What should we do about (Macro) Pru?

Macro Prudential Policy and Credit[1].

Ray Barrell and Dilruba Karim, Brunel University.

: Credit growth is widely used as an indicator of potential financial stress, and it plays a role in the new Basel III framework. However, it is not clear how good an indicator it is in markets that have been financially liberalised. We take a sample of 14 OECD countries

Introduction

Many commentators on the financial crises

(2010) against 38% in the BIS study[4]. This superiority also translates to the type II error rate. At best, the model using the credit-to-GDP gap can identify 57% of crises out-of-sample but more than one in three times the signal will be a false alarm. In contrast, an OECD model which excludes credit can correctly predict 75% of crises out-of-sample with comparatively negligible cost: only 6% of signals will be false alarms.

Besides our own estimates, other papers also do not find conclusive evidence for the role of credit growth in generating financial instability. Mendoza and Terrones (2008) found that credit booms often link to banking crises in emerging market economies but less often in OECD countries. In a study of the Euro area and the US, Kaufmann and Valderrama (2007) note that "the mutually reinforcing effects of lending and asset prices contributing to the build-up of financial imbalances during boom periods is not confirmed in our model" for the Euro area[5]. Boyd et al (2001) investigate the behaviour of credit/ GDP ratios in 22 economies that experienced a single banking crisis and find unusual credit growth in only 6 of them whilst in 10 out of 21 economies rapid credit growth was not always followed by a crisis.

Aside from the methodology, the heterogeneous sample in the BIS countercyclical buffer proposal is potentially problematic in its implementation since the same upper and lower buffer thresholds are applied to the OECD countries and to Latin American countries such as Brazil, Argentina and Mexico and Asian countries such as Indonesia[6]. These banking systems operate very differently with OECD countries being more financially liberalised than others. One objective of this paper is to investigate whether the determinants of banking crises differ between the OECD and emerging economies.

Early Warning Systems for Financial Crises

The literature has developed a number of distinctive multivariate Early Warning Systems (EWS) for banking crises, including logit[7] (Demirguc Kunt and Detragiache, 1998; 2005). Non-parametric signal extraction models (Kaminsky and Reinhart, 1999) differ by being univariate. Davis and Karim (2008) show logit to be the best of the three estimators they consider. Hardy and Pasarbasioglu (1999) and Beck et al. (2006) also demonstrate the merits of logit models. Accordingly we will adopt the logit approach to assess the role of credit and will use a binary banking crisis variable (1 for crisis, zero otherwise) based on the dating of Caprio et al. (2003) and Laeven and Valencia (2010).

There are many potential and competing explanations for financial crises, hence it is essential to estimate the effect of credit growth on banking crisis probabilities alongside a set of crisis determinants traditionally deemed important in the literature. This literature comprises two strands: the first class of logit crisis models estimated by Demirguc-Kunt and Detragiache (1998; 2005) and the second class of logit models by Barrell et al. (2010). The latter append new variables to the Demirguc-Kunt and Detragiache set of determinants for the OECD

[4] Even if we allow for the most generous (3 year) horizon, this model calls 18% more crises correctly.
[5] Although reinforcement occurs to an extent in the US market based banking system.
[6] The research behind counter cyclical buffer proposal also included Islamic banking systems (Saudi Arabia) alongside fundamentally different non-Islamic banking systems.
[7] A non-parametric approach, the binary recursive tree, is discussed in Davis and Karim (2008).

1983, Denmark in 1987, the US in 1988, Italy and Norway in 1990, Finland, Sweden and Japan in 1991, France in 1994, whilst in the UK there are crises in 1984, 1991 and 1995. Laeven and Valencia (2010) classified Belgium, Denmark, France, Germany, the Netherlands, Spain and Sweden in crisis by 2008 and the US and UK in 2007. The authors treat the 2008 crisis in the US and the UK as a continuation of 2007 crisis, while we treat it as separate crises since 2008 was induced by the collapse of Lehman Brothers.

