# Can We Predict the Next Capital Account Crisis?

MARCOS CHAMON, PAOLO MANASSE, and ALESSANDRO PRATI[*]

*This paper uses binary classification trees (BCTs) to predict capital account crises. BCTs successively compare candidate variables and thresholds to split the data into two subsamples, allowing for a large number of indicators to be considered and complex interactions to emerge in a way that standard regressions cannot easily replicate. We identify a robust leading indicator role for three variables (international reserves, current account balance, and short-term external debt) as well as a reserve cover measure that combines them. External indebtedness and domestic GDP growth forecasts are also important predictors of vulnerability. Out of sample, we were able to capture some of the main emerging market crises with relatively few false alarms but the overall out-of-sample performance of our forecasts was mixed. Global cyclical variables help explain vulnerability to crises but they are difficult to predict and, therefore, are of limited use for forecasting purposes.* [JEL F02, F31, F32, F33, F34]

Predicting capital account crises is extremely difficult. Academics and policymakers have identified several factors that contribute to their inception, but the imbalances at the origin of each wave of crises, as well as their propagation mechanism, keep changing over time and, with them, the set of potential crisis indicators. A typical (noncomprehensive) list includes measures of real exchange rate (RER) misalignment; terms of trade shocks; international reserves; external and public debt; monetary and fiscal policy; balance sheet mismatches (currency and maturity) in the corporate, banking, and government sectors; political uncertainty; global cyclical and financial conditions; and general market sentiment.

This unwieldy set of potential indicators makes it difficult to compare their information content using standard regression techniques. To remedy this problem and isolate a parsimonious set of robust leading indicators of crises within a group of more than 100 vulnerability indicators, we used a nonparametric statistical methodology, called Binary Classification Trees (BCTs). BCTs can handle a large set of variables and their interactions, and select critical thresholds. For example, we found that three simple conditions—based on international reserve coverage and the level and change of external debt-to-GDP ratios—selected a subset of observations in which the frequency of capital account crises was 21.3 percent as opposed to the sample average of 6.1 percent.

We applied BCTs to a new data set of 34 capital account crises that took place in 49 countries during the period 1994–2005. Capital account crises, or "sudden stops," were defined as large and sudden reversals in net private capital flows. We dated all crises on the basis of their inception and all indicators were lagged one year so that only precrisis information was used. We also included one-year-ahead forecasts of contemporaneous variables (for example, *World Economic Outlook* (WEO) forecasts of GDP growth or current account balances) and market-based forward-looking indicators such as lagged Emerging Markets Bond Index (EMBI) spreads. We used previous lists of crises and numerical rules to select and date potential capital account crises, which IMF's country desks then revised and validated. This last step was important because numerical rules occasionally identify capital flow reversals that have noncrisis explanations that country desks may provide (for example, the end of a privatization program).

The in-sample fit of the BCT estimated on the entire sample was reasonably good, with four indicators (and their respective thresholds) breaking down the sample into (1) a subsample with a frequency of crises 3.5 times as high as in the overall sample, (2) another subsample with a frequency of crises twice as high as in the overall sample, and (3) three "safe" subsamples with a minimal frequency of crises (about 1 percent).

BCTs yielded mixed results when used out of sample. A BCT based on information up to the year 2000 would have correctly predicted three of the five crises in 2001, including in Argentina, Turkey, and Lebanon. A BCT based on information up to 2001 would have missed all crises of 2002 (Brazil, Colombia, Israel, and Uruguay), whereas a BCT based on information up to

2002 would have perfectly predicted the two crises of 2003 (Dominican Republic and Jamaica).

Would BCTs have predicted the Asian crisis? Given that this crisis took place in 1997 and our sample began in 1994, we could not meaningfully estimate a BCT on the previous three years of data. We could, however, estimate a BCT on a sample that excludes all observations corresponding to East Asian countries and check how it would split the latter into crisis and noncrisis years. On the basis of the level of international reserve coverage (the lagged ratio of international reserves to the sum of current account deficits and short-term external debt), this BCT would have initially classified Thailand, Indonesia, Korea, and the Philippines into a group of countries with a frequency of crises twice as high as in the rest of the sample, and considered Malaysia less vulnerable than the average country in the sample. However, all East Asian countries would have been misclassified as safe once we had taken into account a second set of conditions based on exchange rate overvaluation or fiscal positions, which are good predictors of crises in the subsample without East Asian countries.[1] This out-of-sample exercise reinforces the view that the East Asian crises were somewhat special and would have been difficult to predict.

BCTs also allow us to address the issue of the relative role of global economic conditions and country-specific imbalances in capital account crises. Is it true that during global booms or periods of abundant liquidity in capital markets even countries with serious domestic imbalances can remain unscathed? To answer this question, we estimated a variant of the full-sample BCT including contemporaneous global indicators (which are exogenous to crisis events in individual countries). We found that two gauges of the global conditions each country faces—commodity export prices and import demand by trading partners—contribute to explaining the occurrence of crises. For example, when real commodity prices were at least 13.5 percent below their historical country-specific average, the frequency of crises in countries with low reserve coverage rose from 14 to 22.6 percent; by contrast, when real commodity export prices were higher than this threshold, the frequency of crises dropped from 14 to 2.8 percent. In other words, low commodity export prices are a key trigger of crises in countries with low reserve coverage. This finding is, however, of little use for crisis prediction because forecasts, or lagged values, of global indicators are not as good as their contemporaneous values at separating crisis from noncrisis episodes.

The empirical literature on early warning systems (EWS) shares with this paper the focus on crisis prediction. Frankel and Rose (1996); Kaminsky, Lizondo, and Reinhart (1998); and Berg and Pattillo (1999) wrote seminal EWS papers. In the same spirit of this paper, Berg, Borensztein, and Pattillo

---

[1]That is, precrisis data of East Asian countries point to sound fiscal positions and no exchange rate misalignment (although some indication of the latter would emerge if postcrisis information were used).

(2005) analyzed the in-sample and out-of-sample fit of EWS models, showing that the latter varies substantially by model and forecast horizon. The EWS papers differ from our paper in the empirical methodology, the prevalent focus on currency crises, and the monthly frequency of observations. BCTs can assess the predictive power of a much richer set of indicators and experiment with more interactions than can EWS. Furthermore, the BCT algorithm selects indicators and thresholds by taking into account the preferred trade-off between the cost of missing crises and that of predicting crises erroneously, whereas the EWS probit/logit models can only be estimated independently from that trade-off. These advantages translate into better in-sample crisis prediction performance of BCTs, with 88 percent of crises correctly called (30 out of 34) as opposed to 60–70 percent in typical EWS models. Comparing out-of-sample performance of BCTs and EWS is much more difficult because of the different periods and crises considered. For example, BCTs predicted correctly out of sample the 2001 crises in Argentina and Turkey, which the EWS models considered in Berg, Borensztein, and Pattillo did not try to predict, but BCTs were less successful than EWS models in predicting the Asian crisis.

Few studies have used the BCT methodology. Ghosh and Ghosh (2003), Frankel and Wei (2004), and Kaminsky (2006) applied it to currency crises and assessed its in-sample forecasting performance. Manasse and Roubini (2005) used BCTs to study the determinants of sovereign crises and to predict them. Van Rijckeghem and Weder (2004) developed a nonparametric technique similar to BCTs to study the political determinants of debt crises. Our paper is the first application of BCTs to capital account crises.

## I. Methodology

BCTs are a nonparametric statistical technique suitable for identifying complex interactions among variables with the objective of predicting binary outcomes (in our case, "crisis inception" or "no crisis inception"). BCTs identify the indicators and their thresholds that can better separate the sample into crisis and noncrisis observations. The order in which the indicators are used in each split allows complex interactions to emerge, in a way that would be difficult to replicate in a standard regression approach.

BCTs' classification rules are a collection of inequalities, such as the following: *if* (1) international reserves cover less than 80 percent of the sum of short-term external debt and the current account deficit; (2) external debt is higher than 24 percent of GDP; and (3) external debt is not falling by at least 3 percent of GDP per year, *then* the frequency of crises next year is 21 percent and the observation is classified as "crisis prone."

We computed BCTs using the nonparametric statistical algorithm CART (Classification and Regression Trees; Breiman and others, 1984). In a nutshell, the BCT algorithm computes a score of how well each variable does at separating crisis from noncrisis observations, and splits observations into two groups based on the variable with the highest score. The process

continues for each branch of the data and eventually stops according to the criteria used to measure further improvements. Like other nonparametric methods, BCTs are apt tools for detecting nonlinearities, which is critical when an indicator has information only for values beyond certain thresholds. In theory, standard regression techniques (for example, probit or logit models) could be used for similar purposes. However, even for a very small set of indicators, it would be impossible to experiment with all possible interactions and thresholds in a single regression.[2] Another important drawback of parametric regression approaches is the need to make assumptions about a functional form. Given the lack of well-established theoretical relationships between the variables used for predicting crises and outcomes, it may become more attractive to rely on a complex set of threshold interactions rather than on a parametric functional form.

The key elements of the BCT analysis are a set of rules for (1) splitting each node into two child nodes; (2) assigning each node to a class outcome (for example, crisis vs. noncrisis); and (3) deciding when to stop growing the tree.

The BCT algorithm starts by comparing candidate variables and thresholds to split the sample into two child nodes. All splitting rules are based on whether or not a variable is above or below a threshold. Each split is assigned a score based on how it improves the purity of the classification. A variable and a threshold that perfectly separate all crisis observations from all noncrisis observations would yield the purest possible classification. In practice, however, each possible split classifies observations in two groups that have both crises and noncrises. The BCT algorithm computes a cost that rises with the extent by which the actual classification departs from the perfect classification, and selects the split that minimizes such cost. The tree is grown by repeating this process on the child nodes. The loss function being minimized is

$$Cost = \sum_{i} \left( (N_0(i) + N_1(i)) \widehat{p}_0(i) \widehat{p}_1(i) \right),$$

where $N_0(i)$ and $N_1(i)$ are the number of noncrisis and crisis observations (respectively) in terminal node $i$, and

$$\widehat{p}_1(i) = \frac{\frac{C_1\pi_1}{C_0\pi_0 + C_1\pi_1} \frac{N_1(i)}{N_1}}{\frac{C_0\pi_0}{C_0\pi_0 + C_1\pi_1} \frac{N_0(i)}{N_0} + \frac{C_1\pi_1}{C_0\pi_0 + C_1\pi_1} \frac{N_1(i)}{N_1}},$$

$$\widehat{p}_0(i) = 1 - \widehat{p}_1(i),$$

where $C_0$ is the cost of misclassifying a noncrisis observation, $C_1$ the cost of misclassifying a crisis observation, $\pi_1$ the prior probability that an observation is a crisis, $\pi_0$ the prior probability that it is a noncrisis, and $N_0$ and $N_1$ are the number of noncrisis and crisis observations in the sample.

---

[2]For example, the number of possible interactions of indicators and threshold values in our data exceeds by several orders of magnitude the number of observations.

