25 percent crisis threshold. Second, we changed the implementation of the “windowing” procedure to more closely match the FR intent of ensuring that only the first of a sequence of crises was counted in the sample. See Milesi-Ferretti and Razin (1998), who recommended these two modifications.

⁵¹For the overvaluation variable itself, the correction is the source of the improvement. For the other variables, the changes in sample resulting from the data revision are more important than the data revisions themselves, the changes in the windowing procedure and definition of crisis, or the correction of the overvaluation variable in driving these changes in results.

⁵²For purposes of predicting 1997 outcomes, we also estimate this regression with the government budget as a share of GDP excluded from this regression, because this variable is not available for 1996 as would be required for forecasting 1997. This omission makes little difference.

⁵³Milesi-Ferretti and Razin (1998) raise these sample issues and extract this smaller sample, for which they get improved results compared with FR.

Table A3. Frankel and Rose: Probit Estimates of Probability of a Currency Crash

	Original 1970–92 Sample				1970–96 Sample	
	Model 1A FR benchmark		Model 2A revised		Model 3A FR benchmark	
	dF/dx	$ z ^a$	dF/dx	$ z ^a$	dF/dx	$ z ^a$
Commercial bank share						
of total debt	0.03	0.2	-0.07	-0.6	0.03	-0.3
Concessional share	-0.14	-2.1	-0.12	-1.6	-0.13	-1.9 *
Variable rate share	-0.03	0.2	0.20	1.5	0.13	1.1
Short-term share	0.23	2.0	0.28	2.3 **	0.27	2.2 **
FDI/debt	-0.31	-2.5	-0.53	-3.7 ***	-0.46	-3.3 ***
Public sector share	0.19	2.2	0.18	2.0 **	0.16	1.8 *
Multilateral share	-0.06	-0.8	0.08	1.0	0.06	0.8
Debt/GNP	-0.04	-1.7	-0.02	-1.5	-0.02	-1.4
Reserves/imports	-0.01	-3.4	-0.01	-3.3 ***	-0.01	-3.9 ***
Current account/GDP	0.02	0.2	-0.03	-0.3	-0.02	-0.2
Overvaluation ^b	0.08	2.5	0.15	4.0 **	0.12	3.3 ***
Government budget surplus/ GDP	0.16	1.1	0.10	0.7	0.11	0.8
Domestic credit growth	0.10	3.2	0.08	3.8 ***	0.08	3.7 ***
GDP growth rate	-0.16	-1.3	-0.07	-1.5	-0.08	-1.6
Foreign interest rate	0.80	2.6	0.48	1.4	0.33	1.0
Northern (OECD) growth	-0.85	-1.5	-1.17	-1.9 *	-1.52	-2.5 **
Sample size	780		881		940	
Pseudo R^2	0.17		0.17		0.15	

Goodness of Fit

	Model 1A			Model 2A			Model 3A		
	Tranquil	Crash	Total	Tranquil	Crash	Total	Tranquil	Crash	Total
Cutoff probability of 50 percent ^c									
Predicted tranquility	707	64	771	777	88	865	830	97	927
Predicted crash	4	5	9	7	9	16	5	8	13
Total	711	69	780	784	97	881	835	105	940
Cutoff probability of 25 percent ^d									
Predicted tranquility	678	52	730	743	62	805	792	72	864
Predicted crash	33	17	50	41	35	76	43	33	76
Total	711	69	780	784	97	881	835	105	940

^aOne, two, and three asterisks denote significance at the 10, 5, and 1 percent levels, respectively.

^bDefined as the deviation from the average real exchange rate over the period.

^cA crisis is correctly called when the estimated probability of crisis is above 50 percent if a crisis ensues within 24 months. A tranquil period is correctly called when the estimated probability of crisis is below 50 percent and there is no crisis within 24 months.

^dA crisis is correctly called when the estimated probability of crisis is above 25 percent if a crisis ensues within 24 months. A tranquil period is correctly called when the estimated probability of crisis is below 25 percent and there is no crisis within 24 months.

