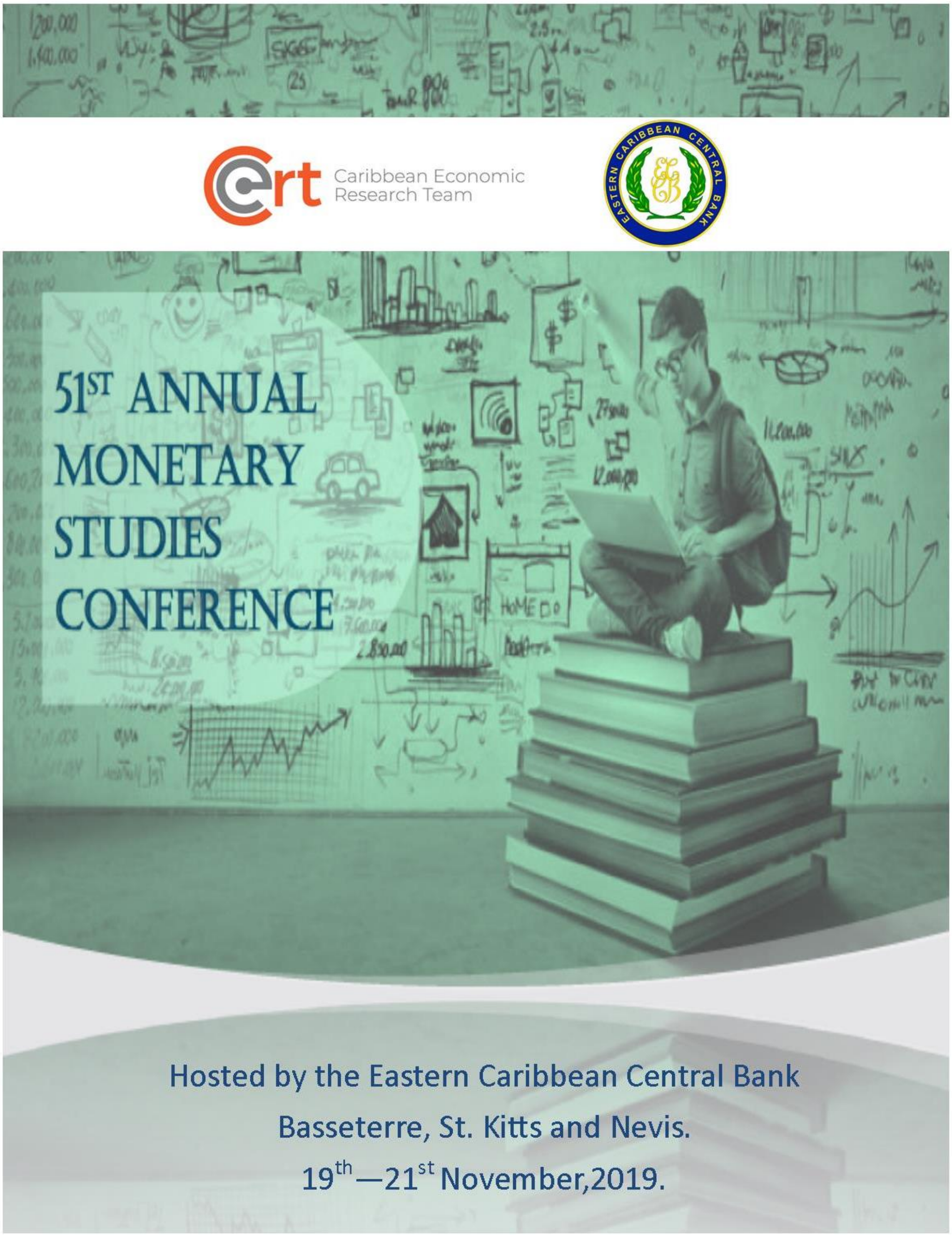

Caribbean Economic Research Team
51ST ANNUAL MONETARY STUDIES CONFERENCE
Hosted by the Eastern Caribbean Central Bank
Basseterre, St. Kitts and Nevis.
19th—21st November,2019.

An Early Warning Model of Financial Stress for the ECCU[1]

Allister Hodge[2]

Abstract

Early warning models are an integral component of macroprudential policy framework as it provides policy makers with a set of indicators alongside judgment in assessing financial stability. This paper attempts to build an early warning model for identifying financial stress in the banking sector of the ECCU region. The paper utilises classification trees and the random forest variant to arrive at a set of early warning indicators to identify banking sector stress. The main variables that emerge from the random forest are, private sector credit growth including nonfinancial corporation and household credit, government credit, liquidity, banking sector efficiency and global credit growth. The thresholds for these variables are also derived from the classification tree.

Keywords: Early Warning Systems, Banking Crises, Decision/Classification Trees, Random Forest.

JEL Classification C40, G01, E44, E61, G21.

---

[1] This Working Paper should not be reported as representing the views of Eastern Caribbean Central Bank. The views expressed are those of the authors and do not necessarily reflect those of Eastern Caribbean Central Bank.

[2] Allister O Hodge. Author affiliation Eastern Caribbean Central Bank. Research Department, Eastern Caribbean Central Bank, P.O. Box 89, Bird Rock Road, St Christopher and Nevis. Email: allister.hodge@eccb-centralbank.org. Tel Phone No. 1-869-465-2537 ext. 3504. I would like to thank the economist at the ECCB research department for their invaluable feedback on this paper. All remaining errors and omissions ae my own.

Table of Contents

Table of Figures

1.0 Introduction

Is it possible to predict the next banking crisis? Almost certainly not. On the one hand, it may be argued that it is econometrically very challenging to predict these very rare events, which in many cases have different causes and consequences. On the other hand, if accurately predicting an emerging banking crisis with some anticipation were feasible, policymakers would ideally be able to take all the necessary measures to avoid its materialization, which would then make the method fails. All crises evolve differently therefore making it extremely difficult to predict them, for example the recent European crisis and USA mortgage crisis evolved from different factors but led to the same effect.

Since the global financial crisis, there has been a renewed interest in using macroprudential policy and tools as a response to address financial instability. Accompanying this growth in macroprudential supervision has been a plethora/burgeoning of approaches used to assess the stability of the financial sector, these range from very simple indicators (credit to GDP gap) to more sophisticated models such as, dynamic stochastic general equilibrium models and agent-based models.

One of the main tools used in macroprudential analysis are, early warning models /indicators (EWI/M). EWM/I have their roots dating back to the seminal contribution of (Kaminsky and Reinhart, 1996) who focused on currency crises. Today, EWMs are an essential component for the implementation of time-varying macroprudential policies, such as countercyclical capital buffers, that can help reduce the high losses associated with banking crises and are hence important to the effective implementation of macroprudential policies (see for example Detken et al. (2014)).

Due to the large costs that economies suffer in a financial crisis, effective early warning tools have substantial value to policy-makers by allowing them to detect underlying economic weaknesses and vulnerabilities, and possibly take pre-emptive policy actions to head off the potential crisis or limit its effects. Therefore, EWMs in this context must not only have sound statistical forecasting power, but also need to satisfy several additional requirements. For instance, signals need to arrive early enough, so that policy measures have enough time to be effective, and they need to be stable as policy makers tend to react on trends.

Considering the fact that EWIs are ultimately judged on their efficacy to be used as macro prudential policy tools, they should satisfy three requirements.

Firstly, timeliness: EWMs must provide signals early enough so that policy actions can be implemented in time to be effective. The timeframe required to do so is dependent on inter alia the lead-lag relationship between changing a specific macro prudential tool and the impact on the policy objective.

The second requirement is stability: the indicator should not flip-flop between signalling a crisis and being *"off"*. EWMs that issue stable signals reduce uncertainty regarding trends and allow for more decisive policy actions.

The final requirement is interpretability: EWI signals should be easy to analyse and interpret and should be making sense for the policy makers to work upon.

Under the ECCB's Strategic Plan 2017-2021, financial stability has been identified one of the four (4) identified pillars which the plan to focus on over the next several years. As such, the Eastern Caribbean Central Bank has developed a macroprudential policy framework to monitor systemic financial stability across the Eastern Caribbean Currency Union (ECCU) [3].

Given the importance of EWM/I in macroprudential analysis this paper attempts to develop EWM for the countries of the ECCU. It is expected that this will contribute to the current macroprudential toolkit used by the (ECCB). It is the hope of this paper that results derived from this can be used to help policy discussions surrounding potential variables to be used in financial stability monitoring.

To knowledge this is the second attempt at developing an EMW for the ECCU region, the first attempt by (Anderson, 2012). In her analysis (Anderson,2012) would have defined a stress event in the banking sector as an infraction in meeting the weekly 6.0 per cent reserve requirement ratio. A logit model was used to estimate the EWM; the variable, which had the highest probability of distress event, was the return on average assets (ROA). However, based on the results this indicates that as return on assets increased the probability of distress increased a counterintuitive result though not discussed in the paper.