The relative misclassification costs and prior probabilities can be used interchangeably in order to make the crisis classification more conservative. Given our subjective preference for reducing the chance of missing crises, we chose parameter values that made the BCT algorithm classify as crisis nodes all those where the crisis frequency was at least twice as high as in the sample.[3] Specifically, to make the BCT algorithm follow this rule, we set the prior probability of a crisis to 6.1 percent, which was the frequency of crises in our sample, and the cost of misclassifying crises as noncrises to 7.7 times that of misclassifying noncrises as crises. Alternatively, we could have obtained the same results by assuming a less asymmetric misclassification cost combined with a larger prior probability that the observation is a crisis.[4] Our conservative approach implies that the crisis nodes of our BCTs tend to be characterized by a low crisis frequency and a relatively higher number of misclassified noncrises. Perturbating the parameters to make our approach more or less conservative did not affect the variable and threshold in the top split of our BCTs but occasionally affected lower-level splits.

The option of choosing misclassification costs at the outset (that is, before running the BCT algorithm) to influence the model choice (that is, the set of indicators and thresholds) is a key difference between BCTs and EWS. In fact, Berg, Borensztein, and Pattillo (2005) used a misclassification cost function (weighting Type I and Type II errors) to identify the probability cutoff point that would best predict crisis and noncrisis events out of sample only after having estimated the probit model. As a result, their cost function cannot influence the choice of the indicators and coefficients in the probit model.

We used our judgment to decide when data were sufficiently partitioned. This decision was critical because in BCTs there is not necessarily an optimal number of splits. In fact, it is always possible to use a very large set of rules to attain a perfect classification. Increasing the number of splits, however, may lead to poor out-of-sample forecasts, similar to what happens in regression analysis when the number of regressors increases.

Although we used judgment in selecting the size of the trees to present, we also took into consideration the results of a technique called "V-fold cross-

---

[3]The rationale for choosing a threshold that is twice the sample frequency of crises is as follows: Crises were relatively rare events in our sample. If we had set, for example, the prior probability of crisis at the sample frequency of 6.1 percent, a very high share of crisis observations in a node (at least 35 percent) would have been necessary to classify it as a crisis node, despite the asymmetric misclassification cost imposed. Because we preferred to err on the side of being conservative, we required a much smaller frequency of crisis observations to classify a node as crisis prone. The threshold around 12.2 percent used for the entire sample allowed us to be conservative while still acknowledging that crises are relatively rare events. We applied this same logic to the choice of the misclassification cost parameter. As in that case, the option of influencing the model selection by choosing the frequency of crises required to classify a node as crisis prone is a feature of BCTs that distinguishes them from probit-based EWS.

[4]For example, if we had set the cost of misclassifying crises as noncrises to twice that of misclassifying noncrises as crises, we would have obtained the same results by setting the prior probability that an observation is a crisis to 20 percent.

validation.'' This technique amounts to using out-of-sample performance as a guide to select the best number of splits. The sample is divided into 10 parts and, then, each 10 percent of the observations is used, in turn, to test the predictive power of 10 ancillary trees estimated on the remaining 90 percent of observations (in a way that each observation is used once and only once in an ancillary test sample). On the basis of the out-of-sample performance of these ancillary trees, the algorithm proposes an optimal level of complexity (measured by the number of terminal nodes) for the tree estimated on the full sample. Section IV discusses the several reasons why we often overrule the V-fold's proposed pruning. The main reason is that, in many instances, the V-fold cross-validation technique suggests trees with no split[5] or including many splits, some of which make no economic sense.

Many indicators have missing values (years, countries, or both).[6] The BCT algorithm does not drop observations for which some indicators are missing, unlike, say, a regression in which missing observations would be dropped. When a variable with missing observations was used to split the data, the missing observations were assigned to the partitioning of the sample that minimizes the cost function. To prevent this default rule from influencing the selection of the best indicators, we penalized indicators with missing observations. In practice, this choice forced indicators with missing observations to be used for splitting smaller partitions of the data (for which their coverage is reasonable) or not to appear at all.[7]

The BCT algorithm can be applied to a very large number of candidate predictors: unlike in a standard regression, the inclusion of irrelevant indicators, which are not used for splitting the data, does not affect the results. However, when an indicator slightly outperforms another as a ''splitter,'' the latter may never appear in the final tree even though its information content is almost as good as that of the top splitter. To avoid drawing the incorrect inference that all omitted indicators are not important, we checked the competing indicators for the top split.

As a robustness test for our selection of indicators and as a benchmark for out-of-sample prediction, we used a new procedure called RandomForests that Breiman (2001)—one of CART's developers—proposed as a way to address the problem of few additional variables or observations substantially changing the BCTs. Adding variables will not change the BCTs if the new variables do not improve any of the splits obtained with the preexisting

---

[5]The V-fold cross-validation methodology assumes that, in the no-split tree, all observations are crises. The no-split tree has, then, a zero Type I error but the highest possible Type II error.

[6]There was substantial variation in data coverage across countries and time. Some indicators were available for only a subset of countries (for example, corporate vulnerability indicators). Others were not available at the beginning of the sample (for example, detailed financial vulnerability indicators or data for transition countries).

[7]If the fraction of missing values of an indicator is 50 percent, its improvement score will be multiplied by 25 percent.

variables. However, if one of the new variables is informative enough to replace a preexisting variable even in a single split, there is a good chance that the branch developing from that split onward will feature a completely new set of variables and thresholds. Similarly, the introduction of additional years or countries to the sample may lead to substantial changes in the optimal tree if and where changes occur. Breiman proposed an algorithm based on a collection of hundreds of trees that classifies and predicts each observation according to the response of the majority of trees. A bootstrap procedure over two dimensions selects the sample and the list of variables used to estimate each tree (hence the algorithm's name, RandomForests).

In our application of the RandomForests algorithm, we grew trees on 1,000 bootstrapped samples allowing three randomly chosen indicators to be used to split the data at each node. By randomizing over the variables, each tree was likely to involve a very large number of different splitters. By randomizing simultaneously over the sample, each tree in the forest analyzed only small portions of data at a time. This process, called "slow learning," highlights different aspects of the data set and reduces the risk of drawing wrong conclusions "too fast" (see Friedman, 1991). If a pattern genuinely exists in the portion of data analyzed, the RandomForests algorithm will detect it repeatedly in different trees; conversely, it will wash out any accidental pattern in the process of averaging the results. Although this algorithm can improve predictive accuracy, it has the important drawback of not allowing the researcher to recover thresholds and variable interactions, because it relies on aggregating many different trees of different shapes. For this reason, we used it only as a robustness check.

## II. The Data Set

The sample covers 49 countries during the 1994–2005 period.[8] The countries in the sample are listed in Appendix I. The coverage focuses on emerging market countries that had significant access to private international financial markets and did not have a substantial net foreign asset position. Very small economies (with GDP below US$7.5 billion at the end of the sample) were not considered no matter what level of income per capita they had.

### Crisis Definition

We define "capital account crises" as sudden stops in capital flows that are likely to be associated with currency, sovereign, banking, or corporate crises. Table 1 lists crisis episodes for our 49-country sample. Only the first year of a capital reversal (the crisis inception) is considered. This selection of the list of crises is the result of a concerted effort by the IMF staff's working group on

---

[8]The period 1994–2005 was chosen because the capital account regime was relatively stable in most countries and because we wanted to have only post-transition years for Central and Eastern European countries.

| Table 1. Capital Account Crisis Episodes by Year of Inception | | | | | | |
|---|---|---|---|---|---|---|
| Year | Countries | | | | | |
| 1994 | Algeria | Bulgaria | Mexico | Turkey | Ukraine | Venezuela |
| 1995 | Argentina | | | | | |
| 1996 | Hungary | | | | | |
| 1997 | Czech Republic | Indonesia | Israel | Korea | Malaysia | Philippines Thailand |
| 1998 | Brazil | Pakistan | Russia | Ukraine | | |
| 1999 | Colombia | Ecuador | Lithuania | Romania | | |
| 2000 | | | | | | |
| 2001 | Argentina | Lebanon | South Africa | Turkey | Venezuela | |
| 2002 | Brazil | Colombia | Israel | Uruguay | | |
| 2003 | Dominican Republic | Jamaica | | | | |
| 2004 | | | | | | |
| 2005 | | | | | | |

Source: Authors' and IMF staff calculations. This list of episodes should not be interpreted as an IMF's official list of capital account crises.

vulnerability indicators. The following two-stage procedure was used. We identified a first set of potential crisis episodes on the basis of various definitions of crises, including two measures of sudden stop in net private capital flows,[9] years of high exchange rate pressure as indicated by EWS, sovereign defaults, IMF programs (only years with positive net disbursements), a banking indicator, and a corporate crisis indicator.[10] Second, we chose the final set of crisis years, taking into consideration comments from IMF desk economists. Their suggestions helped us solve ambiguities on the timing of the crisis inception, discard years that were identified by some crisis indicator but should not have been considered a capital account crisis, and add one episode that no crisis indicator had picked up. Appendix I lists the different crisis indicators, and Appendix II provides country-by-country details on the selected crises.

---

[9]There is no standard definition of "sudden stop." In some cases, a sharp and sudden reversal in capital flows is easy to classify as a sudden stop (for example, Thailand in 1997). In other instances, a steady decline takes place over a prolonged period, resulting in a crisis (for example, Venezuela from 1998 to 2000). In this latter case, it is not straightforward to determine the inception year. Footnotes 3 and 4 in Appendix I describe the numerical rules used to address this issue in a systematic manner. Somewhat related rules were used by Catão (2006).

[10]This initial selection of potential capital account crisis years was based mainly on the sudden stop indicators. The other indicators helped to select potential crisis episodes that did not translate into a substantial deterioration in net private capital flows or to fine-tune the year of inception of the crisis. Sovereign crises are from Manasse and Roubini (2005), updated with the sovereign debt default indicator of Celasun, Debrun, and Ostry (2006). The banking crisis indicator is based on Demirguc-Kunt and Detragiache (1998), updated by the IMF's Monetary and Financial Department (MFD). The corporate crisis indicator is based on the Corporate Vulnerability Utility (CVU), developed by the IMF's Research Department.

There are 554 country-year observations, of which 34 (6.1 percent) correspond to the year of inception of a crisis, and the rest to noncrisis years. In fact, we dropped from the sample the observations corresponding to years immediately following a crisis because their characteristics are clearly different from those of noncrisis years. At the same time, these postcrisis years should not be confused with the crisis-inception years because they may be easier to predict using previous-year indicators that already reflect the impact of a crisis. Of course, dropping only the first year after a crisis is a relatively arbitrary way of dealing with this problem. Nonetheless, dropping additional postcrisis years did not change the results.

## Indicators

The IMF's WGVI also suggested the core set of indicators used in this paper.[11] They cover four sectors:

- *External sector:* (1) gross international reserve coverage (relative to maturing external debt and the current account deficit); (2) current account balance (in percent of GDP); (3) real exchange rate (RER) overvaluation; (4) rigidity of the nominal exchange rate regime; and (5) external debt (in percent of GDP).
- *Fiscal sector:* (1) overall balance; (2) primary balance, including the gap between primary balance and debt-stabilizing primary balance; (3) public debt (in percent of GDP); (4) maturity of public debt; and (5) foreign currency debt in percent of total debt.
- *Financial sector:* (1) capital adequacy; (2) return on assets; (3) nonperforming loans as a share of total loans; (4) growth in private sector credit (as a ratio to GDP); and (5) the share of foreign currency loans.
- *Corporate sector:* (1) default probability (extracted from a Black-Scholes-Merton formula); (2) interest coverage ratio; (3) debt-to-assets ratio; (4) real return on assets; and (5) a valuation measure based on the price-to-earnings ratio.