The results are broadly similar, as regression 1 of Table 4 shows. The most notable changes are that the ratio of reserves to imports is no longer significant whereas the current account and the fiscal balance now are.

The main text discusses our consideration of some alternative explanatory variables. Regression 2 of Table 4 includes the ratio of the reserves to M2 and the degree of openness of the economy, as a result of this specification search. This model suggests that the probability of a crash increases when concessional debt and FDI are small and public sector debt large as a share of total external debt, the ratio of reserves/M2 is low, the current account deficit is large, the real exchange rate is overvalued, domestic credit growth is high, foreign interest rates are high, and the country is not open to trade.

Model 3A of Table A3 is close to the original FR specification, with some corrections and minor revisions, while model 2 of Table 4 is our augmented specification using a more homogeneous sample. The diagnostic statistics show that, in-sample, these models rarely generate a predicted probability of crash above 50 percent. Model 3A correctly predicts only 8 out of the 105 crashes; model 2 (Table 4) does better, predicting one-third of the crashes in the sample. When an estimated probability of above 25 percent followed by a crash is considered success, the results look better. Model 2, for example, generates a probability above 25 percent before 63 percent of crises. About half of warnings defined this way (41 out of 79) were not followed by a crash.

The FR models thus show some promise for predicting crises based on this in-sample assessment. There is a fair amount of parameter stability across samples, and many sensible variables are significant predictors of crisis. The overall explanatory power is fairly low, though our modifications lead to some improvement here.

Sachs, Tornell, and Velasco (1996) Cross-Country Regressions

The text discussed reproduction of the original STV benchmark regression, using their data,⁵⁴ as well as results using revised data and estimating over the common 23-country sample (Table 5, regressions 1, 2, and 3, respectively). We also considered a revised specification based on a different definition of the real exchange rate (Table 5, regression 4).

Table A4 shows some further variants of the STV regressions for the 1994–95 sample. Regression 5 is another variant on the definition of the real exchange rate variable, measuring *RER* as the level of the real exchange rate in 1994 compared with its average over the 1986 to 1989 period. It is also quite similar to the benchmark specification in Table 5, regression 3.

The definitions of low reserves and weak fundamentals in terms of which quartile of the sample the country finds itself are somewhat arbitrary. For this reason, STV vary the definition of low reserves and weak fundamentals so that countries in different fractions of the sample qualify. For example, regression 6 of Table A4 reproduces the STV results for the

⁵⁴Regression 1 differs slightly from the published STV benchmark, mainly because we have corrected an error in the calculation of *RER* for Taiwan Province of China, in STV. The resulting differences are statistically, numerically, and economically small. In addition, the data used both in the STV benchmark and regression 1 differ slightly from that described and published in STV. First, the data published in STV (but not that used in their regressions) contain several typographical errors, which we have corrected with the help of the authors. Second, here and in the STV regression the lending boom variable was calculated differently for Peru than for the other countries and as defined in the appendix of STV. Specifically, LB is defined as the growth from 1990 through 1994 in the ratio of domestic credit to the private sector to GDP. For Peru, however, the base year actually used is apparently 1991. This is presumably because the hyperinflation and stabilization of 1989/1990 led to a tiny base of credit/GDP and would have resulted in a large outlier for Peru if calculated as defined in STV. Third, the measure of reserves for South Africa apparently includes gold reserves, as is standard for that country but contrary to the description in the appendix of STV.