Sahely and Polius (2003) used discriminant analysis to map periods of financial stress for the banking sector of the ECCU. The results from the discriminant analysis suggest that loan quality, liquidity, earnings, growth in imports and the ratio of domestic credit to (GDP) were

[3] The countries of the ECCU are Anguilla, Antigua and Barbuda, Dominica, Grenada, Montserrat, St Kitts and Nevis, Saint Lucia and St Vincent and the Grenadines. Anguilla and Montserrat are dependents of the United Kingdom whilst the other six countries are independent. Jointly the countries share a common currency the Eastern Caribbean Dollar (XCD), they operate under a quasi-Currency Board Arrangement with the XCD fixed at $2.7 XCD to $1.00 USD. The Eastern Caribbean Central Bank is the Central Bank .

the best candidate variables in explaining financial stress in the banking sector. For the full sample of banks, the predictor variable with the greatest influence was liquidity as measured by the loans to deposit ratio. Loan quality and growth of the banking sector relative to the economy were also found to be strong predictor variables

This paper makes use of classification trees (CART) to develop an EWM of banking stress for the ECCU. CART is a statistical methodology which retains the advantages of the two approaches traditionally used in the early warning literature, i.e. the signalling and the discrete choice (logit/probit) approach or in some parlance non-parametric and parametric approach. The EWM allows for identification of early warning indicators of banking stress in the region. The methodology developed in this paper allows not only for the identification of relevant indicators but also relevant thresholds for the EWIs.

The rest of the paper is organised as followed, section two (2) review of literature, section three (3) methodology, section four (4) discussion of data and stylised facts, section five presents the results and section six (6) concludes the paper with some policy recommendations.

2.0 Literature Review

The work on EWM is quite extensive and dates back to the 1960's; when the use of discriminant analysis was the dominant methodology being used, this technique persisted up the 1980's. The methodology was first advanced by (Beaver, 1966) and then closely followed by (Altman, 1968). Since then these models have been supplanted by the popular binary dependent variable models namely, logit and probit models and their more advanced counterparts namely the multinomial and dynamic versions of these models. (Pattillo and Berg, 1998), (Demirgüç-Kunt and Detragiache, 1999) were perhaps some of the first authors to have used the probit logit approach, this was followed by using multinomial models (Bussiere and Fratzscher, 2003 and 2006), which generalise the discrete choice from two (yes/no) to more states, such as crisis, post-crisis, and tranquil periods.

In parallel, the simple yet intuitive signal extraction approach based on the work of (Kaminsky and Reinhart, 1996) which aims to simply find thresholds on individual indicators also gained popularity. However, over the years as the methodologies in economics have evolved and computing power increasing coupled with increases in data points so did the EWM.

A new group of flexible and non-linear machine learning techniques have been introduced to various forms of financial stability surveillance; these include the use of Binary Classification

Regression tree models. Recent literature indicates that these novel approaches hold promise for systemic risk identification because of their ability to identify and map complex dependencies. The premise of difference in performance relates to how methods treat two aspects: individual versus multiple risk indicators and linear versus non-linear relationships. While the simplest approaches linearly link individual indicators to crises, the more advanced techniques account for both multiple indicators and different types of non-linearity, such as the mapping of an indicator to crises and interaction effects between multiple indicators.

In the 1990s a wide-ranging methodological debate started, including studies on banking and balance-of-payments problems (Kaminsky and Reinhart, 1996) and currency crashes (Frankel and Rose, 1996). This is the signalling approach, which is also a non-parametric approach. This is a rather simple approach, which involves checking for a signal from several indicators at least 24 months (2years) ahead of a crisis. The model depends on an indicator breaching some threshold either some upper or lower percentile depending on the type of data.

Research papers utilising classification trees in the context of EWMs are still relatively scarce. In examining banking crises in 50 emerging market and developing countries during 1990–2005 (Duttagupta and Cashin, 2008) used a binary classification tree to arrive at their results. By using the BCT, the authors were able to derive key indicators and their threshold values at which vulnerability to banking crisis increases. The three conditions or banking crises, (i) very high inflation, (ii) highly dollarized bank deposits combined with nominal depreciation or low liquidity, and (iii) low bank profitability. They highlight that foreign currency risk, poor financial soundness, and macroeconomic instability are key vulnerabilities triggering banking crises

(Manasse, Savona and Vezzoli, 2013) utilising a variant of a classification and regression tress called ***CRAGGING (CRoss-validation AGGregatING)*** to build an early warning model for banking crises in emerging markets. According to the authors their method was able to dominate more traditional methods such as a Stepwise Logit, a Classification Tree, and an *"Average"* model, and we find that our model. They used 504 variables to classify banking crisis. Based on their results they identified two types of banking crises in emerging markets: a ***"Latin American type",*** resulting from the combination of a (past) credit boom, a flight from domestic assets, and high levels of interest rates on deposits; and an ***"Asian type",*** which is characterized by an investment boom financed by banks' foreign debt.

In a somewhat related paper (Holopainenc, and Sarlin, 2016) conduct a horse race of conventional statistical methods and machine learning methods as early-warning models. Their results confirm that machine-learning methods such as k-nearest neighbours and neural networks, and particularly by model aggregation, approaches through ensemble learning tend to outperform conventional statistical models.

Lastly recent papers by (Alessi and Detken, 2017) and (Joy et al, 2017) are more recent of the papers which use classification trees in the determination of banking crises. However, the more recent papers of Joy et al and Alessi and Detken offer an improvement on the original contribution by Cashin et al in that these papers avoid of the overfitting problem with classification tress by using the random forest methodology.

Another strand of the literature comes from renowned econometrician Peter Phillips in two papers, Phillips et al. (2011, hereafter PWY) and Phillips et al. (2015, hereafter PSY). While the procedure is not necessarily ground breaking since it relies on the already developed methodology of Dickey and Fuller (1979). The procedure set out by PWY and PSYT set out to use an advanced version of the ADF test to detect the presence of financial bubbles by using a recursive and rolling sample procedure and in so doing using a right-hand side test of ADF test the presence of financial bubbles. Broadly applied their procedure seeks to provide policy makers with a tool that can be used in the identification of financial bubbles ex-ante before it is too late. Therefore, this work can easily be extended to the area of macro-prudential monitoring to develop an EWM.

3.0 Methodology

Firstly (i), defining what a crisis or stress period is for the ECCU region; secondly (ii) explaining and describing the methodology that will be used in the paper to arrive at the results, thirdly (iii) data gathering and possible cleaning of the data that may be necessary.

(i) Definition of Stress Event

Several definitions of banking crisis exist in the literature. In a nutshell these definitions focus on four broad definitions, these include the asset quality, liquidity, capital and government assistance. For example, (Demirguc-Kunt and Detragiache (1999)) take the asset quality approach by focusing on an increase in NPLs above 10.0 per cent. The liability approach focuses on the liability side of banks' balance sheet. The essence of this approach is the incidence of bank runs. And the government assistance approach identifies banking crises as at least one of the policies such as (i) large-scale nationalization of banks, (ii) deposit freezing,

(iii) bank closure, and (iv) bank recapitalization is undertaken, this definition has been used by (Kaminsky and Reinhart (1999)) and (Laeven and Valencia (2008), (2010), (2012)). Lastly (Caprio and Klingebiel (2003) focus on 'episodes in which much or all of bank capital was exhausted'. (Alessi and Detken (2017)) provide another definition which states a banking crisis is defined by "*significant signs of financial distress in the banking system as evidenced by bank runs in relevant institutions or losses in the banking system (nonperforming loans above 20 per cent or bank closures of at least 20 per cent of banking system assets); or significant public intervention in response to or to avoid the realization of losses in the banking system*".

In the ECCU, there have been limited occurrences of bank failures[4], hence the focus of this paper is on episodes of financial stress which also incorporates banking failures as well.

A financial stress is defined as:

i. The failure of one or more commercial bank in a country;
ii. The ratio of NPLs rises above 10.0 per cent for a period exceeding 8 consecutive quarters this follows (Demirguc-Kunt and Detragiache (1999));
iii. Profits fall by at least 0.75 standard deviation below its historical mean/ average;
iv. The growth rate of lending remains negative for 8 consecutive quarters or more;
v.

The periods of stress a lagged between 4 to 16 quarters preceding the stress event. The intuition behind this approach is that it allows policy makers sufficient time to respond to a stress event before it hits (ring the alarm). For this reason, we follow Bussiere and Fratzscher (2006) and define our dependent variable, $Y_{i,t}$ as a forward-looking variable

$$Y_{i,t} = \{ \begin{matrix} 1 \; if \; FS_{i,t+k} = for \; k \, \epsilon [5,12] \\ 0 \qquad\qquad\qquad otherwise \end{matrix}$$

Where $FS_{i,t+k}$ signifies that country i experienced a financial stress at time$t + k$. Thus, the dependent variable takes the value one (1) during the 5 to 12 quarter period preceding a financial stress. Like Behn et al. (2013), I omit all observations in which a country is classified to have experienced a financial crisis, as well as the 6 quarters succeeding a crisis. This is done to avoid the post-crisis bias, as discussed in e.g. (Bussiere and Fratzscher, 2006).