Whenever data coverage was incomplete, we used close substitutes of these indicators. For example, we used only short-term debt as opposed to short-term debt plus maturing medium- and long-term debt in computing reserve cover. We also constructed a number of alternative measures of financial sector soundness from Boyd, De Nicoló, and Jalal (2006). Note that, as previously discussed, the nature of BCTs is such that including additional

---

[11]IMF's country desk economists provided the historical data going back to 1994 that were necessary to construct vulnerability indicators of the external and fiscal sectors. IMF's monetary and capital markets department provided most financial sector data (with measures of capital adequacy and nonperforming loans beginning in 2000), and Boyd, De Nicoló, and Jalal (2006) provided data (extracted from BankScope) on return on assets, equity-to-asset ratio, and loan-to-asset ratio. The corporate vulnerability utility team provided corporate sector indicators.

variables with limited explanatory power does not change the results (unlike in a regression, where degrees of freedom would be affected).

Country-specific measures complemented these sectoral indicators:

- *Macroeconomic conditions:* One-year-ahead WEO forecasts of (1) real GDP growth and (2) CPI inflation.
- *Global demand conditions:* (1) One-year-ahead WEO forecasts of growth in import demand by trading partners and (2) levels and changes of commodity price indices faced by each particular country.[12] Both measures are country-specific.
- *History of defaults:* (1) Number of sovereign defaults since 1900 and since 1950 and (2) share of time under a sovereign default since 1900 and since 1950.
- *EMBI spreads.*

We did *not* include country-invariant global macroeconomic and capital markets conditions (for example, global growth or U.S. interest rates) because they could end up playing the role of yearly dummies.[13] If included, however, they would not show up in any tree.[14] Given that predicting capital account crises was the main goal of our exercise, we used lagged values for all variables (for example, indebtedness at time $t-1$ to predict a crisis at $t$). Moreover, lagged values are more likely to convey useful information; contemporaneous ones would be affected by the inception of a crisis (for example, low levels of reserves could be a consequence of a crisis rather than one of the underlying vulnerabilities that allowed a crisis to happen). The only classification tree that considers contemporaneous variables is the one discussed in Section V, where contemporaneous global conditions are used to compare country-specific vulnerabilities under favorable and unfavorable global scenarios. It is worth noting that 30 out of 34 crises in our sample are concentrated in six years: 1994, 1997, 1998, 1999, 2001, and 2002. When presenting the results, we briefly discuss the implications of including a dummy variable for those crisis-prone years, which would act as a rough measure of global financial conditions (under perfect foresight).

We also considered political economy–related indicators. These indicators tended to have very limited explanatory power, possibly because they adjusted sluggishly, with abrupt movements in political variables often occurring after a crisis.

---

[12]The IMF Research Department Commodities Unit constructed these data.

[13]For example, there is a large concentration of crises in the late 1990s, when oil prices were relatively low. In some preliminary versions with an oil price indicator, low oil prices seemed to be harmful for the average country in the sample, most likely because of the association between cheap oil and crises in the late 1990s.

[14]It is possible that the lack of cross-country variation adversely affected EMBI spreads' explanatory power in BCTs. Calvo, Leiderman, and Reinhart (1992) found that push factors, such as international interest rates and the U.S. business cycle, explained part of capital flows to Latin America in the early 1990s.

The forecasting nature of the exercise also requires that any ex post measure of RER overvaluation be excluded as long as it uses unavailable future information to compute real exchange trends. To overcome this problem, we experimented with a number of possible approaches to estimating overvaluation at time $t$ using only information available up to that period (rolling RER trends). However, those estimates turned out to be very noisy and not informative. For example, a sound economic expansion with rapid productivity growth resulted in an appreciating trend of the RER just like that of a country teetering on the brink of crisis. Oftentimes, the extrapolation of a trend following a large depreciation suggested that the currency was overvalued, even when the RER was broadly in line with its equilibrium value (or had overshot it). Using rolling Hodrick-Prescott filters instead of rolling linear trends did not improve matters much. In the end, our preferred method for determining the level of overvaluation without using ex post information was to compare the RER with its long-term average since 1960 (where data availability permitted).

The exchange rate regime (for example, Reinhart and Rogoff's (2004) de facto classification) did not have much predictive power either. This result may be partly due to the fact that, even in countries that pegged the exchange rate and experienced a crisis, no crisis took place during most of the peg years, thus diluting the positive relationship between crisis observations and exchange rate rigidity. Also, the EMBI spreads never appeared in any BCT, which is somewhat surprising because in principle they should be a forward-looking market-based indicator; their limited sample coverage prior to the late 1990s may contribute to this result.

### III. A Baseline Tree That Explains In-Sample Crises

#### Baseline Tree

Figure 1 shows the baseline tree. The BCT algorithm used only variables dated one year prior to that of the outcome we were trying to predict (crisis and noncrisis). The variable that best split the sample into crisis-prone and non-crisis-prone observations was the lagged reserve cover, measured as the ratio of gross international reserves to the sum of the short-term external debt (from the Bank for International Settlements) and the current account deficit (set to zero if a surplus). For example, a reserve cover of 100 percent would allow a country to finance its entire current account deficit plus all short-term external debt maturing within a year by bringing its stock of international reserves to zero. The BCT algorithm selected a threshold of 81 percent, which partitioned the sample into 164 crisis-prone observations with a lower lagged reserve cover (of which 23, or 14 percent, were crises; top-left branch) and 390 observations with a higher lagged reserve cover (of which 11, or 2.8 percent, were crises; top-right branch).

The dominant role of reserve cover in predicting capital account crises is very robust. All the BCTs presented in this paper had either lagged reserve cover or its components (the lagged current account balance and the ratio of short-

Figure 1. Binary Classification Tree Based on 1994–2005 Sample and Crisis Episodes

```
                          ┌─────────────────────────┐
                          │        Sample           │
                          │ 554 obs. o/w 6.1% crisis │
                          └─────────────────────────┘
              Reserve Cover ≤ 81%              Reserve Cover > 81%
        ┌─────────────────────────┐      ┌─────────────────────────┐
        │ 164 obs. o/w 14% crisis  │      │ 390 obs. o/w 2.8% crisis │
        └─────────────────────────┘      └─────────────────────────┘
                                     WEO Forecasted Real GDP    WEO Forecasted Real GDP
    External Debt/GDP ≤ 24%            Growth ≤ 3%                Growth > 3%
         External Debt/GDP > 24%
┌──────────────────────┐  ┌────────────────────────┐
│30 obs. o/w 0% crisis │  │134 obs. o/w 17.2% crisis│
└──────────────────────┘  └────────────────────────┘
```

| 54 obs. o/w 13% crisis | 336 obs. o/w 1.2% crisis |
|---|---|
| Colombia   2002 | Israel 1997, 2002 |
| Venezuela  1994, 2001 | Malaysia 1997 |
| Lebanon   2001 | Czech Rep. 1997 |
| Bulgaria    1994 | |
| Ukraine    1994 | |
| Hungary   1996 | |

Change in External Debt/GDP ≤ –3.3%          Change in External Debt/GDP > –3.3%

| 26 obs. o/w 0% crisis |
|---|

| 108 obs. o/w 21.3% crisis | |
|---|---|
| Algeria 1994 | Mexico 1994 |
| Argentina 1995, 2001 | Pakistan 1998 |
| Brazil 1998, 2002 | Philippines 1997 |
| Colombia 1999 | Romania 1999 |
| Dominican Rep. 2003 | Russia 1998 |
| Ecuador 1999 | South Africa 2001 |
| Indonesia 1997 | Thailand 1997 |
| Jamaica 2003 | Turkey 1994, 2001 |
| Korea 1997 | Ukraine 1998 |
| Lithuania 1999 | Uruguay 2002 |

Notes: All variables used are lagged, corresponding to the value in the previous year. Reserve cover is the ratio of gross international reserves to the sum of the short-term external debt (from the Bank for International Settlements) and the current account deficit (zero if it indicates a surplus).

term external debt to reserves) at the top. It is not surprising that countries can forestall capital account crises by accumulating large stocks of international reserves, containing current account deficits, and limiting short-term debt. What is new, however, is how well simple thresholds on the values of these variables, or for the reserve coverage measure that combines them, can forecast crises.

The share of countries with reserve cover above the estimated safe threshold rose from about 40 percent in 1994 to 80 percent in 2005. This increase is consistent with the lack of crises in recent years. Table 2 shows that, since 2002, between one-third and one-half of the countries in the sample have had a reserve cover ratio above 200 percent, which is well in excess of the 81 percent threshold selected by BCT. This suggests that motives for reserve accumulation other than crisis prevention might be at play.

Countries with a low reserve cover are not necessarily doomed. In our sample, only one in seven countries with a reserve cover below 81 percent experienced a crisis. The level of external debt in relation to GDP (left branch of the tree in Figure 1) helped to sharpen crisis prediction in instances of low reserve cover. When lagged external debt was below 24 percent of GDP, no crisis took place even

| Table 2. Number of Countries by Reserve Cover Range over Time | | | | | |
|---|---|---|---|---|---|
| | Reserve Cover | | | | |
| Year | 0–0.5 | 0.5–0.812 | 0.812–1 | 1–2 | >2 |
| 2000 | 7 | 11 | 3 | 15 | 13 |
| 2001 | 2 | 9 | 8 | 16 | 14 |
| 2002 | 6 | 6 | 5 | 15 | 17 |
| 2003 | 3 | 7 | 2 | 17 | 20 |
| 2004 | 4 | 5 | 6 | 10 | 24 |
| 2005 | 3 | 6 | 2 | 21 | 17 |

Note: Reserve cover is defined as the ratio of gross international reserves to the sum of the short-term external debt (from the Bank for International Settlements) and the current account deficit (zero if it is a surplus).

though reserve cover was below the threshold. The few low-reserve-cover countries in this situation (30 out of 164) had a relatively high fraction of short-term debt or sizable current account deficits but incurred no crisis. Conversely, when external debt was above 24 percent of GDP, the frequency of crises among low-reserve-cover countries rose to 17.2 percent (one in six). It is worth noting that, although this 24 percent threshold may seem low, it was based on only external debt as opposed to the entire stock of public debt (which would also include domestic public debt).[15] At the same time, our external debt measure included the external debt of the private sector. As a result, this split captured an external sector vulnerability rather than fiscal vulnerability.

In our sample, countries with low reserve cover and high external debt could still have escaped a crisis if external debt had fallen by at least 3.3 percent of GDP in the previous year. Smaller reductions or increases in the external debt-to-GDP ratio isolated, instead, a crisis-prone group of 108 observations with 23 crises (21.3 percent or one in five).