Table A4. STV: 1994/5 Regressions

Regression Number	Number of Countries	R^2	\bar{R}^2	Results ^{a, b}						
				STV hypotheses: $\beta_2 = 0$	$\beta_2 + \beta_4 = 0$	$\beta_2 + \beta_4 + \beta_6 < 0$	$\beta_3 = 0$	$\beta_2 + \beta_5 = 0$	$\beta_2 + \beta_5 + \beta_7 > 0$	
With quartiles										
5	23	0.43	0.21	2.22 (2.566)	-0.30 (0.134)	-2.69 (1.141)	0.46 (0.710)	-5.72 (5.779)	1.80 (1.351)	
With halves										
6	20	0.68	0.54	-0.52 (1.545)	0.48	-6.58	0.26 (0.547)	-1.80	3.32	
7	20	0.66	0.50	-0.20 (0.952)	0.15 (0.881)	2.51 (4.657)	0.17 (0.399)	2.06 (3.832)	4.94 (2.209)	
8	23	0.64	0.51	-0.26 (1.054)	1.50 (1.216)	-14.45 (3.992)	0.93 (0.724)	5.50 (9.222)	1.93 (0.731)	

^aCoefficients in bold are significant at the 5-percent level. Bolded coefficients are significantly inconsistent with the STV hypothesis. Figures in parentheses are standard errors.

^bThe β s are coefficients from the regression $IND = \beta_2 RER + \beta_3 LB + \beta_4 RER \cdot DLR + \beta_5 RER \cdot DWF + \beta_6 RER \cdot DWF + \beta_7 LB \cdot DWF$, where RER is the degree of real depreciation, LB is a measure of the lending boom, DLR is a dummy variable for countries with low reserves, and DWF is a dummy for countries with weak fundamentals (see text for explanations).

case where “low reserves” is defined as having a reserves/M2 ratio in the bottom *half* of the sample, while “weak fundamentals” is having low reserves or an exchange rate depreciation in the lower *half* of the sample. The main results continue to hold. Regressions 7 and 8 of Table A4 present the reestimation of regression 5 with revised data and correcting the Taiwan Province of China crisis variable. Unlike with the quartile regressions, this changes the results: most important, *RER* with low reserves and weak fundamentals ($\beta_2 + \beta_4 + \beta_6$) now has the wrong sign, though it is insignificant.⁵⁵

A number of the STV results are not robust to the data revisions that have taken place since their estimations and to the addition of three countries to the sample. The fit of the models is generally poorer and the main hypotheses receive mixed support at best.

REFERENCES

- Adams, Charles, Donald J. Mathieson, Garry Schinasi, and Bankim Chadha, 1998, *International Capital Markets* (Washington: International Monetary Fund, September).
- Berg, Andrew, 1999, “The Asian Crisis: Causes, Responses, and Outcomes” (unpublished; Washington: International Monetary Fund).
- and Catherine Pattillo, 1999, “Predicting Currency Crises: The Indicators Approach and an Alternative,” *Journal of International Money and Finance*, forthcoming.
- , 1998, “Are Currency Crises Predictable?: A Test,” IMF Working Paper 98/154 (Washington: International Monetary Fund).
- Blanco, Herminio, and Peter M. Garber, 1986, “Recurrent Devaluation and Speculative Attacks on the Mexican Peso,” *Journal of Political Economy*, Vol. 94 (February), pp. 148–66.
- Bussière, Matthieu, and Christian Mulder, 1999, “External Vulnerability in Emerging Countries: The Trade-Off Between Fundamentals and Liquidity” (unpublished; Washington: International Monetary Fund).
- Calvo, Guillermo A., and Enrique G. Mendoza, 1996, “Mexico’s Balance-of-Payments Crisis: a Chronicle of a Death Foretold,” *Journal of International Economics*, Vol. 41 (December), pp. 235–64.
- Corsetti, Giancarlo, Paolo Pesenti, and Nouriel Roubini, 1998a, “Paper Tigers? A Model of the Asian Crisis,” NBER Working Paper No. 6783 (Cambridge, Massachusetts: National Bureau of Economic Research).
- , 1998b, “What Caused the Asian Currency and Financial Crisis? Part I: A Macroeconomic Overview,” NBER Working Paper No. 6833 (Cambridge, Massachusetts: National Bureau of Economic Research).
- , 1998c, “What Caused the Asian Currency and Financial Crisis? Part II: The Policy Debate,” NBER Working Paper No. 6834 (Cambridge, Massachusetts: National Bureau of Economic Research).
- Flood, Robert, and Nancy Marion, 1998, “Perspectives on the Recent Currency Crisis Literature,” NBER Working Paper No. 6380 (Cambridge, Massachusetts: National Bureau of Economic Research).
- , 1994, “The Size and Timing of Devaluations in Capital-Controlled Developing Economies,” NBER Working Paper No. 4957 (Cambridge, Massachusetts: National Bureau of Economic Research).