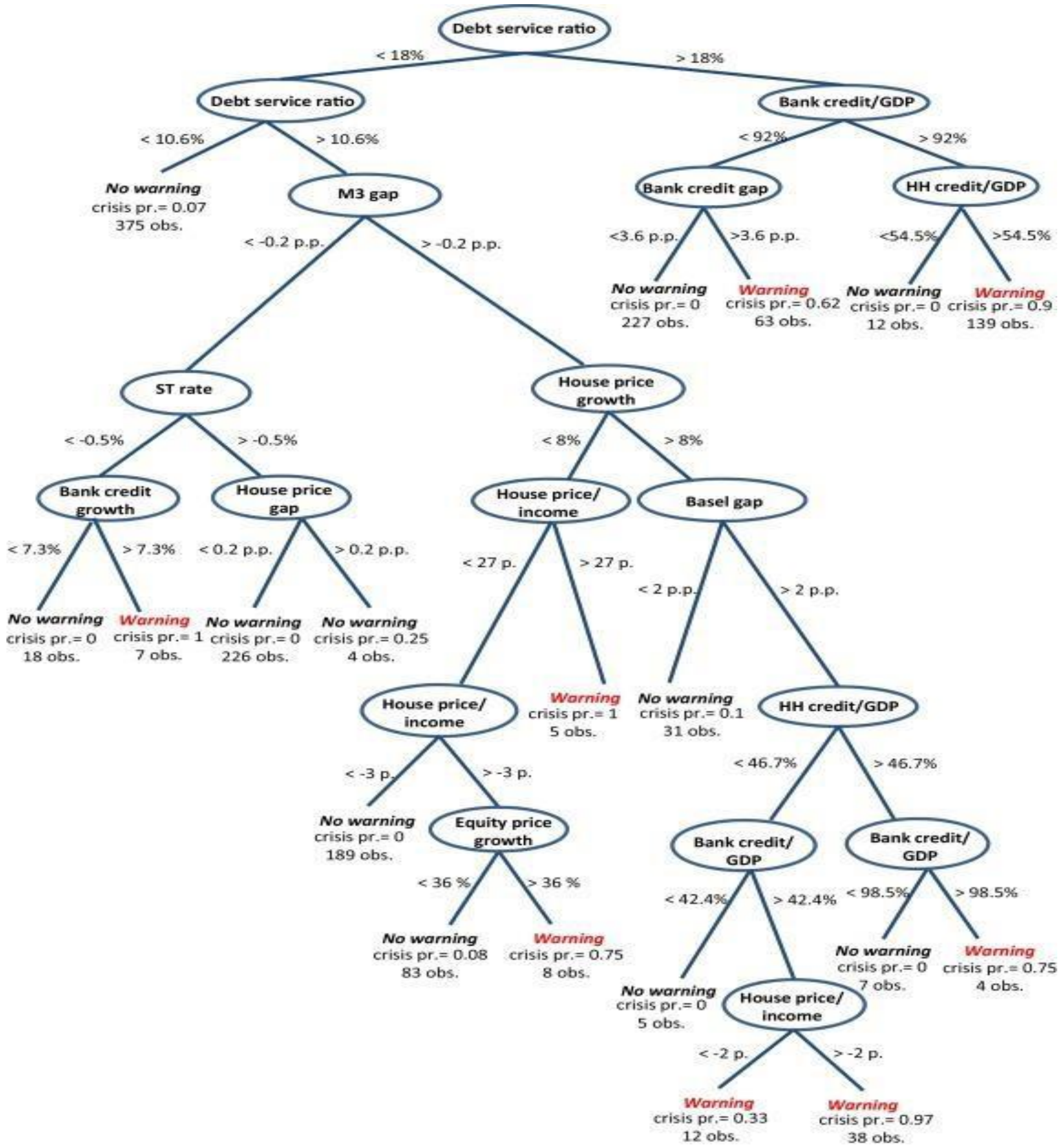
4 Notable bank failures in the region include Bank of Commerce St Kitts and Nevis in the early 1980's caused by a bank run. Bank of Antigua in Antigua and Barbuda in 2010 due to a bank run, ABI Bank Antigua and Barbuda in 2012 due to insolvency, Caribbean Commercial Bank and National Bank Anguilla in 2013 both due to insolvency. Bank of Montserrat required significant capital injection by the Government in 1996.

(ii) Modeling Strategy

In developing the EWM, classification and regression trees and the random forest variant and logit regression were used. The logit model is well known binary dependent variable model and the RF being a machine learning classification system. The random forest is machine-learning methodology built on the idea of classification and regression trees.

To explain the method, classification trees are a statistical procedure, which seeks to partition data into nodes based on the relevant criterion being assessed. Therefore, for example, this paper is assessing predictors of banking sector stress; there are two outcomes, which are ***"stress"*** and ***"no stress"***. In this case a CART will seek to assess the data based on these criteria and spilt the data based on these two criteria, i.e. stress or no stress. In short, the classification-tree approach searches through different possible splits for all explanatory variables and selects those splits that best separate stress episodes from no-stress episodes. In doing this process, the algorithm actually searches for the threshold at which to makes these splits. It is for this reason why the CART has some advantages over the signals and logit approach. For example, when using the logit model, the model would have been able to produce the threshold at which a stress episode is triggered rather it would produce a probability or marginal impact. Additionally, with a CART the method can handle non-stationary variables and is not sensitive to outliers, which for a logit model can be problematic as this can inflate the standard errors.

Figure 1: Example of a Classification Tree (borrowed from (Alessi and Detken, 2017))



Despite its many strengths, the classification tree approach, if employed without due care; can produce results susceptible to *"over-fitting"*. That is, the approach can create overly complex trees that do not generalise well from the training data. Additionally, classification trees are known not to be particularly robust when additional predictors or observations are included.

To overcome this problem this paper uses the random forest algorithm, which allows the researcher to choose the best-fit tree with the variables found to be most important (Breiman et al., 1984).

Unlike the classification tree algorithm, which fits just a single tree to the data, a random forest fits many hundreds possibly thousands of trees (depending on the desired accuracy) from randomly permuted sub-samples of the data. The random forest methodology is part of the broader aspect of ensemble learning falling under the category of "bagging".

The random forest algorithm works in the following fashion:

1. Suppose there are N observations and M variables in training data set. First, a sample from the training data set is taken randomly with replacement.
2. A subset of M variables is selected randomly, and whichever variables gives the best split is used to split the node iteratively.
3. The tree is grown to the largest.
4. The above steps are repeated, and prediction is given based on the aggregation of predictions from n number of trees. During this step K-fold, cross validation is used to improve the predictive accuracy or reduce the error rate of the model/tree.

Around two thirds of the data are selected in each sub-sample. The remaining data (called out-of-bag data) are preserved in order to establish variable importance and out-of-sample error rates. Each iteration issues a vote on variable importance and majority votes are used to yield the final, pooled, variable-importance rankings.

Random forests offer two key advantages in this context: **(i)** insensitivity to outliers **(ii)** handles missing values and maintains accuracy for missing data, **(iii)** handles higher dimensionality data very well and **(iv)** avoidance of over-fitting. Variable-importance rankings are, as a result, far more robust than those issued by a single classification tree. Each of the trees in the forest is an out-of-sample exercise, as the observations that are not used to grow the tree (so called out-of-bag observations) can be put down the tree to get a classification. It is therefore possible to compute the total misclassification error of the forest.

A natural question then, why not just use the RF to determine a decision tree of interactions and thresholds? The answer is straightforward, apart from what they can tell us about variable importance, random forests are difficult to interpret. By their very nature of being based on a majority-voting procedure from a multiplicity of sub-samples, they cannot be used to backward-induce a single tree of interaction effects. So, once we have used random forests to inform us about variable importance, and select the most relevant variables, we return to the standard classification tree to shed light on interaction effects.

A second step we use to avoid over-fitting is pruning. After our random forest, procedures have been run, and once we are in a position to construct a standard classification tree, avoidance of over-fitting is achieved by growing an overly large tree and then pruning its unreliable branches. Pruning can be regarded as a search problem, where one looks for the best-pruned tree.

Step 2:

Once the CART procedure has been completed, this paper goes on to estimate a logit model[5], using the variables chosen by the random forest model as being important. The goal behind estimating the logit model is to obtain probability of a stress event. The logic behind logit models are well established so there is no need to go over them here.

$$Prob(y_{it} = 1 \mid \Phi(\alpha_i + \beta' X_{it}) = \frac{e^{\alpha_i + X'_{it-1}\beta}}{1 + e^{\alpha_i + X'_{it-1}\beta}}$$

In plain english, this is read as the probability of banking sector of country $\boldsymbol{i}$ in period $\boldsymbol{t}$ being in stress period conditional on a set of variables in country $\boldsymbol{i}$ in period$\boldsymbol{t}$. In this specification $Prob(y_{it} = 1 \mid X_{it-1})$ is defined the probability that a given country $\boldsymbol{i}$ in a given quarter $\boldsymbol{t}$ is in a pre-crisis state. The logit models we estimate may also include a set of country dummy variables, α_i and X_{it-1} represents a matrix of variables that are regressors and β' being the coefficients which are associated with the chosen variables.

To estimate the logit model, I use the LASSO (Least Absolute Shrinkage and Selection Operator) logistic regression (Tibshirani [84]) which attempts to select the most relevant predictor variables for inference and is often applied to problems with a large number of predictors. The method maximizes the log likelihood subject to a bound on the sum of absolute values of the coefficients $max_{\beta^l} = (\beta|y) - \lambda\Sigma_i|\beta_i|$, for which the $|\beta_i|$ is penalized by the L_i norm. This implies that the LASSO sets some coefficients to equal zero and produces sparse models with a simultaneous variable selection. The optimal penalization parameter _ is oftentimes chosen empirically via cross-validation. Thus, I follow the literature and use the logistic LASSO method as the modelling technique in combination with cross-validation to set the penalty parameter that determines the complexity of the model.
By estimating the logit model, I can recover the marginal contribution of each variable to the probability of creating a stress event. In addition, to recover the probabilities of a stress event.

