Appendix

Classification Tree Analysis. Parameters Selection

1. *Default method used to split nodes in the tree*: Gini splitting.
2. *Minimum size node to split*: 20 observations. This constraint results in a smaller tree (fewer nodes), which gives greater predictive accuracy than larger trees. There is a trade–off between the size of the tree and the misclassification rate (Bayes probability error). As the tree grows (more nodes) the method gets a better description of each crises but prediction is poorer.
3. *Probability of crisis*: 0.3071. When the target variable is categorical (as it is in our case) we can assume prior probabilities but rather than weighting categories (crisis and non-crisis) arbitrarily we use frequency distribution in our data set. From the MTI classification we have 43 crises years, which means that probability of crisis is 30.7%.
4. *Misclassification costs*: 1. This means that the cost of misclassify a *crisis year* as a *non-crisis year* is the same as committing the mistake of classifying a *non-crisis year* as a *crisis year*.
5. *Cross-validation*: 10 fold. V-fold cross validation is a technique to determine the optimal tree size. V-fold cross validation performs independent tree size test without requiring separate test datasets and without reducing the data used to build the tree to determine the optimal tree size. The learning dataset is partitioned into some number of groups called "folds". For instance, assume 10 partitions (the default number) are created. In each case, 90% of the data is used to build a test tree and 10% is held back for independent testing, that is, the classification error for this 10% is computed. This process is repeated 10 times, building 10 separate test tree. Then, their classification error rate as a function of tree size is averaged. This averaged error rate for a particular tree size is the cross validation cost. The tree size that produces the minimum cross validation cost is the optimal tree size.
6. Tree pruning control: we prune the tree to the minimum cross-validated error.

   The classification tree method applied to currency crises has a couple of advantages when compared to other traditional methods. First, it does not impose the same functional form to all crises such as logit and probit models[1]. Second, the probability of crisis augments as the number of variables indicating vulnerability increases. For example, an expanding domestic credit may be explosive in the context of fixed exchange rate with capital inflow reversal.

The classification also identifies which subset of early warnings is most important to classify and predict crises. Interestingly, decision tree also handle missing data values in the sample using surrogate splitters, which are back-up rules that closely mimic the action of primary splitting rules.

[1] Cerro and Meloni (2013) estimate a Logit function