Returning to the top of the tree in Figure 1 and moving down the right branch, we notice that a high reserve cover needs to be combined with a strong growth outlook to shield countries from capital account crises (that is, to reduce the crisis frequency to about 1 percent). By contrast, when the previous-year WEO real growth forecast was below 3 percent, crises took place with a frequency of 13 percent (seven crises out of 54 observations) even at high levels of reserve cover. Although 3 percent may look like a relatively high threshold for GDP growth, as many as 336 observations ended up in the "safe" node with a higher forecasted GDP growth. This pattern reflects the relatively high growth rates of emerging market countries: in our sample, the first quartile of the distribution of real growth forecasts was as high as 3.5 percent (Table 3).

---

[15]Reinhart, Rogoff and Savastano (2003) also found a "safe" threshold for the external debt-to-gross national product ratio as low as 15 percent for some developing countries.

The in-sample fit of the baseline tree is very good. It predicted correctly 30 out of 34 crises (88.2 percent) and wrongly predicted 132 crises out of 520 noncrisis observations (a misclassification rate of 25.4 percent). The optimal tree based on the V-fold cross-validation technique would have had 30 terminal nodes. This alternative tree would have predicted all the crises, and misclassified only 3.7 percent of noncrisis observations.

If we included the dummy for crisis-prone years (under perfect foresight) discussed in Section II, that indicator would appear as the top split in our tree. The risky branch of that split (the crisis-prone years) would grow similarly to how our baseline tree grew. That tree would miss even more crises (eight instead of four) but it would have a lower noncrisis misclassification rate (12.50 percent). Depending on how we value this trade-off, such a tree may be preferable. The gains from including the proxy for the global financial environment would, however, be small.

It is useful to compare our estimates with those of EWS models. Two EWS models estimated on the period 1985–97: the Developing Country Studies Division (DCSD) and the Kaminsky, Lizondo and Reinhart (KLR) models considered in Table 4 of Berg, Borensztein, and Pattillo (2005) had a poorer in-sample prediction rate (63 and 60 percent, respectively). However, their false alarms as a percentage of total alarms were lower than those of the BCT of Figure 1 (64 and 71 percent as opposed to 81 percent). The different

| Table 3. Descriptive Statistics of Key Indicators | | | | | | |
|---|---|---|---|---|---|---|
| Variable | Number of Observations | Mean | Standard Deviation | 25th Percentile | Median | 75th Percentile |
| Reserve cover[1] | 521 | 1.836 | 2.344 | 0.691 | 1.133 | 2.188 |
| External debt/GDP | 496 | 48.194 | 22.999 | 32.549 | 45.948 | 59.919 |
| Change in external debt/GDP | 448 | −0.179 | 9.715 | −3.578 | −0.268 | 2.735 |
| WEO real GDP growth forecast | 526 | 4.366 | 1.910 | 3.500 | 4.500 | 5.463 |
| Current account/GDP | 544 | −0.024 | 0.061 | −0.052 | −0.026 | 0.002 |
| Short-term debt/reserves | 522 | 1.027 | 2.053 | 0.321 | 0.556 | 0.980 |
| Fiscal balance/GDP | 495 | −3.375 | 3.835 | −5.195 | −3.034 | −1.177 |
| Exchange rate misalignment[2] | 546 | 0.522 | 23.646 | −15.284 | 0.782 | 17.829 |
| Deviation of commodity prices from past average[3] | 542 | −9.084 | 22.914 | −21.981 | −12.982 | −4.039 |
| Change in real import demand by trading partners[3,4] | 542 | 8.817 | 4.841 | 5.673 | 9.269 | 12.017 |

[1]Reserve cover is defined as the ratio of gross international reserves to the sum of the short-term external debt (from the Bank for International Settlements) and the current account deficit (zero if it indicates a surplus).
[2]Misalignment relative to average real effective exchange rate from 1960 to previous year.
[3]Contemporaneous values.
[4]Excludes oil.

frequency of the data (monthly in the case of EWS) and sample periods suggest that these comparisons should be interpreted with caution.

## The Importance of Different Classes of Indicators

The robustness of reserve cover as crisis indicator is confirmed by the fact that the main competitors for the top split are its components (the ratio of short-term external debt to international reserves and the current account-to-GDP ratio) or close substitutes of its components (the WEO forecast of the current account-to-GDP ratio and the ratio of short-term external debt to GDP). Other competitors are indicators that appear further down the baseline tree, such as the change in the external debt-to-GDP ratio and the WEO growth forecast. There are also no surprises among the competitors of the second-level indicators, with the change in the government debt-to-GDP ratio emerging as a possible competitor of the external debt indicators and different WEO vintages of GDP growth forecasts as competitors for the right-hand split.

The lack of an exchange rate overvaluation measure in the baseline tree may look surprising in view of the prominent role this variable played in EWS. This result may reflect, however, the similar information content of current account balances and exchange rate overvaluation measures. The inherent difficulty in constructing an ex ante indicator of overvaluation using only precrisis information can also explain why our simple overvaluation measure—which compares the RER with its past long-term average—turns out to have less information content than do current account balances. Nonetheless, our overvaluation measure is in a second group of competitors for the top split (ranking between fifth and tenth) and emerges as a second-level splitter in the BCT estimated on the subsample that excludes East Asian countries (see Section IV). Finally, a mixed alternative overvaluation measure—in which the simple deviation of the RER from its long-term average is replaced with its deviation from the equilibrium RER computed according to the IMF's Consultative Group on Exchange Rate Issues (CGER) methodology for all countries in the sample for which the latter is available—would feature as a second-level splitter in the baseline tree in place of external debt.[16]

These results shed some light on whether traditional domestic macroeconomic variablessuch as growth, RERs, current account deficits, international reserves, and fiscal variablescontain enough information to

---

[16]We did not use the mixed exchange rate overvaluation measure as our main overvaluation measure because the equilibrium real exchange rate is computed as a function of variables, such as net foreign assets, relative productivity growth in the traded and nontraded goods sectors, and terms of trade, using parameters that are estimated over the 1973–2004 period and, therefore, estimated on ex post information for most of our sample. At the same time, relying on such parameters does not create as many problems as proxying equilibrium real exchange rates with country-specific trends because the parameter estimates are not country-specific but are panel estimates, which are identical for all CGER countries. Moreover, we computed the only country-specific parameter (the fixed effect in the equilibrium real exchange rate equation) in a rolling fashion using only ex ante information.

predict future crises, or whether microeconomic indicators of imbalances in financial and corporate sectors are needed to improve predictability. In the baseline tree, traditional macroeconomic variables trump financial and corporate sector indicators despite the rich set of candidates for the latter that the BCT algorithm took into consideration (see Section II). We also verified that this result was not due to the larger number of missing observations that characterize some financial and corporate sector indicators, by re-running BCTs on subsets of observations for which measures of capital adequacy, return on assets, corporate debt-to-asset ratios, and the EMBI index were not missing. In all these instances, only macroeconomic indicators still showed up in the BCTs. The little information content of financial and corporate indicators may, then, reflect the lag with which balance sheet data record financial and corporate vulnerabilities. Furthermore, we suspect such vulnerabilities play a major role in determining how disruptive capital account crises can be but may play a more limited role in determining whether a capital account crisis takes place to begin with.

## IV. Out-of-Sample Forecasts

The good in-sample fit of the baseline tree is encouraging but does not actually answer the question in the title of this paper. To have a clue about BCTs' ability to predict crises, we had to consider out-of-sample forecasts. For this section, we estimated BCTs using data up to 2000, 2001, and 2002 to predict crises in, respectively, 2001, 2002, and 2003. We focused on these years because they were the last with crises in our sample. We also present out-of-sample forecasts for East Asian countries based on a sample that excludes them.

Larger trees improve the in-sample fit but may actually worsen out-of-sample performance (as is the case with standard regressions). The V-fold cross-validation technique described in Section I suggests an optimal "pruning" of trees, which we often override using our judgment. The first reason for overriding is that the V-fold cross-validation technique often recommends an uninformative tree with no splits or with too many splits, including some based on statistical correlations that make no economic sense. Second, the tree size preferred by the V-fold cross-validation technique is based on the out-of-sample performance of a set of trees that—having been estimated on random subsamples of data—might have little or no resemblance to the tree whose out-of-sample performance we want to assess. Third, the V-fold cross-validation technique may lead to "overfitting."[17] Despite these reservations, to be fully transparent about our methodology, we always report the size that the V-fold cross-validation

---

[17]Consider a V-fold simulation in which the observation for Indonesia 1997 is randomly selected for out-of-sample testing whereas the observation for Thailand 1997 is used in sample. Because those two crises shared similar features, the estimated tree would choose a set of indicators that can predict Thailand 1997 quite well (and, therefore, probably also Indonesia 1997) but it would be as though we had known ahead of time that a crisis was going to take place in Thailand in 1997.

would recommend and the associated misclassification rates, together with those of our preferred tree.

In discussing each out-of-sample forecast, we also report out-of-sample results based on the RandomForests algorithm. Despite the drawback of not yielding specific rules and thresholds, the forecasting ability of the RandomForests algorithm is a benchmark against which we can compare that of the out-of-sample trees pruned using our judgment. Overall, the RandomForests algorithm did not yield forecasts consistently superior to those of our trees.

## Out-of-Sample Prediction of 2001 Crises

Figure 2 shows the tree estimated on data up to 2000 and the nodes in which each country was predicted to end up in 2001. The V-fold cross-validation suggested a tree with no splits. We chose to grow the left branch of the tree and have three terminal nodes because the additional split made economic sense and raised the crisis frequency from 12.5 to 16.1 percent.

The variable that best split the 1994–2000 sample into crisis-prone and non-crisis-prone observations was the current account balance over GDP. The BCT algorithm selected a threshold of −2.9 percent of GDP, which partitioned the sample into 168 crisis-prone observations with a lower current account balance (of which 21, or 12.5 percent, were crises; top-left branch) and 152 observations with a higher current account balance (of which 2, or 1.3 percent, were crises; top-right branch). The ratio of short-term external debt to reserves (left branch of the tree in Figure 2) further split the crisis-prone node. When this ratio was below 41 percent, crises were relatively rare (one crisis out of 44 observations, or 2.3 percent). High ratios of short-term debt to reserves—combined with large current account deficits—raised, instead, the frequency of crises to 16.1 percent (20 crises out of 124 observations), characterizing the crisis-prone node of this tree. What is interesting is that the two most informative indicators selected using the 1994–2000 sample were the two components of the reserve cover ratio used in the first split of the baseline tree.

The tree in Figure 2 would have successfully predicted the crises in Argentina, Lebanon, and Turkey, but it would have failed to predict those in South Africa and Venezuela. Although an error of 40 percent can be seen as high, we find it reassuring that it would have predicted the major crises.[18] The false positives corresponded to 33 percent of noncrisis observations, which was reasonable given the nature of the exercise and the fact that we wanted to err on the side of being conservative. It is worth noting that one of the false positives (Brazil) had a crisis in 2002 and two (Dominican Republic and Jamaica) had a crisis in 2003.

---

[18]The forecasting performance would not have improved if we had included a dummy for crisis-prone years as a measure of the global financial environment. In fact, such a tree would have made us miss the crisis in Lebanon.