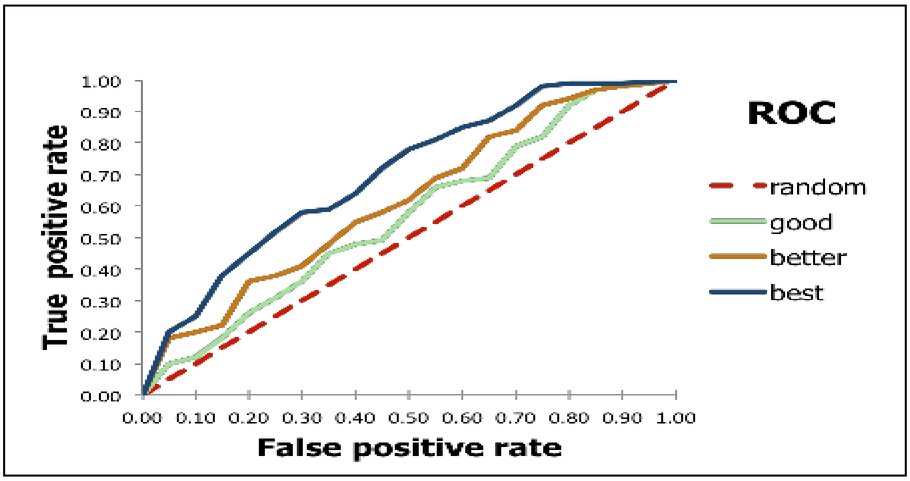
[5] In future work it would be interesting to use a Logit model with LASSO to see how much the results change.

Step 3:

The goal of any EWM is to reduce the type II errors or rather the production of false positives since this carries a large cost to the policy maker. A type II error refers to a situation where a signal is sent from an indicator, but it turns out to false. The alternative is the that no signal is issued but a stress event occurs, this is referred to as a type I error. Therefore, the goal to minimise the ratio of Type I error to Type II errors.

The assess the ability of the models to predict banking stress the receiver operating characteristic curve (ROC) and the area under the ROC (AUROC) metrics are used along with the policy maker loss function see (Alessi and Detken,2017). The term ROC reflects the origins of the tool in the analysis of radar signals during World War II, although it has a long tradition in other sciences (e.g. Swets and Picket, 1982)). Its applications to economics are more recent (e.g. Cohen et al, 2009), (Berge and Jorda, 2011), (Jorda et al, 2011)). The AUC summarises the trade-off between correct and false signals for all different operator (policymaker) preferences. This type of evaluative model has been used by (e.g. Detken et al., 2014), (Drehmann and Juselius, 2014), (Drehmann and Tsatsaronis, 2014), (Giese et al., 2014)) in evaluating the credit to GDP gap in signalling financial crisis across a broad number of countries. The upshot of this technique is that it does not require any assumption on possible threshold values and weights of the signal ratio against the noise ratio. The AUROC summary statistics are bounded between 0 and 1, whereas higher AUROC values reflect more informative models. A value of 1 would represent a perfect fit, whereas a value below 0.5 corresponds with an uninformative specification.

Figure 2: Example of a ROC Curve



The usefulness approach follows the work of (Alessi and Detken, 2017) hinged on earlier work by (Demirgüc-Kunt and Detragiache, 1999) and (Bussière and Fratzscher, 2008). The loss function of (Alessi and Detken, 2017) is defined as follows:

$$L(\theta) = \theta T_2 + (1-\theta)T_1 = \theta \frac{C}{A+C} + (1-\theta)\frac{B}{B+D}$$

where the right-hand side is a weighted average of the **Type I and Type II errors**, T_1 and T_2, respectively. The weights $\theta$and $(1-\theta)$ in the loss function reflect the policy maker's assumed preferences for **Type I and Type II errors** which is unknown at least ex-ante. A parameter θ value higher than **0.5** means that the policymaker is more averse to missing a signal of an upcoming crisis than to receiving a false alarm. For the most part, as commonly done in the literature, θ is set at $\theta = 0.5$, but also try $\theta = 0.6$ as a robustness check. $\boldsymbol{A}$ is the number of periods in which an indicator provides a correct signal (the crisis starts within 1 to 3 years of issuing the signal), and $\boldsymbol{B}$ the number of periods in which a wrong signal is issued. $\boldsymbol{C}$ is the number of periods in which a signal is not generated during a defined period from the onset of the crisis (1 to 3 years). Finally, $\boldsymbol{D}$ denotes the number of periods in which a signal is correctly not provided. In other words, $\boldsymbol{A} = \boldsymbol{TP}$, number of true positives; $\boldsymbol{B} = \boldsymbol{FP}$, number of false positives; $\boldsymbol{C} = \boldsymbol{FN}$, number of false negatives; and $\boldsymbol{D} = \boldsymbol{TN}$, number of true negatives.

The relative usefulness statistic is defined as

$$U_r = \frac{\min(\theta, 1-\theta) - L}{\min(\theta, 1-\theta)}$$

A relative usefulness $\boldsymbol{U_r}$ of 1 would mean that the indicator is able to perfectly forecast all the pre-crisis windows and produces no false positives. A relative usefulness of 0 or less means that the indicator is not useful.

4.0 Data

The data covers information on the aggregate banking sector for each ECCU country and covers the period 1999Q1 up to 2014Q. The key variables used in this study includes credit variables, macroeconomic variables and other international variables see **Appendix 1, Table 1** for a full definition of the variables used.

Credit Variables

The consensus view in the EWM literature is that rapid credit growth sows the seeds for future financial crises (see e.g. Reinhart and Rogoff (2008), (Schularick and Taylor ,2012) and (Mendoza and Terrones ,2008).

The credit variables are presented in three (3) ways, first the year on year (y-o-y) rate of growth, secondly as ratio of GDP and thirdly as deviations of such ratio from its trend (*i.e. the "credit to gap"*), computed with a backward-looking slowly-adjusting (λ= 200,000 to 260,000) one sided HP filter. The last procedure assumes that the financial cycle is twice as long as the business cycle, meaning that financial cycle adjusts after the business cycle. First broad private sector credit is presented (**pcredit**) and the private sector credit variable is further decomposed into credit to non-financial corporations (**NFC**) and credit to households (**house**). Credit covers both loans and debt securities and measures the amount of outstanding debt at the end of the quarter.

Household credit is further disaggregated between housing loans and consumer durables. In the case if the ECCU housing prices are non-existent and thus I take housing loans as a sign of overheating in the housing market and upward pressure on valuations. Additionally, use is made of the sectoral break down of loans to the economy, (i.e. credit to construction (**const**), tourism, wholesale and retail sectors(**dist**)) which can be viewed as lending to NFC.

I also use global credit data for spill over effects in the ECCU and a measure of capital flows to the ECCU. The logic is that when credit is overheating at the global level it is also leads to greater capital flows to regions like the ECCU where the capital flows are likely to create overheating symptoms in the local economies and an overvaluation of asset prices particularly real estate prices. Finally, use is made of the concentration of the loan portfolio as measured by the Hirschman Herfindahl index (**HHI**). The intuition behind using this metric is that increasing concertation in the loan portfolio can weaken the banking sector is the commonly exposed to the same sector. Net foreign assets as a percent of GDP is used as another risk measure.

Banking Sector Variables

The banking sector variables used include capital to total assets (**eqta**), this variable has been shown by (Barrell et al. ,2010) and (Behn et al., 2013) to be an important predictor of financial crises. The loans to deposit ratio (**liq**) is used as measure of liquidity, it well known that tightening liquidity can lead to a systemic banking crisis. I also use the aggregate banking system cost to income ratio as an indicator (**eff**). The intuition behind this is that during good times banks may increase their cost faster than their income where most ok this may be spend on expensive expansions such as new branches and offices as well as on compensation to management and other staff.

Macroeconomic Variables

The macroeconomic variables we examine are real GDP (**rgdp**) y-o-y growth and foreign direct investment and current account in percentage of GDP. Money aggregates, in terms of real y-o-y rate of growth and gap and deseasonalised tourism growth are also considered. This paper also considers government debt to GDP ratio and overall fiscal balance as a percent of GDP as well as Government borrowing from commercial banks as a potential variable as well. The rationale for using the Government fiscal position is due to sovereign bank nexus, for example a deterioration of a country's fiscal position and creditworthiness may reduce sovereign bond prices, generating losses and weakening banks' capital position through their holdings of sovereign paper. This may undermine the system's ability to provide credit to the private sector, which, in turn, would lead to lower economic activity and a further deterioration of the fiscal position. Additionally, sovereign debt restructuring due to a weak fiscal position can affect the banking sector through both the holding of Government debt instruments and loans.