We undertake three sets of experiments. The first is designed to directly test the BIS hypothesis on countercyclical buffers and uses the Hodrick Prescott filtered gap between credit and GDP using the same parameters as they do. We then look at the ratio of credit to GDP and then finally the growth in this ratio. We do not include them in the same model in order to clarify their role individually. The results of the sequential elimination process are reported in Table 2. We report on elimination until the variables included all have z statistics that are significant at the conventional 1 step 5% level. However, we should note that we have performed a sequence of tests, and we should be raising our standard in order to take account of this. Hence a probability of 0.116 for GDP growth lagged two periods in the first two models or a probability of 0.109 for the growth

Table 2: OECD General to Specific Estimation, 1980 – 2008.

Regression Number	1	2	3	4	5	6	7	8
Liquidity Ratio(-2)	-0.112 (0.006)	-0.113 (0.006)	-0.11 (0.007)	-0.114 (0.006)	-0.107 (0.007)	-0.126 (0)126 (0)112		

there is no crisis[14] is 35% and the false call rate when there is a crisis[15] is 30%. The overall successful call rate (both crisis and no crisis called correctly) is 66%, with 16 out of the 23 (or 70 %) crisis episodes captured correctly at a cut-off point of 0.061[16]. These results stand up well against the wider literature. For example, Demirgüc-Kunt and Detragiache (2005) had a type II error of 32% and a type I error of 39%, with an overall success rate of 69% at a threshold of 0.05 for their most preferred equation.

During the subprime period there is only one genuine false call in Canada, and a failure to call Germany, where the purchase of low quality US ABS to hold on balance sheet was the

from 1980 – 1989. We use the Bank of International Settlements (BIS) Regulatory Capital Ratios reported by the Banker to construct our regulatory capital variable[18].

The BIS Capital Ratio is a comparable measure across banks that were required to calculate capital adequacy according to BIS rules. However coverage may be an issue because not all banks in our emerging market countries will have entered the top 1000 global bank list. Nevertheless, it is reasonable to assume that where a bank did enter the list, it would have been systemically important (in the "too-big-to-fail" sense) and thus its capital ratio would be correlated with the health of the financial system. Hence although our capital data may not contain all the variance associated with a particular banking system, it should be broadly representative of its capital soundness. From 1998 onwards, we revert to the IMF's Global Financial Stability Reports to obtain capital adequacy ratios for the entire banking system. Like The Banker, these data are risk weighted according to BIS regulatory requirements.

We use the Barrell et al (2010) definition of liquidity and the IMF's International Financial Statistics database to create the variable. This is a narrow liquidity definition because of the exclusion of claims on the private sector. During the Asian crises, capital flight would have

panel we include the growth of credit to GDP and we show 6 sequential variable deletions culminating in a 5 variable equation. We include all three variables from the first model, and also include an indicator of foreign exchange cover which may be relevant in credit constrained countries. Given the discussion above of the impact of a sequence of tests on significance levels it would be possible to eliminate the current account from this sample as well, making the separation of causes much clearer between the two groups of countries. However, we leave the variable in our final equation. The variable deletions themselves are of interest since they suggest changes in GDP growth, inflation, domestic credit and the exchange rate do not significantly affect crisis probabilities. The third specification in its final form is able to identify 71% of crises at a threshold based on the sample proportion of period where crises start. This is associated with a cost of 36% false alarms so that our emerging market model marginally outperforms the OECD model in terms of crisis prediction but is fractionally worse in terms of false alarms (36% as opposed to 35%). We should note that three of the 'unforeseen' crises occurred in Argentina where the factor driving problems were often political not economic[19]. We do not replicate table 3 as the models all arrive at different solutions, and the ROC curve analysis below helps us arbitrate between them

Table 4: Latin America and Asia General to Specific Estimation, 1980 – 2008.