Figure 2. Binary Classification Tree Based on 1994–2000 and Out-of-Sample Predictions for 2001
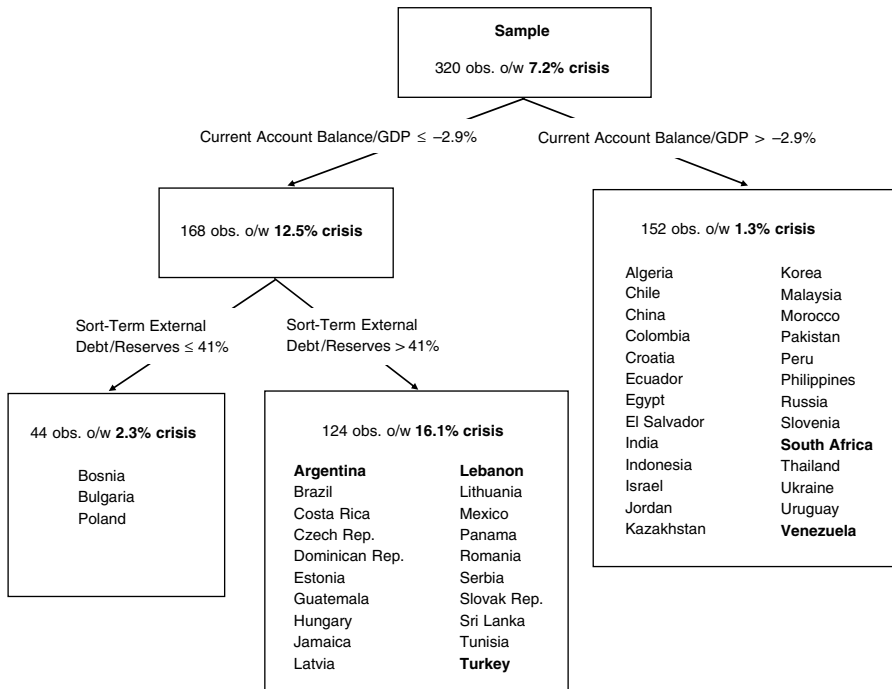


Notes: All variables used are lagged, corresponding to the value in the previous year. Sample frequencies reported correspond to in-sample values for 1994–2000. Countries are listed based on their out-of-sample classification for 2001, with crisis episodes in bold.

These out-of-sample results cannot be easily compared with those of the EWS models considered by Berg, Borensztein, and Pattillo (2005), who tested the DCSD and KLR models on the out-of-sample period from January 1999 to December 2000, which includes only three crises (Brazil, Colombia, and Zimbabwe) and excludes all 2001–03 crises, including Argentina and Turkey. In these EWS modes, the percentage of crises correctly called was measured as the number of observations for which the estimated probability of crisis was above the cutoff probability and a crisis ensued within 24 months, as a share of all observations for which a crisis ensued within 24 months. Using this measure, Berg, Borensztein, and Pattillo found that DCSD correctly predicted 31 percent of the precrisis months whereas KLR correctly predicted 58 percent of precrisis months. These results are comparable to the 60 percent out-of-sample prediction rate of the BCT in Figure 2 (three out of five crises).

The RandomForests algorithm also predicted the crises in Argentina, Lebanon, and Turkey and missed those in South Africa and Venezuela, but issued more false alarms (misclassifying 50 percent of noncrises). Thus,

for this out-of-sample exercise, the forecasting performance of the RandomForests algorithm was worse than that of the tree in Figure 2.

## Out-of-Sample Prediction of 2002 Crises

Figure 3 repeats the same exercise, this time estimating a tree on the 1994–2001 sample to predict crises in 2002. The V-fold cross-validation again suggested a tree with no splits. We chose instead a tree with the same top split as the 1994–2000 tree based on the current account balance. This split failed to predict the crises in Colombia, Israel, and Uruguay, all of which had a lagged current account balance above the threshold of −2.9 percent. The precrisis large current account deficit of Brazil placed it, instead, in the top-left node with a crisis probability of 12.6 percent. The splits based on the ratio of short-term external debt to reserves and the previous-year WEO real growth forecast, which did a good job in isolating crisis observations in sample, would have, however, misplaced Brazil in a relatively safe node with a crisis frequency of only 4.4 percent.

To put this result in perspective, we believe it worth mentioning that the tree estimated on the 1994–2001 sample raised the threshold on the ratio of short-term external debt to reserves from the 41 percent level of the 1994–2000 tree to 125 percent. This change considerably improved the in-sample prediction (with the crisis frequency in the crisis-prone node rising to 33 percent) but it made us miss the crisis in Brazil, which had a ratio of short-term external debt to reserves of 119 percent. Moreover, the crises in Colombia and Uruguay had a strong contagion component for which we lacked a proper indicator, and the crisis in Israel was to a large extent related to non-economic considerations. The false alarms corresponded to only 2.5 percent of the noncrisis observations.[19]

The RandomForests algorithm predicted the crisis in Brazil and missed those in Colombia, Israel, and Uruguay, with a crisis misclassification rate of 75 percent and a noncrisis misclassification rate of 25 percent. Given our preference to err on the side of caution, we would have preferred this performance to that of the tree in Figure 3.

## Out-of-Sample Prediction of 2003 Crises

Figure 4 shows the tree estimated on data up to 2002 and the nodes in which each country was predicted to end up in 2003. The V-fold cross-validation again suggested a tree with no splits. The tree in Figure 4 has the same left branch of the baseline tree estimated on the full sample, differing from it only in the lack of the split based on the previous-year WEO real growth forecast on the right branch. This tree perfectly predicted the two crises of 2003

---

[19]The forecasting performance would have improved if we had included a dummy for crisis-prone years as a measure of the global financial environment. Such a tree would have predicted the crisis in Brazil but would still have missed the remaining three crises.

**Figure 3. Binary Classification Tree Based on 1994–2001 and Out-of-Sample Predictions for 2002**

```
                              Sample
                      369 obs. o/w 7.6% crisis

   Current Account Balance/GDP ≤ −2.9%    Current Account Balance/GDP > −2.9%

        191 obs. o/w 12.6% crisis                 178 obs. o/w 2.2% crisis

  Short-Term External        Sort-Term External
  Debt/Reserves ≤ 125%       Debt/Reserves > 125%
```

| 178 obs. o/w **2.2% crisis** | |
| --- | --- |
| Algeria | Morocco |
| Chile | Pakistan |
| China | Panama |
| **Colombia** | Peru |
| Egypt | Philippines |
| El Salvador | Poland |
| India | Russia |
| Indonesia | Slovenia |
| **Israel** | Sri Lanka |
| Jordan | Thailand |
| Korea | Ukraine |
| Malaysia | **Uruguay** |
| Mexico | |

```
        158 obs. o/w 8.2% crisis          33 obs. o/w 33.3% crisis
                                                  Estonia

  WEO Forecasted Real GDP    WEO Forecasted Real GDP
      Growth ≤ 2.5%              Growth > 2.5%

   21 obs. o/w 33.3% crisis
```

| 137 obs. **o/w 4.4% crisis** | |
| --- | --- |
| Bosnia | Hungary |
| **Brazil** | Jamaica |
| Bulgaria | Kazakhstan |
| Costa Rica | Latvia |
| Croatia | Lithuania |
| Czech Rep. | Romania |
| Dominican Rep. | Serbia |
| Ecuador | Slovak Rep. |
| Guatemala | Tunisia |

Notes: All variables used are lagged, corresponding to the value in the previous year. Sample frequencies reported correspond to in-sample values for 1994–2001. Countries are listed based on their out-of-sample classification for 2002, with crisis episodes in bold. Countries that experienced a crisis in 2001 are excluded.
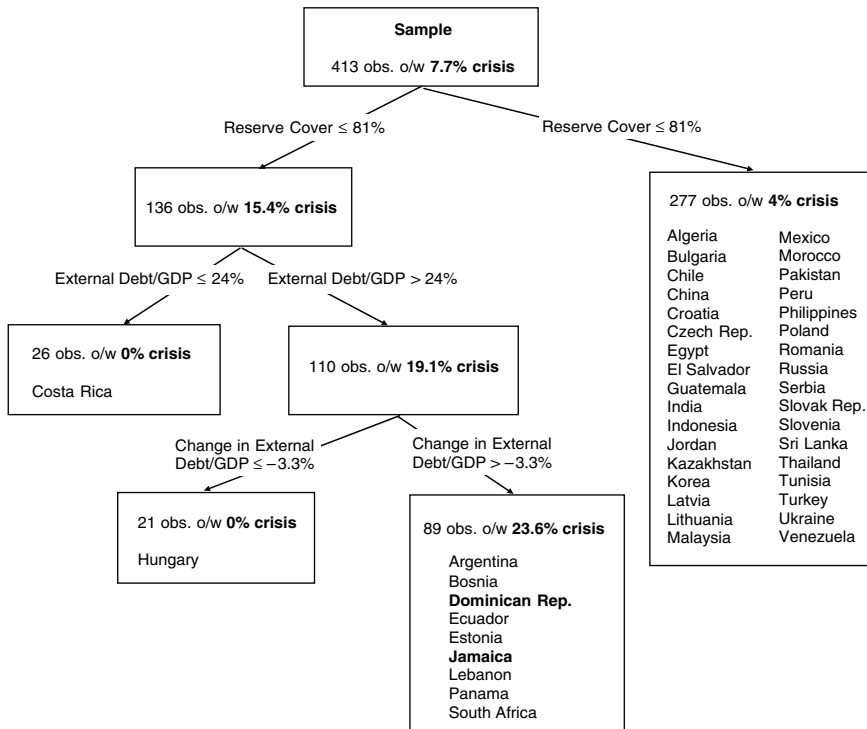
(Dominican Republic and Jamaica). False alarms corresponded to 16 percent of the noncrisis observations.

The RandomForests algorithm predicted the crisis in Jamaica and missed the one in the Dominican Republic (thus a 50 percent misclassification of crises), and misclassified 28 percent of the noncrisis observations. In this case, the performance of the RandomForests algorithm was worse than that of the tree in Figure 4.