Temporal properties of the data

The presence of unit roots in the data can affect standard inference when doing time series econometric analysis; see e.g. the seminal paper by Granger and Newbold (1974). This problem can also arise when doing econometric work with binary dependent variables as in the case of doing the logit model in this paper, see Park and Phillips (2000). Thus, for the reliability of drawing inferences from the model, it is important to establish the temporal properties of the data series considered in the empirical analysis. To explore this, both panel data unit root test, namely the Im-Pesaran-Shin (IPS) (see Im et al. (2003)) and a Fisher-type tests (see Choi (2001) for a discussion) for testing for unit roots in unbalanced panels. These tests are executed using STATA 14.2.

Results from the unit root tests are presented in **Table 3** in Appendix. While the results are not unambiguous for all the series, we follow the literature and continue our analysis under the modelling assumption that all the series are stationary.

5.0 Results

The goals of using the random forest methodology is to select those variables, which have good predictive ability in predicting a banking sector stress. The number of trees used in constructing the forest amounts to 30,000 trees, repeated k-fold cross validation is used in training the model; to assess the accuracy of the method the ROC is used. The variables selected from the random forest are used to construct the classification tree.

The results from the random forest are shown in **Figure (3),** the ten most important variables are chosen for this study**.** The vertical axis lists the most important variables by order of importance from highest to lowest, the horizontal axis gives the score for relative importance of each variable.

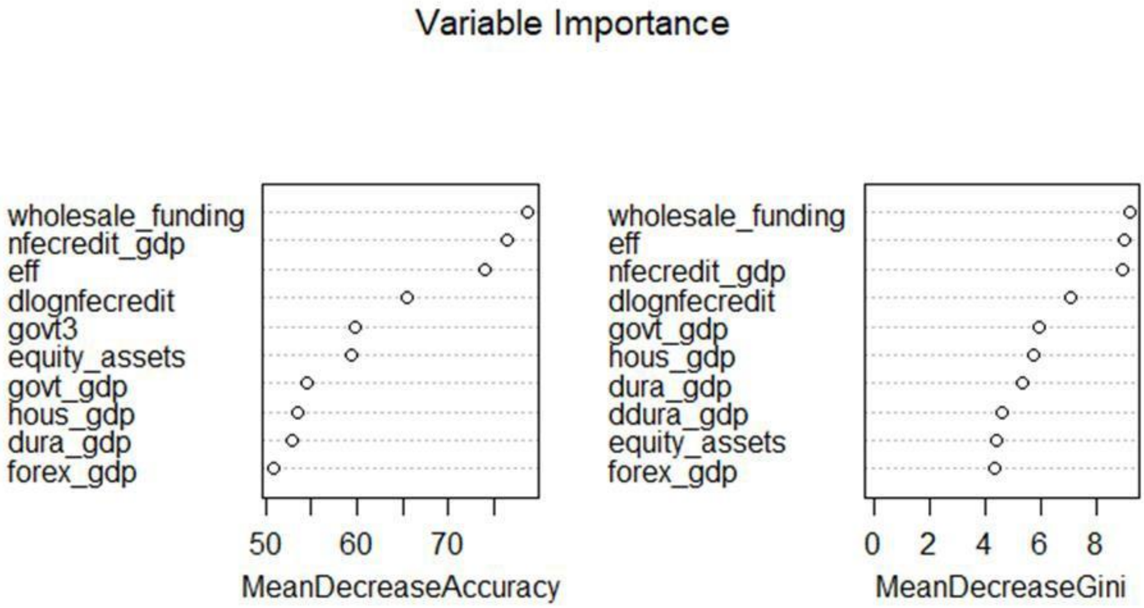
The ten most important variables are cost to income (***eff***) which is used as a proxy for bank management efficiency, government borrowing from the banking sector to GDP(***govt_gdp***), the ratio of wholesale funding relative to total funding (***wholesale_funding***), the growth rate of credit to non-financial enterprises (***dlognfecredit***), the ratio of NFE credit to GDP (***nfecredit_gdp***), credit for house and land acquisition to GDP (***house_gdp***), credit for consumer durable to GDP (***dura_gdp***), the rate of growth of private sector credit(***dlogpredit***) and lastly another Government lending variable which is the ratio of government loans to total assets of the banking sector (***govt3***).

Government lending as an important variable highlights the importance of the sovereign bank nexus. What we come away here is that credit to NFEs, credit for house and land acquisition and consumer durables, funding type and government lending[6] are important predictors of financial stress in the ECCU banking sector.

However, this methodology does not allow us to get the thresholds at which these variables affect stress, for that use is made of the CART tree. The results from the classification tree is shown in **Figure 4** on the ensuing page.

Figure 3: Top (10) Important Variables from the Random Forest[7]

[6] There have several cases of Government debt restructuring in the ECCU, this includes the Government of Antigua and Barbuda, Dominica and St Kitts and Nevis and a soft restructuring by Grenada and St Vincent and the Grenadines.
[7] Importance of the variables ranked from top to bottom with the top variable being the most important and the last variable not the least important but in this case of the top ten, the least important.



Our baseline tree for banking sector stress is shown in Figure 4. Lending to NFE (***dlognfrecredit***)emerges as the main splitter at the top of the tree (Node 1), with an estimated threshold value of 18.0%. Following the main right branch of the tree (where NFE lending growth is above the threshold of 18.0%) lending for consumer durables as a percent of GDP (***dura_gdp***) with an estimated threshold value of 5.3% emerges as a second split along this left branch of the tree.

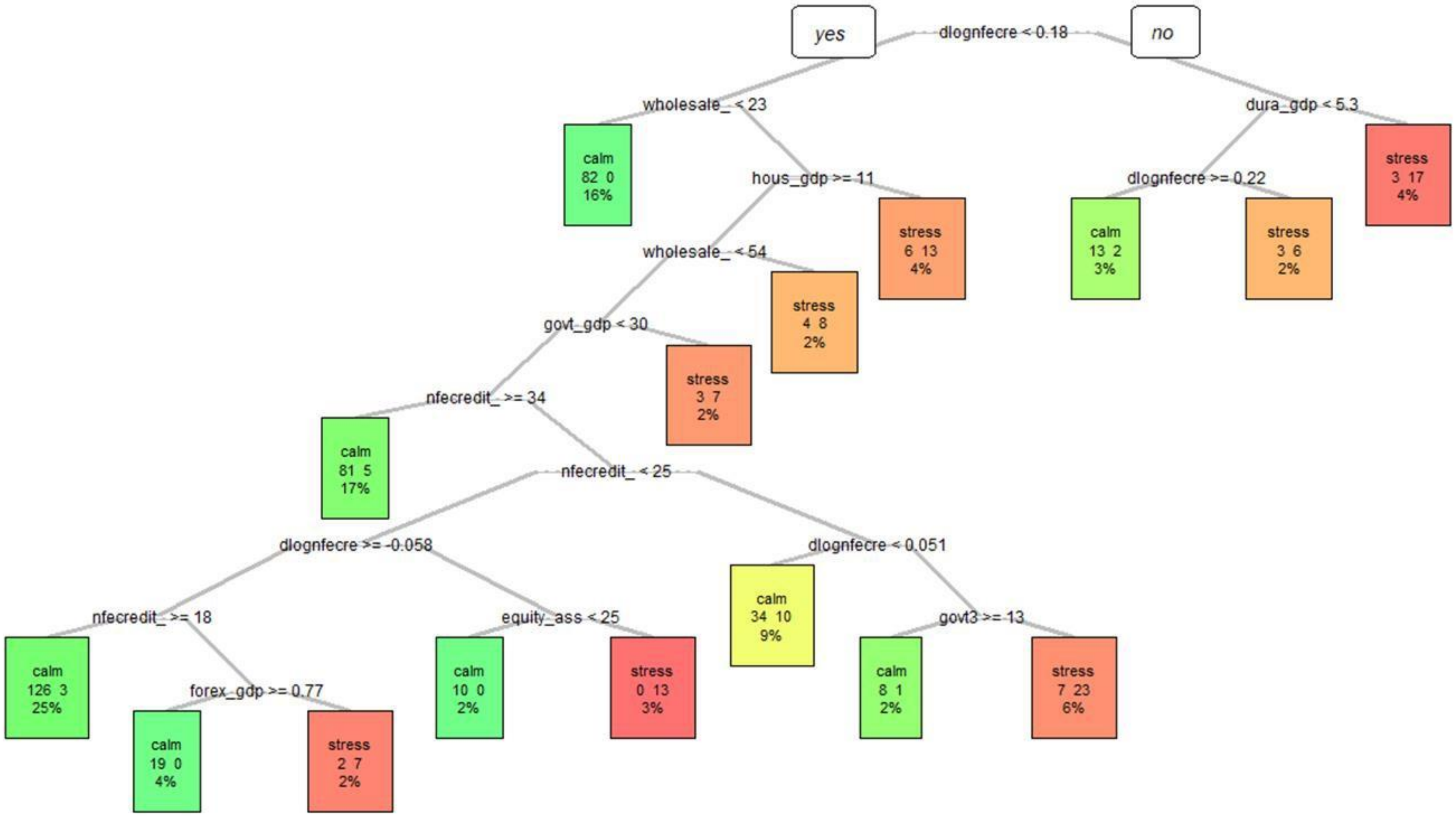
When durable goods lending is less the 5.3% (i.e. the right branch of ***dura_gdp***), the growth rate of nfe lending appears as a conditional splitter (Node 5) with an estimated threshold value of 22.0%. When durables good lending and nfe credit growth are less than their estimated thresholds this increases the probability of financial stress in the banking sector, this is right branch of the node associated with nfe credit. The probability of this occurring is 67.0% and happens in about 2% of the observations.