⁵⁵In this case, part of the reason for the difference is that, even using the (typo-corrected) STV data, we were not able to reproduce regression 5.

- Frankel, Jeffrey A., and Andrew K. Rose, 1996, "Currency Crashes in Emerging Markets: An Empirical Treatment," *Journal of International Economics*, Vol. 41 (November), pp. 351–66.
- Furman, Jason, and Joseph Stiglitz, 1998, "Economic Crises: Evidence and Insights from East Asia," *Brookings Papers on Economic Activity: 1*, Brookings Institution, pp. 1–136.
- Goldfajn, Ilan, and Rodrigo Valdés, 1997, "Are Currency Crises Predictable?" *European Economic Review*, Vol. 42 (May), pp. 873–85.
- Goldstein, Morris, 1998, "Early Warning Indicators and the Asian Financial Crisis" (unpublished; Washington: Institute for International Economics).
- International Monetary Fund, 1998, *World Economic Outlook* (Washington: International Monetary Fund, May).
- Kaminsky, Graciela, 1998a, "Currency and Banking Crises: A Composite Leading Indicator," IMF Seminar Series No. 1998–6, February.
- , 1998b, "Financial Crises in Asia and Latin America: Then and Now," *American Economic Review, Papers and Proceedings*, Vol. 88 (May), pp. 44–48.
- , and Carmen M. Reinhart, 1996, "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems," International Finance Discussion Paper No. 544 (Washington: Board of the Governors of the Federal Reserve System, March).
- Kaminsky, Graciela, Saul Lizondo, and Carmen Reinhart, 1998, "Leading Indicators of Currency Crises," *Staff Papers*, International Monetary Fund, Vol. 45 (March), pp. 1–48.
- Krugman, Paul, 1979, "A Model of Balance-of-Payments Crises," *Journal of Money, Credit and Banking*, Vol. 11 (August), pp. 311–25.
- Lane, Timothy, and others, 1999, "IMF-Supported Programs in Indonesia, Korea and Thailand: A Preliminary Assessment," available via Internet: <http://www.imf.org/external/pubs/ft/op/opasia/asia1.pdf> (Washington: International Monetary Fund).
- Milesi-Ferretti, Gian Maria, and Assaf Razin, 1998, "Current Account Reversals and Currency Crises: Empirical Regularities," IMF Working Paper 98/89 (Washington: International Monetary Fund).
- Radelet, Steven, and Jeffrey Sachs, 1998a, "The Onset of the East Asian Financial Crisis," NBER Working Paper No. 6680 (Cambridge, Massachusetts: National Bureau of Economic Research).
- , 1998b, "The East-Asian Financial Crisis: Diagnosis, Remedies, Prospects," *Brookings Papers on Economic Activity: 1*, Brookings Institution, pp. 1–90.
- 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.
- Sachs, Jeffrey, 1997, "What Investors Should Learn from the Crisis That Has Forced Thailand to Seek an IMF Loan," *Financial Times* (London), July 30.
- Tornell, Aaron, 1998, "Common Fundamentals in the Tequila and Asian Crises" (unpublished; Cambridge, Massachusetts: Harvard University).
- Werner, Alejandro, 1996, "Mexico's Currency Risk Premia in 1992–94: A Closer Look at the Interest Rate Differentials," IMF Working Paper 96/41 (Washington: International Monetary Fund).