## Out-of-Sample Prediction of the 1997 East Asian Crisis

Figure 5 estimates a tree based on a sample that excludes all observations corresponding to East Asian countries (China, Indonesia, Korea, Malaysia, Philippines, and Thailand) over the 1994–2005 period and checks how well it would have predicted the 1997 crisis. This was not a perfect out-of-sample

Figure 4. Binary Classification Tree Based on 1994–2002 and Out-of-Sample Predictions for 2003

**Sample**

413 obs. o/w **7.7% crisis**

Reserve Cover ≤ 81%          Reserve Cover ≤ 81%

136 obs. o/w **15.4% crisis**

277 obs. o/w **4% crisis**

| | |
|---|---|
| Algeria | Mexico |
| Bulgaria | Morocco |
| Chile | Pakistan |
| China | Peru |
| Croatia | Philippines |
| Czech Rep. | Poland |
| Egypt | Romania |
| El Salvador | Russia |
| Guatemala | Serbia |
| India | Slovak Rep. |
| Indonesia | Slovenia |
| Jordan | Sri Lanka |
| Kazakhstan | Thailand |
| Korea | Tunisia |
| Latvia | Turkey |
| Lithuania | Ukraine |
| Malaysia | Venezuela |

External Debt/GDP ≤ 24%     External Debt/GDP > 24%

26 obs. o/w **0% crisis**

Costa Rica

110 obs. o/w **19.1% crisis**

Change in External          Change in External
Debt/GDP ≤ −3.3%            Debt/GDP > −3.3%

21 obs. o/w **0% crisis**

Hungary

89 obs. o/w **23.6% crisis**

Argentina
Bosnia
**Dominican Rep.**
Ecuador
Estonia
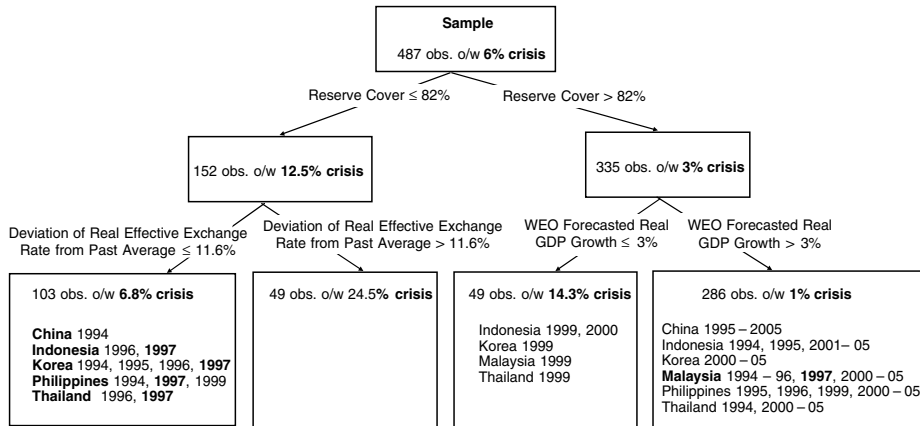**Jamaica**
Lebanon
Panama
South Africa

Notes: All variables used are lagged, corresponding to the value in the previous year. Sample frequencies reported correspond to in-sample values for 1994–2002. Countries are listed based on their out-of-sample classification for 2003, with crisis episodes in bold. Countries that experienced a crisis in 2002 are excluded.

test because it included post-1997 information from non-East Asian countries that was unavailable at the time. Nonetheless, it was our only option, considering that estimating a BCT on the short 1994–96 sample would not have been meaningful. The V-fold cross-validation suggested a tree with no splits. Instead, we chose one with four.

The top split in this tree is the same as in the baseline tree (reserve cover at 82 percent). Given their high short-term external debt in relation to reserves prior to 1997, all East Asian countries except Malaysia would have ended up in the top-left node with a crisis frequency of 12.5 percent. However, the second-level split of the tree would have erroneously classified Indonesia, Korea, Philippines, and Thailand as non-crisis-prone in 1997. For these countries, in fact, there was no sign—based on our measure—of RER overvaluation in the year prior to the crisis (that is, the RER was less than 12 percent above its country-specific average from 1960 to the previous year). The first-ranked competitor of exchange rate misalignment (the government overall balance-to-GDP ratio) would have

Figure 5. Binary Classification Tree Based on 1994–2005, Excluding East Asia and Out-of-Sample Predictions for East Asia



Notes: All variables used are lagged, corresponding to the value in the previous year. Sample frequencies reported correspond to in-sample values for 1994–2005, excluding East Asia. East Asian observations are listed based on their out-of-sample classification, with crisis episodes in bold. Countries that experienced a crisis in previous year are excluded.

also mispredicted the crisis in these four East Asian countries on the heels of their strong precrisis fiscal position. These results highlight the fact that, in the subsample without East Asia on which the tree is estimated, crises were unlikely to occur in countries where the reserve cover was low but the RER was not misaligned or the fiscal position was strong. To predict correctly the crises in Indonesia, Korea, Philippines, and Thailand, we would have needed to use the second-ranked competitor of the exchange rate overvaluation measure (the lagged level of external debt). The estimated tree also failed to predict Malaysia because its reserve cover was relatively high and the WEO real growth forecast for 1997 was above 3 percent.

In sum, if we had stopped at the first split, a tree estimated on a sample without East Asia would have correctly predicted four crises (Indonesia, Korea, Philippines, and Thailand) and missed only Malaysia (only a 20 percent crisis misclassification rate) and had false alarms for 15 percent of the noncrisis observations. Using, instead, the entire tree, we missed all crises and misclassified 7.5 percent of the noncrisis observations. The V-fold cross-validation suggested no splits in the tree, classifying all observations as noncrises.[20]

---

[20]The forecasting performance would have improved if we had included a dummy for crisis-prone years as a measure of the global financial environment. Such a tree would have predicted the crises in Indonesia, Korea, and Thailand (but would still have missed the ones in Malaysia and the Philippines).

Berg, Borensztein, and Pattillo (2005, Table 4) verified how EWS models would have predicted the East Asian crisis out of sample by estimating the DCSD and KLR models on monthly data over the period December 1985–April 1995 and using them to check for how many months over the period May 1995–December 1996 the estimated probability of crisis would have been above the cutoff probability for the East Asian countries that experienced a crisis in 1997. As in previous instances, the comparison with the BCT results is difficult because the crisis definitions do not match (for example, EWS models do not consider the Philippines a crisis country). Moreover, Berg, Borensztein, and Pattillo ran a proper out-of-sample exercise based only on precrisis information, whereas we used postcrisis experience in other countries to estimate the BCT in Figure 5. In this case, the out-of-sample performance of the EWS was quite good, with the percentage of crises correctly called in 24 months at 84 percent for DCSD and 75 percent with KLR. A BCT based only on the first split would have yielded similar results with a prediction rate of 80 percent (four crises out five), whereas the entire tree of Figure 5 would have been much inferior to EWS models.

The RandomForests algorithm predicted the crisis in Korea and missed all four other East Asian crises, with a crisis misclassification rate of 80 percent and a noncrisis misclassification rate of 29 percent. This performance is marginally preferable to that of the entire tree in Figure 5.

## V. Global Conditions vs. Country-Specific Indicators

This section addresses the question of the role of global economic conditions in capital account crises. So far, we focused on predicting crises and considered only lagged values of variables to study their leading indicator properties and to avoid endogeneity problems caused by contemporaneous domestic indicators (for example, an association between low contemporaneous reserve cover and crises would be no evidence of the indicator role of reserve cover because reserves typically drop during crises). We now include contemporaneous values of measures of global conditions that each country faces, such as an export-weighted index of real commodity prices and an index of import demand by trading partners, which are exogenous to contemporaneous developments in individual countries. By isolating subsets of observations with a higher frequency of crises than in BCTs based only on lagged indicators, we found that these indicators played an important role in improving the in-sample classification.

Figure 6 shows a variant of the baseline tree of Figure 1 that allows for contemporaneous global indicators. Lagged reserve cover remained the most important variable with an unchanged threshold of 81 percent. However, on the left (risky) branch of the tree, the real level of commodity export prices replaced external debt over GDP in splitting observations with low reserve cover. When the real level of commodity export prices was more than 14

**Figure 6. Binary Classification Tree Based on 1994–2005, Including Contemporaneous Global Demand Variables**
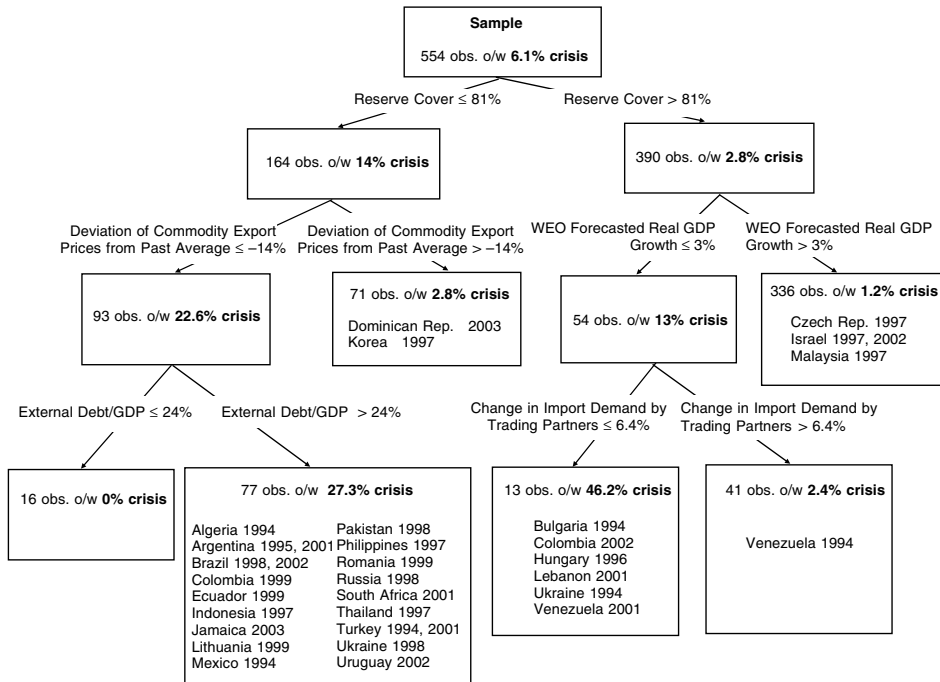


Notes: All variables used are lagged, except for the ones for real commodity prices and import demand which are contemporaneous. Import demand by trading partners indicates the change in the import volume of goods excluding oil.

percent below its past country-specific average, 22.6 percent of observations were crises (21 out of 93), whereas only 2.8 percent of observations were crises (2 out of 71 when commodity export prices were above this threshold). A level of external debt to GDP above 24 percent further raised the frequency of crises to 27.3 percent.

Returning to the top of the tree and moving down the right (safe) branch characterized by high reserve cover, we find the same split of the baseline tree based on the previous-year WEO real growth forecast. But the contemporaneous growth in (nonoil) import demand by trading partners allowed us to further split the node with a weak growth outlook. A strong contemporaneous growth in import demand by trading partners (above 6.4 percent) could offset the impact of the low growth outlook forecasted in the previous year, reducing the frequency of crises to 2.4 percent (one crisis out of 41 observations). On the other hand, if import demand by trading partners is weak, the low growth outlook would translate into a very high frequency of crises equal to 46.2 percent (six out of 13 observations).

The in-sample fit of this tree is good. We failed to predict only seven crises out of 34 (an error of 20.6 percent) and we wrongly predicted 62 crises out of 520 noncrisis observations (an error of 11.9 percent). Therefore, false alarms in percentage of total alarms dropped to 69.7 percent from the 81.5 percent of the BCT in Figure 1. Misclassification rates would not have improved if we had included a dummy for crisis-prone years as a measure of global financial environment.

## VI. Conclusions

This paper described the use of BCTs to study the in-sample and out-of-sample forecasting properties of a large set of indicators of capital account crises. Although BCTs were previously used in few studies of currency and sovereign debt crises, this was the first application of BCTs to capital account crises (''sudden stops''). Our results shed light on the relative importance of leading indicators of crises and on their interaction, suggesting that BCTs could be a useful complement to existing crisis prediction methods. Our estimates compared well with previous EWS, resulting in superior in-sample prediction. The out-of-sample comparison was more difficult because of the different crisis definition (capital account vs. currency crises) and data frequency (annual vs. monthly data), but BCTs seemed to do just as well as EWS out of sample or even better, 2002 aside.