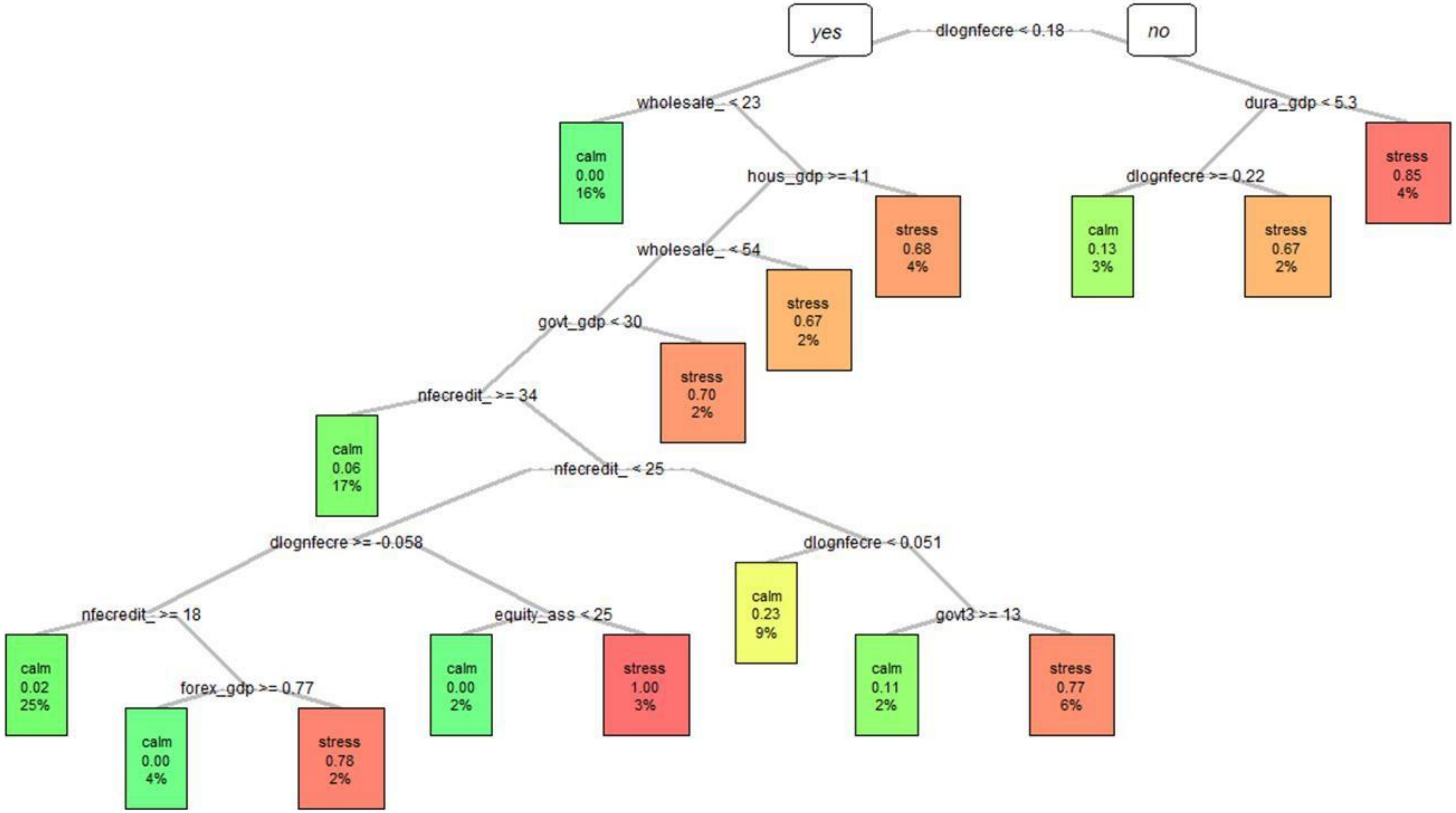
Moving toward the left side of the tree where growth rate of NFE lending is below the estimated threshold of 18.0% wholesale funding (***wholesale_funding***) emerges as the next main splitter with an estimated threshold of 23.0 per cent. When wholesale funding is greater than the estimated threshold of 23.0% housing loans as percent of GDP (***hous_gdp***) emerges the main splitter with a threshold of 11.0%. The branch terminates when housing loans to GDP is less than 11.0 om the right side of this node. Hence when wholesale funding is greater than 23.0% and housing loans is below 11.0% then there is estimated probability of banking sector financial stress of 68.0%, this occurs in 13 of 19 cases found.

As we travel down the tree on the right side of the housing loans node where it is above the threshold of is greater than or equal to 11.0% then wholesale funding emerges as the next splitter this time with an estimated threshold of 54.0 per cent. This node terminates on the right-hand side when wholesale funding is greater than 54.0%. This suggest that when housing loans is greater than 11.0% of GDP and wholesale funding is greater than 54.0% there is an estimated probability of banking sector financial stress of 67.0% or 8 of the 12 episodes found. Below the wholesale funding node on the left side is commercial bank lending to Government as a percent of GDP emerges as the next split with a threshold of 30.0%. This branch terminates on the right-hand side when government borrowing exceeds 30.0%, with a probability of 70.0%

Variables with counterintuitive results include equity assets greater than 25.0 and household credit below 11.0%

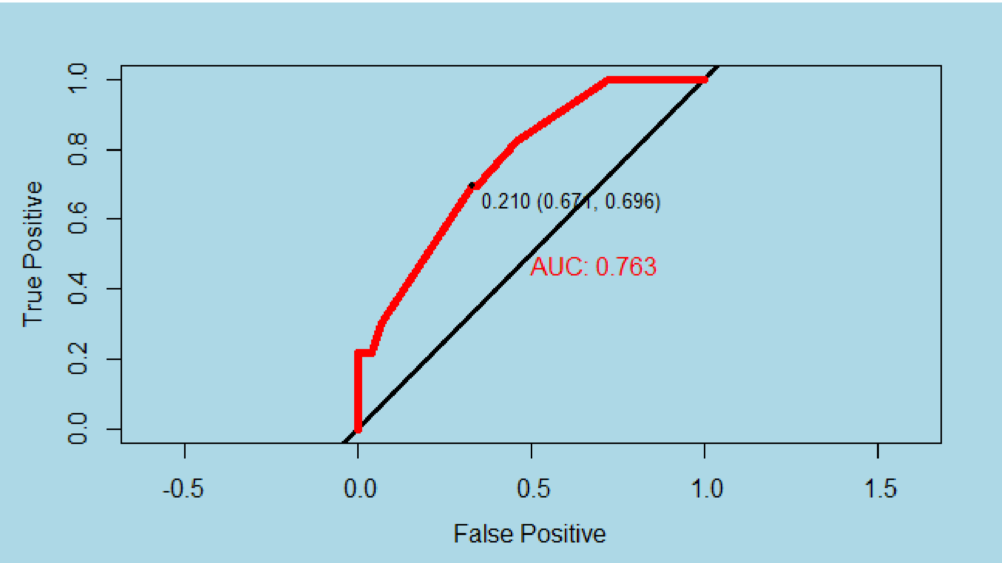
Figure 4: Early warning model Classification Tree




yes
dlognfecre < 0.18
no
wholesale_ < 23
dura_gdp < 5.3
calm
0.00
16%
hous_gdp >= 11
dlognfecre >= 0.22
stress
0.85
4%
stress
0.68
4%
calm
0.13
3%
stress
0.67
2%
wholesale_ < 54
stress
0.67
2%
govt_gdp < 30
stress
0.70
2%
nfecredit_ >= 34
calm
0.06
17%
nfecredit_ < 25
dlognfecre >= -0.058
dlognfecre < 0.051
calm
0.23
9%
nfecredit_ >= 18
equity_ass < 25
govt3 >= 13
calm
0.02
25%
calm
0.00
2%
stress
1.00
3%
calm
0.11
2%
stress
0.77
6%
forex_gdp >= 0.77
calm
0.00
4%
stress
0.78
2%

Post evaluation of the classification tree is given by the **Figure 5**, which depicts the AUROC. The AUROC is recorded at 76.3 per cent, which suggest that classification is able to detect a stress event using the aforementioned variables in 76.3 per cent of the cases. While not necessarily the best performing statistic in terms of is ability to detect stress, its performance is admirable since in roughly three quarters of the cases it able to distinguish a stress from non-stress event. From a policy perspective perhaps, a decision will not be taken immediately using this methodology, but it will alert policy makers of the need to remain vigilant and stand ready to act.

Figure 5: ROC Curve Early Warning Classification Tree



Logit Model

The results from the logit model yield an AUROC with lower performance when compared the decision tree but still able to classify stress in 2/3 (two third) of the cases. The ROC metric is roughly 69.0 per cent, which suggest that the logit model is capable of classifying stress events in 69.0 per cent of the cases using the variables from the random forest (**see Figure 6**). Once again, the credit variables turned up to be quite significant, increasing the odds of a financial stress (**see Table 1**). Household lending turns out to be significant predictor of stress in the banking sector raising questions surrounding increasing lending of banks to the household sector when households are vulnerable to shocks.