The interaction of international reserves, current account balances, and short-term external debt constituted the backbone of the BCTs presented in this paper. The evidence supporting the leading indicator role of these three variables was robust. A measure of reserve cover that combined them (the lagged ratio of international reserves to the sum of current account deficits and short-term external debt) was the best splitter of the full sample into crisis and non-crisis-prone observations, whether we allowed for contemporaneous global cyclical conditions or not. In some instances, splits based only on reserve cover or its components led to better out-of-sample prediction performance than did fully fledged trees. Lagged reserve cover was the preferred leading indicator of crises also in a subsample that excluded the East Asian countries and in all subsamples that included years from 2002 on. In earlier subsamples (up to 2000 and 2001), the current account balance-to-GDP ratio became the top splitter, but a combination of high current account deficits and a high ratio of short-term external debt to reserves characterized crisis-prone observations. This latter evidence suggests that the concurrent buildup of reserves and tranquil international financial markets of recent years is not the only reason BCTs prefer these three indicators to others.

If our estimates were taken at face value, they would suggest a much stronger role for macroeconomic variables than for financial sector variables. But caution should be used in reading this suggestive evidence. First, in this paper we used BCTs mostly as a forecasting tool. The

predominant role of macroeconomic variables is, therefore, only a sign that they are better leading indicators at a one-year horizon, which is not that surprising in view of the balance sheet, and therefore backward-looking nature, of most financial sector indicators, although the few market-based forward-looking indicators that we consider, such as EMBI spreads, also do not seem to have much information content. Second, the evidence of this paper leaves it open that, although financial sector variables are not good leading indicators of crisis inception, they might still play a pivotal role in determining how disruptive capital account crises might be once they occur. In other words, financial sector weakness is a vulnerability that raises crisis risk only when it assumes a macroeconomic dimension (for example, short-term foreign indebtedness of banks and corporations needs to be high not only in relation to other countries but also in relation to international reserves). Further research on the crisis role of financial variables is clearly needed.

BCTs have a clear advantage in permitting analysis of a large number of potential indicators and their interactions but they also have important limitations. One of their unappealing features is potential instability (consider, for example, how our trees changed between Figures 3 and 4 only as a result of an additional year of data). In this paper, we used Breiman's RandomForests algorithm—which compares forecasts from a multitude of trees estimated using randomized samples and indicator sets—to verify whether the classification of each observation remained stable as trees changed and to check the robustness of our results. We found that the out-of-sample performance of our preferred trees was comparable to that of the RandomForests algorithm.

Another limitation of the BCT algorithm is that it considers each split sequentially without taking into account how it will affect further splits down the tree. That is, in deciding which split to use, the BCT algorithm searches for the indicator and threshold that yield the largest improvement in partitioning a given subsample without considering how difficult it would be to further partition the resulting two subsamples. To remedy this problem, the BCT algorithm should choose splits in a forward-looking manner, but the associated computational costs would quickly become prohibitive.

Finally, several crisis episodes have an important contagion component. It is difficult to quantify contagion, let alone predict it. But factoring contagion considerations in crisis prediction models is of critical importance (see, for example, Kaminsky and Reinhart, 2000). We plan to experiment with possible contagion indicators in future work.

## APPENDIX I

See Table A.1.

| | | Table A.1. Capital Account Crises, 1994–2004 | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Country[1] | (1) Capital account Crisis[2] | (2) Sudden Stop 2[3] | (3) Sudden Stop 1[4] | (4) EWS[5] | (5) Sovereign Default[6] | (6) IMF Program[7] | (7) Banking crisis[8] | (8) Corporate Crisis[9] |
| Algeria | 1994 | 1995 | 1995, 2004 | 1993−94 | 1995−96 | 1994−97 | | … |
| Argentina | 1995, 2001 | 2001 | 2001 | 1995, 2001 | 2001−04 | 1992−96, 2000−01 | 1995, 2001−02 | 2002 |
| Bosnia and Herzegovina | | … | … | | … | 1995, 1998−2002 | | … |
| Brazil | 1998, 2002 | 2002 | 2002 | 1998, 2000 | 1983−94 | 1998−99, 2001−03 | 1994−99 | 1995, 1998, 2001−02 |
| Bulgaria | 1994 | 1996 | 1996 | 1994 | 1994 | 1991−94, 1997−00, 2003 | | … |
| Chile | | 1998, 1999 | 1998, 1999 | 1995−96, 1998−99 | | | | |
| China | | 1998 | 1998 | | | | | |
| Colombia | 1999, 2002 | 1998, 1999 | 1998, 1999 | 1993, 1995 | | | 1999−00 | 2000 |
| Costa Rica | | | | | | | 1994−97 | … |
| Croatia | | | | 1999 | | 1994−95, 1997 | | … |
| Czech Rep. | 1997 | 1997 | 1997 | 1993 | | | | |
| Dominican Rep. | 2003 | 1999, 2003 | 1999, 2003 | 1994, 1998−2000 | 1981−02 | 1998, 2003−05 | | … |
| Ecuador | 1999 | 2000, 2004 | 2000, 2004 | 1998−2000 | 1982−95, 1999−00 | 1994, 2000−03 | 1995−02 | … |
| Egypt | | 1999 | 1999 | 1993 | | | | |
| El Salvador | | 2001 | 2001 | 1995, 2002 | 1981−96 | | | … |
| Estonia | | | | | | 1995 | | … |
| Guatemala | | | 2002 | 2002 | 1995 | | | … |
| Hungary | 1996 | 1996 | 1996 | 2003 | | | | 1995 |
| India | | | | 1996−97 | | | 1991−94 | 1993 |
| Indonesia | 1997 | 1998, 1999 | 1998, 1999 | 1997−98 | 1998−2000, 2002 | 1997−00, 2003 | 1992−95, 1997−02 | 1997−98, 2000 |

## Table A.1 *(concluded)*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Israel | 1997, 2002 | 2002 | 2002 | 1993, 1998 | | | | … |
| Jamaica | 2003 | 2002 | 2002 | 1994, 2003 | | | 1996−00 | … |
| Jordan | | 2002 | 2002 | 1993, 1995 | | 1994−99, 2002 | | … |
| Kazakhstan | | | | 1994 | | 1993−96, 1998 | | … |
| Korea | 1997 | 1997 | 1997 | 1995−97 | | 1997−98 | 1997−02 | 1996−98 |
| Latvia | | | | 1994−95, 2003−04 | | 1992−94 | | … |
| Lebanon | 2001 | 2000 | 2000 | | | | | … |
| Lithuania | 1999 | 2000 | 2000 | | | 1992−97 | | … |
| Malaysia | 1997 | 1998 | 1998 | 1994, 1997−98 | | | 1997−01 | 1998 |
| Mexico | 1994 | 1995 | 1995 | 1994 | | 1995 | 1994−97 | 1998 |
| Morocco | | | | | | | | |
| Pakistan | 1998 | | 1993, 1995−96, 1999−2000 | 1998−99 | 1994, 1999, | 2001−02 | | |
| Panama | | 2000 | 2000 | | 1983−96 | 1994, 1996, 1998 | | … |
| Peru | | 1997 | 1997 | 1998 | 1983−97 | 1997 | | |
| Philippines | 1997 | 1998 | 1998 | 1993, 1997 | | 1997−00 | | 1997, 2001 |
| Poland | | | | 1993−94, 2001 | | 1994 | | 2001 |
| Romania | 1999 | 1999 | 1999 | 1999 | | 1994, 1997, 2003 | | … |
| Russia | 1998 | 1999, 2000 | 1999, 2000 | 1997−98 | 1995−00 | 1994−98 | | 1998 |
| Serbia and Montenegro | | … | … | … | … | 2000−03, 2005 | | … |
| Slovak Rep. | | | | | 1993−94 | | | |
| Slovenia | | 2003 | 2003 | 2000 | … | | | |
| South Africa | 2001 | 2000 | 2000 | 1996, 1998 | | | | |
| Sri Lanka | | 1995 | 1995, 2003 | 1993, 1995, 1997−98 | | 1992−94, 2001−03, 2005 | | … |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Thailand | 1997 | 1997 | 1997 | 1995, 1997 | | 1997−99 | 1997−02 | 1996−97 |
| Tunisia | | 1995, 2000 | 1995, 2000 | 2000 | | | 1991−95 | … |
| Turkey | 1994, 2001 | 2001 | 2001 | 1993−94, 2000−01 | | 1994−95, 1999−02 | 1994, 2000−02 | 1994, 1998, 2000−02 |
| Ukraine | 1994, 1998 | 1995 | 1995 | 1994 | 1998−00 | 1994−98 | | … |
| Uruguay | 2002 | 2002 | 2002 | 2002 | 2002 | 1998, 2002−04 | 2002 | … |
| Venezuela | 1994, 2001 | 1999, 2000 | 1999, 2000 | 1994, 2002 | 1995−97 | | 1993−97 | 1994−95, 1998, 2002 |

[1]Whenever a country is not in the sample used to identify a particular type of crisis, "…" is placed in the cell. Blank cells mean there is no crisis for that particular country.

[2]Capital account crises were obtained primarily from sudden stops, but adjusted using information from other crisis definitions when applicable. The inception year of the crisis is reported. The final classification was made after accounting for the country desks' comments.

[3]Sudden stop 2: a sudden stop 1 where net private capital flows/GDP have declined by at least 3 percent from the previous year and at least 2 percent from two years before.

[4]Sudden stop 1: a year in which one of the following holds (where means and standard deviations are computed based on the 1993-2004 values deflated by the U.S. CPI):
- net private capital flows are at least 1.5 standard deviations below their mean and have declined by at least 0.75 standard deviation from the previous year;
- net private capital flows have declined by at least 1.5 standard deviations from the previous year and at least 0.75 standard deviation from two years before; or
- net private capital flows have declined by at least 0.75 standard deviation from the previous year and at least 1.5 standard deviations from two years before.

The source of net private capital flows data is the World Economic Outlook database.

[5]EWS is an early warning system where the classification is based on 2 –standard deviation threshold for exchange rate pressure indicator. Source: IMF Monetary and Capital Markets Department

[6]Source: Manasse and Roubini (2005), updated with sovereign debt default indicator from Celasun, Debrun, and Ostry (2006).

[7]Records only the years, for each country, when total disbursements were bigger than total repayments (principal, charges, and interest).

[8]Source: Demirguc-Kunt and Detragiache (1998), updated by IMF Monetary and Capital Markets Department.

[9]Source: IMF Research Department (based on Corporate Vulnerability Utility).

# APPENDIX II

## Country-by-Country Details on the Selection of Crisis Episodes

*Algeria*

- *1994* chosen as the inception year because of the EWS and IMF program indicators, although sudden stop and sovereign default indicators pointed to 1995.

*Argentina*

- *1995* chosen because of EWS and banking crisis indicators and a moderate decline in net private capital flows.
- *2001* chosen because of sovereign default, banking crisis, sudden stop, and EWS indicators.

*Brazil*

- *1998* chosen because of EWS and IMF program indicators and a moderate cumulative decline in net private capital flows.
- *2002* chosen because of sudden stop and IMF program indicators.

*Bulgaria*

- *1994* chosen because of sovereign default and EWS indicators, with sudden stop indicators suggesting 1996 instead.

*Chile*

*No capital account crises were identified in this period*, even though both sudden stop and EWS indicators suggested a crisis in 1998. Capital outflows during that year were exacerbated by a domestic portfolio reshuffling away from dollar liabilities and toward dollar assets abroad, resulting from liberalization of the capital account and the elimination of the exchange rate band, rather than a loss of access to international capital markets. Thus, that episode was not considered a capital account crisis.

*China*

*No capital account crises were identified in this period*, even though sudden stop indicators suggested a crisis in 1998. It is possible that the observed decline in net private capital flows was due to data problems (much of the flows take the form of trade credit and errors and omissions given the existence of exchange and capital controls), and there was not enough disruption in the economy to justify classifying that year as a capital account crisis.

*Colombia*

- *1999* chosen because of sudden stop indicators.
- *2002* chosen because of contagion from Brazil (the sovereign spread reached 1,100 basis points during that year) and a moderate decline in net private capital flows.

*Czech Republic*

- *1997* chosen because of sudden stop indicators.

*Dominican Republic*

- *2003* chosen because of sudden stop indicators and real depreciation. Note that the sudden stop indicators also suggested a crisis in 1999, which was ruled out because growth remained high, inflation low, and there was no significant change in either the exchange rate or reserves during that year.

*Ecuador*

- *1999* chosen mainly because of sovereign default (EWS suggested 1998–2000, and sudden stop indicators suggested 2000 instead). Note that sudden stop indicators also suggested a crisis in 2004. That decline in net private capital flows was due to the completion of a pipeline causing foreign investment to decline to levels close to its historical average.

*Egypt*
*No capital account crises were identified in this period*, even though sudden stop indicators suggested a crisis in 1999. The decline in net private capital flows following 1998 was the result of a change in the privatization policy, which reduced the supply of assets available to foreigner investors.

*El Salvador*
*No capital account crises were identified in this period*, even though sudden stop indicators suggested a crisis in 2001. The change in net flows during that year was due to currency substitution and reclassification of assets associated with dollarization.

*Guatemala*
*No capital account crises were identified in this period*, even though EWS pointed to exchange rate pressures in 1995 and the sudden stop indicators suggested a crisis in 2002. The sharp decline in net private capital flows in 2002 was partly due to the disbursements of a Eurobond.

*Hungary*

- *1996* chosen because of sudden stop indicators.

*India*
*No capital account crises were identified in this period.* Despite a moderate decline in net private capital flows in 1995 and EWS pointing to currency pressures in 1996–97, none of these episodes could be described as a capital account crisis.

*Indonesia*

- *1997* chosen even though net private capital flows deteriorated substantially only in 1998–99. Choice of 1997 is supported by EWS, IMF involvement, banking crisis, and corporate crisis indicators.

*Israel*

- *1997* chosen because of substantial decline in net private capital flows, with EWS pointing to exchange rate pressures in 1998.
- *2002* chosen because of sudden stop indicators, possibly reflecting an escalation of conflict in the occupied territories.

*Jamaica*

- *2003* chosen as the crisis year, even though sudden stop indicators suggested 2002. The crisis and RER depreciation occurred in the first quarter of 2003, which is picked up by EWS.

*Jordan*

*No capital account crises were identified in this period.* The decline in net private capital flows in 2002 was perceived to be driven, at least in part, by the buildup to the Iraq War.

*Korea*

- *1997* chosen because of sudden stop indicators, currency crisis, and IMF program.

*Lebanon*

- *2001* chosen as the inception of the crisis even though sudden stop indicators suggested 2000. It was only in 2001 that enough disruption was created to justify a capital account crisis classification (it became very difficult for the government to roll over its debt and it eventually required exceptional financing under Paris II in 2002).

*Lithuania*

- *1999* chosen because of contagion from Russia, even though sudden stop indicators suggested 2000 as the crisis year.

*Malaysia*

- *1997* chosen even though net private capital flows deteriorated substantially only in 1998. Choice of 1997 was supported by EWS and banking crisis indicators.

*Mexico*

- *1994* chosen mainly because the currency crisis took place late that year, with net private capital flows deteriorating substantially beginning only in early 1995.

*Pakistan*

- *1998* chosen mainly because of sovereign default, with EWS pointing to currency pressures in the following two years.

*Panama*

*No capital account crises were identified in this period*, despite sudden stop indicators suggesting 2000 as a crisis year. There was a reduction in short-term capital inflows through the banking system during 2000, reflecting temporary concerns about Panama being on the Financial Action Task Force list of noncooperative countries (it was later removed from the list in 2001), which was followed by a rebound.

*Peru*

*No capital account crises were identified in this period*, despite sudden stop indicators suggesting 1997. The observed decline in net private capital flows during that year may have been partly due to a debt-restructuring operation.

*Philippines*

- *1997* chosen even though net private capital flows deteriorated substantially only in 1998. Choice of 1997 was supported by EWS, IMF program, and corporate crisis indicators.

*Poland*

*No capital account crises were identified in this period*, despite sudden stop indicators suggesting 1994. The decline in net private capital flows in that year may have been a result of data problems because there were no capital account pressures in the mid-1990s.

*Romania*

- *1999* chosen because of sudden stop and EWS indicators.

*Russia*

- *1998* chosen mainly because of currency crisis even though net private capital flows deteriorated substantially only in 1999–2000. Choice of 1998 was also supported by EWS and corporate crisis indicator.

*Slovenia*

*No capital account crises were identified in this period*, despite sudden stop indicators pointing to crisis in 2003. The decline in net private capital flows in 2003 was a correction following two years (2001–02) of exceptional inflows.

*South Africa*

- *2001* chosen despite sudden stop indicators suggesting 2000 instead. Although the decline in net flows started in 2000, the large RER depreciation occurred only in 2001, which is widely perceived as the crisis year. The exchange rate pressures indicated by EWS in 1996 and 1998 did not generate enough disruption to be considered capital account crisis episodes.

*Sri Lanka*

*No capital account crises in this period*, despite sudden stop indicators suggesting 1995 and 2003 as crisis years. None of these episodes created enough disruption to warrant being classified as a capital account crisis.

*Thailand*

- *1997* chosen because of sudden stop, EWS, IMF program, banking crisis, and corporate crisis indicators.

*Tunisia*

*No capital account crises were identified in this period*. The drop in net private capital flows picked up by the sudden stop indicator in 2000 is a result of the large privatization inflows during 1999.

*Turkey*

- *1994* chosen because of EWS, IMF program, banking crisis, and corporate crisis indicators.
- *2001* chosen because of the same set of indicators of 1994 plus sudden stop indicators.

*Ukraine*

- *1994* chosen because of EWS indicator with sudden stop indicators pointing to 1995.
- *1998* chosen because of sovereign default.

*Uruguay*

- *2002* chosen because of sudden stop, EWS, sovereign default, IMF program, and banking crisis indicators.

*Venezuela*

- *1994* chosen because of EWS, with sovereign default indicator pointing to 1995.
- *2001* chosen as the year the crisis gained momentum, despite sudden stop indicator suggesting 1999 instead.

## REFERENCES

Berg, Andrew, Eduardo Borensztein, and Catherine Pattillo, 2005, "Assessing Early Warning Systems: How Have They Worked in Practice?" *IMF Staff Papers*, Vol. 52 (December), pp. 462–502.

Berg, Andrew, and Catherine Pattillo, 1999, "Are Currency Crises Predictable? A Test," *IMF Staff Papers*, Vol. 46 (June), pp. 107–138.

Boyd, John, Gianni De Nicoló, and Abu M. Jalal, 2006, "Bank Risk-Taking and Competition Revisited: New Theory and New Evidence," IMF Working Paper 06/297 (Washington, International Monetary Fund).

Breiman, Leo, 2001, "Random Forests," *Machine Learning*, Vol. 45 (October), pp. 5–32.

———, Jerome Friedman, Richard Olshen, and Charles Stone, 1984, *Classification and Regression Trees* (London, Chapman & Hall).

Calvo, Guillermo, Leonardo Leiderman, and Carmen Reinhart, 1992, "Capital Inflows and Real Exchange Rate Appreciation in Latin America: The Role of External Factors," IMF Working Paper 92/62 (Washington, International Monetary Fund).

Catão, Luis, 2006, "Sudden Stops and Currency Drops: A Historical Look," IMF Working Paper 06/133 (Washington, International Monetary Fund).

Celasun, Oya, Xavier Debrun, and Jonathan D. Ostry, 2006, "Primary Surplus Behavior and Risks to Fiscal Sustainability in Emerging Market Countries: A 'Fan-Chart' Approach," *IMF Staff Papers*, Vol. 53 (December), pp. 401–425.

Demirguc-Kunt, Asli, and Enrica Detragiache, 1998, "The Determinants of Banking Crises in Developing and Developed Countries," *IMF Staff Papers*, Vol. 45 (March), pp. 81–109.

Frankel, Jeffrey A., and Andrew K. Rose, 1996, "Currency Crashes in Emerging Markets: An Empirical Treatment," *Journal of International Economics*, Vol. 41 (November), pp. 351–366.

Frankel, Jeffrey A., and Shang-Jin Wei, 2004, "Managing Macroeconomic Crises," NBER Working Paper No. 10907 (Cambridge, Massachusetts, National Bureau of Economic Research).

Friedman, Jerome H., 1991, "Multivariate Adaptive Regression Splines," *Annals of Statistics*, Vol. 19, No. 1, pp. 1–67.

Ghosh, Swati R., and Atish R. Ghosh, 2003, "Structural Vulnerabilities and Currency Crises," *IMF Staff Papers*, Vol. 50 (December), pp. 481–506.

Jeanne, Olivier, and Romain Ranciere, 2006, "The Optimal Level of International Reserves for Emerging Market Countries: Formulas and Applications," IMF Working Paper 06/229 (Washington, International Monetary Fund).

Kaminsky, Graciela, 2006, "Currency Crises: Are They All the Same?" *Journal of International Money and Finance*, Vol. 25 (April), pp. 503–527.

———, and Carmen M. Reinhart, 2000, "On Crises, Contagion, and Confusion," *Journal of International Economics*, Vol. 51 (June), pp. 145–168.

———, Saul Lizondo, and Carmen M. Reinhart, 1998, "Leading Indicators of Currency Crises," *IMF Staff Papers*, Vol. 45 (March), pp. 1–48.

Manasse, Paolo, and Nouriel Roubini, 2005, "'Rules of Thumb' for Sovereign Debt Crises," IMF Working Paper 05/42 (Washington, International Monetary Fund).

———, and Axel Schimmelpfennig, 2003, "Predicting Sovereign Debt Crises," IMF Working Paper 03/221 (Washington, International Monetary Fund).

Reinhart, Carmen, and Kenneth Rogoff, 2004, "The Modern History of Exchange Rate Arrangements: A Reinterpretation," *Quarterly Journal of Economics*, Vol. 119 (February), pp. 1–48.

———, Kenneth Rogoff, and Miguel Savastano, 2003, "Debt Intolerance," *Brookings Papers on Economic Activity: 1*, Brookings Institution, pp. 1–74.

Van Rijckeghem, Caroline and Beatrice Weder, 2004, "The Politics of Debt Crises," CEPR Discussion Paper No. 4683 (London, Centre for Economic Policy Research).