However, one particular variable is quite good at increasing the odds of a crisis and that variable is global credit. This suggest that overheating credit at the global level and capital inflows is an important predictor of banking sector stress at 4 to 12 quarters ahead. This suggest that policy makers will do well at monitoring global credit variables and capital inflows in the ECCU. As mentioned before capital flows help to create macroeconomic balances in the economies that once unwound create problems for the economies during the adjustment phase.

Figure 6: ROC Curve Logit Model

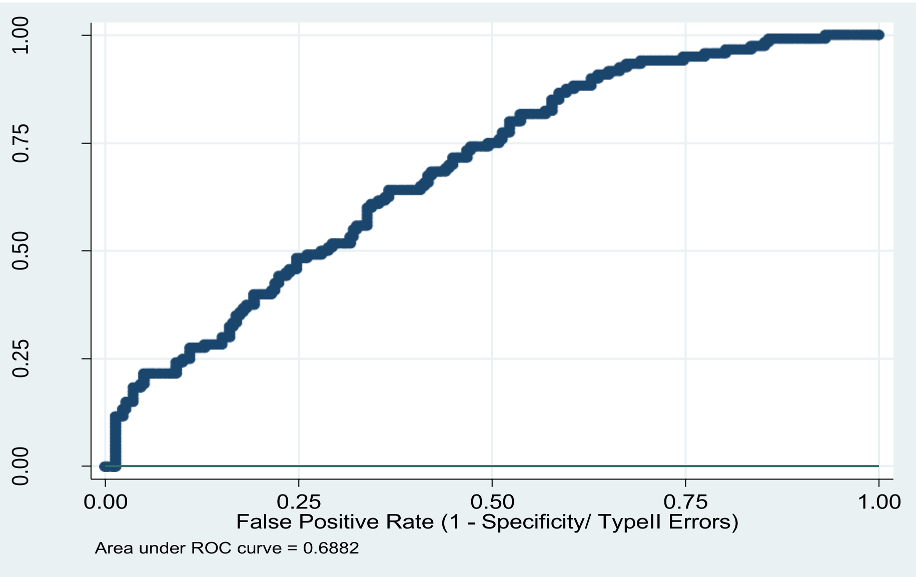


Table 1: Logit Early Warning Model

	(1) Logit Model
crisis4_12	
eff1	-0.0124
	(0.0140)
liq1	0.0197
	(0.0275)
govt1	0.00517***
	(0.00198)
NFE credit growth (yoy)	0.262
	(2.531)
NFE credit growth (yoy)	1.980
	(2.981)
Private credit growth (yoy)	-4.264
	(3.884)

Household credit growth (yoy)	-23.41***
	(4.558)
hcredit_gdp_gap1	0.599***
	(0.152)
Glo. Real credit growth (yoy)	16.33
	(13.30)
Non-core/assets growth(yoy)	11.37***
	(3.121)
Constant	-2.035
	(1.452)
Pseudo R-Squared	0.326
AUROC	0.860
$U_r = \theta = 0.90$	0.440
$TPR = \theta = 0.90$	0.743
$FPR = \theta = 0.90$	0.301
Country Fixed Effects	Yes
Countries	7
Crisis	29
Observations	338

Standard errors in parentheses
***Author's Calculation
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

6.0 Conclusion

As work continues finding appropriate early warning indicators in the macroprudential sphere this paper presents develops an early warning model that helps select potential variables and under what scenarios that these variables can be used in macroprudential supervision /surveillance. The variables which turn as having some significance in the case of the ECCU are the usual credit variables that have been found globally, however in the case of the ECCU it is growth rates that turn out to hold information rather than the credit to GDP gaps. Moreover, global credit cycle is also found to be important, this can be attributed that a heightened state of credit globally can set of stress events in the region due to capital flows. These capital flows are typically associated with rising asset prices and loss in competitiveness due an increase in prices in the non-traded sectors of the economy and upward movement in wages.

In terms of the performance of the models as judged by the ROC, the models do a good job of classifying stress in roughly 70.0 per cent of the time. Though not the best performance it is still good, what this means is that an analyst reporting to the Governor will be correct 70.0 per cent of times is suggesting to the Governor the need to be cautious with the current state of affairs, at least one year in advance of the occurrence of a stress.

Given the results from the model and the current state of the ECCU financial cycle, an incipient stress event appears to be far away.

Recommendations

Given the importance of both household and NFC credit in predicting/classifying banking sector stress a number of recommendations can be put forward to monitor these risks;

- **Firstly,** the derivation of household debt servicing ratio and household debt to disposable income as metrics to monitor household debt. Information to support the construction of this index can be initially gleaned from the National Insurance/Social Security Systems of the region;
- **Secondly,** collecting information on unemployment by sector. This however may be of more medium nature given the lack of / scarcity already of unemployment figures for the region (except Saint Lucia);
- **Thirdly,** private sector non-financial corporations ideally one would want the same debt service indicators, however given that most businesses in the region are privately held this information would be difficult to obtain (though it is likely that information can be collected form the Inland Revenue department). A more realistic approach would be to simply monitor developments in the sector via the national accounts.
- **Fourthly,** developing macroprudential tools, like loan to value (LTV) to help curtail credit growth when its growing too rapidly**.**

References

i. Alessi, Lucia, and Carsten Detken. "Identifying Excessive Credit Growth and Leverage." Journal of Financial Stability, 2017. doi: 10.1016/j.jfs.2017.06.005.

ii. Babecký, Jan, Tomáš Havránek, Jakub Matějů, Marek Rusnák, Kateřina Šmídková, and Bořek Vašíček. "Banking, Debt, and Currency Crises in Developed Countries: Stylized Facts and Early Warning Indicators." Journal of Financial Stability 15 (2014): 1-17. doi: 10.1016/j.jfs.2014.07.001.

iii. Berge, Travis J., and Óscar Jordá. "Evaluating the Classification of Economic Activity into Recessions and Expansions." American Economic Journal: Macroeconomics 3, no. 2 (2011): 246-77. doi:10.1257/mac.3.2.246.

iv. Breiman, Leo Et Al. Classification and Regression Trees. S.L.: Crc Press, 2017.

v. Bussiere, Matthieu, and Marcel Fratzscher. "Towards a New Early Warning System of Financial Crises." Journal of International Money and Finance 25, no. 6 (2006): 953-73. doi: 10.1016/j.jimonfin.2006.07.007.

vi. Chakraborty, Chiranjit and Andreas Joseph. "Machine Learning at Central Banks". Bank of England Working Paper. No. 674 (2017).

vii. Demirgüç-Kunt, Asli, and Enrica Detragiache. "Monitoring Banking Sector Fragility: A Multivariate Logit Approach with an Application to the 1996-97 Banking Crises." Policy Research Working Papers, 1999. doi:10.1596/1813-9450-2085.

viii. Drehmann, Mathias, and Mikael Juselius. "Evaluating Early Warning Indicators of Banking Crises: Satisfying Policy Requirements." International Journal of Forecasting 30, no. 3 (2014): 759-80. doi: 10.1016/j.ijforecast.2013.10.002.

ix. Giese, Julia et al. "The Credit-To-GDP Gap and Complementary Indicators for Macroprudential Policy: Evidence from The U.K." International Journal of Finance and Economics. Vol, Issue 1.

x. Kaminsky, Graciela L., and Carmen M. Reinhart. "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems." American Economic Review 89, no. 3 (1999): 473-500. doi:10.1257/aer.89.3.473

xi. Jordà, Òscar, Moritz Schularick, and Alan Taylor. "Financial Crises, Credit Booms, and External Imbalances: 140 Years of Lessons." IMF Economic Review, 59 (2010): 340-78. doi:10.3386/w16567.

xii. Joy, Mark, Marek Rusnák, Kateřina Šmídková, and Bořek Vašíček. "Banking and Currency Crises: Differential Diagnostics for Developed Countries." International Journal of Finance & Economics 22, no. 1 (2016): 44-67. doi:10.1002/ijfe.1570.

xiii. Manasse, Paolo, Roberto Savona, and Marika Vezzoli. "Rules of Thumb for Banking Crises in Emerging Markets." SSRN Electronic Journal, 2013. doi:10.2139/ssrn.2236733.

xiv. Pattillo, Catherine A., and Andrew Berg. "Are Currency Crises Predictable? a Test." IMF Working Papers 98, no. 154 (1998): 1. doi:10.5089/9781451857207.001.

xv. Tanaka, Katsuyuki, Takuji Kinkyo, and Shigeyuki Hamori. "Random Forests-based Early Warning System for Bank Failures." Economics Letters 148 (2016): 118-21. doi: 10.1016/j.econlet.2016.09.024.

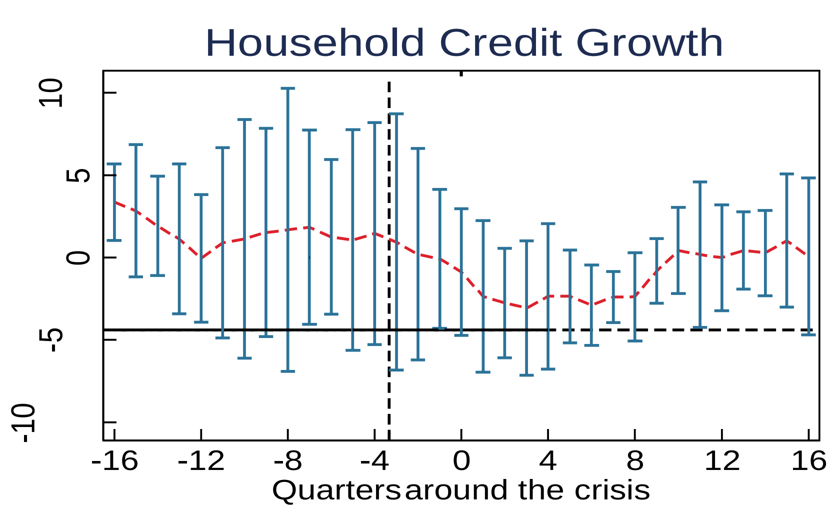
xvi. Phillips, Peter CB, Shu-Ping Shi, and Jun Yu. "Testing for Multiple Bubbles 1: Historical Episodes of Exuberance and Collapse in the S&P 500." International Economic Review 56, no. 4 (November 2015).

xvii. Phillips, Peter CB, Yangru Wu, and Jun Yu. "Explosive Behavior in the 1990s NASDAQ: When Did Exuberance Escalate Asset Values." International Economic Review 52, no. 1 (February 2011): 201-26.

Appendix

Figure 7 displays the behaviour of household credit growth around a stress event, the graph suggests that household credit peaks roughly 8 quarters prior to the onset of a stress event. Given the significant concentration the banking sector loan/lending portfolio to the household sector this renders them highly vulnerable to shocks to households. For example, the closure or lay off by a major corporation in the island can have large reverberations to the banking sector. Therefore, the monitoring of household leverage as suggested by the BIS in term of the debt service ratio and household debt to disposable income therefore become important variables to monitor.

Figure 7: Behaviour of Household Credit around stress



As in the case of the household sector, lending to NFE/C also peaks significantly prior to stress event (see Figure 8). However, the peak is much larger than the growth in household credit. The peak occurs around between 9-11 quarters prior the onset of a stress event.

Figure 8: NFE/C Credit Growth around a stress

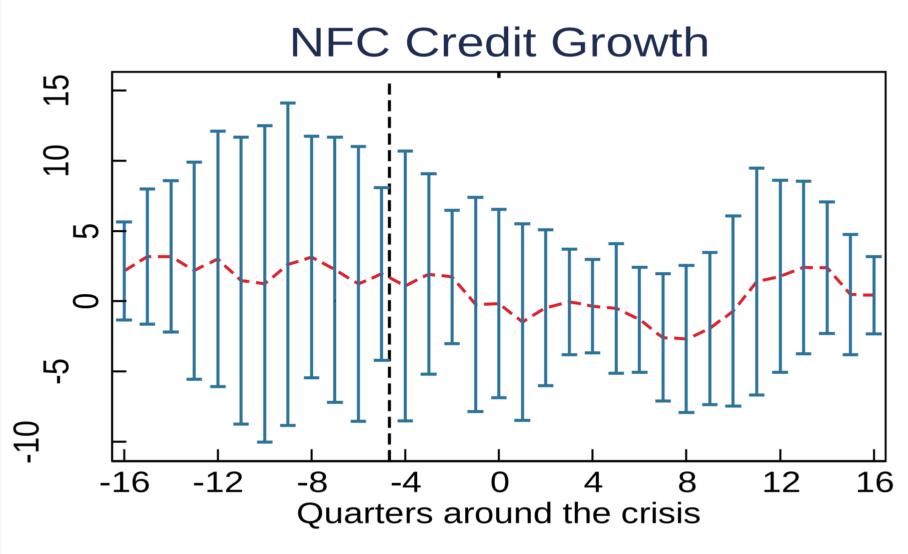


Table 2: A list of indicators.

Variable name	Definition	Transformation and additional info
Private Sector Credit	Total Credit to the private sector, households, private businesses	As per cent GDP y-o-y growth rates Gap using $\lambda = 200{,}000$
Household Credit	Total Credit to Households resident and non-resident	As per cent GDP y-o-y growth rates Gap using $\lambda = 200{,}000$
Non-financial corporation Credit	Total Credit to NFC resident and non-resident	As per cent GDP y-o-y growth rates Gap using $\lambda = 200{,}000$
Foreign Currency Loans	Foreign Currency loans resident and non-residents	As per cent GDP y-o-y growth rates Gap using $\lambda = 200{,}000$
Government Loans	Central Government Credit resident and non-resident	As per cent GDP y-o-y growth rates Gap using $\lambda = 200{,}000$
Credit to Construction Sector		As per cent GDP y-o-y growth rates Gap using $\lambda = 200{,}000$
Credit to Tourism Sector		As per cent GDP y-o-y growth rates Gap using $\lambda = 200{,}000$
Credit for House Construction and Land Purchase		As per cent GDP y-o-y growth rates

		Gap using $\lambda = 200{,}000$
Credit for Consumer Durables		As per cent GDP y-o-y growth rates Gap using $\lambda = 200{,}000$
Net Foreign Assets to GDP	Measure of Risk Taking	Ratio
Hirshman-Herfindahl Index (HHI)	Measure of Concentration	Index
Inflation Rate		Growth rate
Tourism Earnings		Growth rate
GDP growth		Growth rate
Government Debt to GDP		Ratio
Government Overall Balance to GDP		Ratio
M2 to GDP		Ratio
Loans to Deposit Ratio	Measure of Liquidity and Mismatch	Ratio
Risk Weighted Assets to Total Assets	Measure of Risk Taking	Ratio
Global Credit	Measure of Spill over	As per cent GDP y-o-y growth rates Gap using $\lambda = 200{,}000$

Table 3: Results from unit root tests[8]

Variable name	IPS Test[9]	Fisher Type Test	
Private Sector Credit			
As per cent GDP	-0.06(0.47)	9.82(0.78)	
y-o-y growth rates	-0.94(0.17)	16.40(0.30)	
GDP Gap	-1.17(0.12)	24.50(0.04)	

[8] Notes: The table shows results for the Im-Pesaran-Shin (see Im et al. (2003)) and the Fisher-type (see Choi (2001) panel unit-root tests. The table also reports the results from country-specific Augmented Dickey-Fuller tests (see Dickey and Fuller (1979)). For all tests,an initial lag length of 4 of was chosen, and the optimal lag truncation was decided on based on SBC information criteria. Only an intercept was included in the ADF-regressions, and – as a cut-off for the country-specific unit root tests – we used critical values from the Dickey-Fuller distribution consistent with a 5.0% significance level.

[9] H0: all panels have unit root against H1: some panels are stationary

Household Credit			
As per cent GDP	1.45(0.93)	6.53(0.95)	
y-o-y growth rates	-1.20(0.12)	19.17(0.16)	
GDP Gap	-4.20(0.00)	26.97(0.02)	
Non-financial corporation			
As per cent GDP	-0.61(0.27)	11.28(0.66)	
y-o-y growth rates	-2.50(0.00)	15.63(0.34)	
GDP Gap	-2.09(0.02)	24.08(0.04)	
Foreign Currency Loans			
As per cent total loans	-0.6128 0.2700	14.8924 0.3856	
y-o-y growth rates			
Government Loans		24.22(0.04)	
As per cent GDP			
y-o-y growth rates			
Gap using			
Credit to Construction Sector			
As per cent GDP	1.29(0.90)	6.47(0.95)	
y-o-y growth rates	-4.89(0.00)	33.96(0.00)	
Gap	-3.99(0.00)	30.16(0.01)	
Credit to Distributive Sector			
As per cent GDP	-11.70(0.00)	14.05(0.45)	
y-o-y growth rates	-14.71(0.00)	40.415(0.00)	
Gap using	-3.99(0.00)	26.56(0.02)	
Credit for House Construction and Land Purchase			
As per cent GDP	1.21(0.89)	7.94(0.89)	
y-o-y growth rates	-2.04(0.02)	31.03(0.01)	
Gap using	-5.23 (0.00)	37.30(0.00)	
Credit for Consumer Durables			
As per cent GDP	-2.08(0.02)	23.38(0.05)	
y-o-y growth rates	-4.83(0.00)	27.11(0.02)	
Gap using	-3.73(0.00)	38.42(0.00)	
Net Foreign Assets to GDP	0.51(0.70)	8.69(0.85)	
Hirshman-Herfindahl Index (HHI)	3.08(0.90)	8.92(0.84)	
GDP growth	-6.94(0.00)	37.81(0.00)	
Government Debt to GDP			
Government Overall Balance to GDP	-1.66(0.05)	20.07(0.13)	
M2 to GDP	1.14(0.87)	5.82(0.97)	
Loans to Deposit Ratio	1.57(0.94)	6.69(0.95)	
Risk Weighted Assets to Total Assets	-2.97(0.00)	18.38(0.19)	
Global Credit			
As per cent GDP	2.08(0.98)	5.63(0.97)	

y-o-y growth rates	-3.23(0.00)	32.14(0.00)	
Gap	-5.20(0.00)	38.02(0.00)