	1	2	3	4	5	6	7	8
Liquidity Ratio(-2)	-0.054	-0.055	-0.049	-0.049	-0.048	-0.048	-0.048	-0.054
	(0.001)	(0.001)	(0.001)	(0)	(0)	(0)	(0)	(0)
Capital Adequacy Ratio(-2)	-0.176	-0.175	-0.213	-0.226	-0.224	-0.227	-0.242	-0.249
	(0.003)	(0.002)	(0)	(0)	(0)	(0)	(0)	(0)
Current Account Balance (% of GDP)(-2)	-0.095	-0.094	-0.082	-0.079	-0.08	-0.078	-0.07	-0.08
	(0.048)	(0.042)	(0.06)	(0.067)	(0.063)	(0.068)	(0.084)	(0.057)
Exchange Rate(-2)	0	0	-0.001	-0.001	-0.001	-0.001	-0.001	
	(0.285)	(0.283)	(0.236)	(0.217)	(0.216)	(0.209)	(0.176)	
ΔGDP(-2)	-0.054	-0.053	-0.034	-0.034				
	(0.306)	(0.29)	(0.486)	(0.4003>[illegible].335631.3332TD0Tc((0)74.2(.40003>6.8<000300030003				

Regression Number	1	2	3	4	5	6	7
Liquidity Ratio(-2)	-0.053 (0.002)	-0.053 (0.002)	-0.047 (0.002)	-0.047 (0.002)	-0.046 (0.002)	-0.049 (0)	-0.052 (0)
Domestic Credit/ GDP(-2)	-0.019	79 ظ	Amqì r0.047		ێ	۶	ك

Am' DYA3sÐÀ7 0. Amp0$,$ ÎBÐ0Ê2•F8 £/DYA3xb 0 0

The second specification suggests that the ratio of credit to GDP may be important in these countries. The final specification suggests the most important determinants of combined Latin American and Asian crises are: changes in domestic credit/ GDP, bank capital adequacy and liquidity and adds in the ratio of M2 to foreign reserves and the current account balance. An improvement in the M2 to reserves ratio, the capital and liquidity soundness of banks and the current account reduces the likelihood of systemic bank failures while an increase domestic credit relative to GDP raises the failure probability. This latter result is significant in terms of our objectives, as this variable was eliminated in the OECD sample. It suggests that curbing the growth in credit to GDP may have some benefits in emerging markets that have been financially liberalised more recently than the OECD.

Credit Constraints, Financial Liberalisation and the Policymaker's Options

The level or growth of the ratio of credit to GDP appears to be a significant determinant of crises in Latin America and East Asia, but it does not influence the probability of a crisis in OECD countries. In general we may say that OECD financial markets have been largely deregulated in the last 25 years, and hence there have been few constraints on borrowing. Barrell and Davis (2007) look at the impact of financial liberalisation on consumption and generally conclude that it was removed by the mid 1980s, and perhaps a little later in some Scandinavian countries. They give the Swedish liberalisation date as 1985 and Abaid et al (2008) show that although there was also a round of liberalisation in Finland in the mid 1980s, financial liberalisation actually peaked in 1993. Jonung (2008) notes how liberalisation in these economies fundamentally affected credit availability. The financial markets in our sample of East Asian and Latin American economies still exhibit significant ggests ycnre have 5ak.2(stemic ohoro wd[strug.ro)]TJ98.17 0 TD.0005 T5.0745 Trecerief andce's Optpact u

or diagnosing it once symptoms have manifested. In the former case, the model's ability to discriminate between patient classes is important, but this metric may not yield the best diagnostic model (Cook, 2007). The latter leads to curative care which may be costly, hence the value of many different indicators should be examined to make a diagnosis. Conversely, for relatively little cost, the policymaker may wish to immunise the population against the disease. The trade-off between type I and II errors becomes important because a model which raises crisis prediction accuracy necessarily generates a higher false call rate and could elicit unnecessary costs of intervention. Hence an immunisation model should optimise the trade-off between model accuracy and the number of instruments; if the set of instruments can be reduced without compromising the informational content of a model then the toolkit becomes simpler and less costly. Immunisation models can be selected on the basis of their Receiver Operating Curve characteristics which we discuss next.

Model Selection and the use of ROC Curves

Receiver operating characteristic (ROC) curves test the "skill" of binary classifiers and hence can be used to discriminate between competing models. In the context of logit estimators, probabilistic forecasts can be classified for accuracy against a continuum of thresholds. This generates a true positive rate and true negative

at the cost of high false positive rates may lead to "tail events" being missed with commensurate economic costs.

Since the true positive and false positive rates are functions of the threshold, a policy makers' risk attitude to crises may influence the choice of threshold and thus optimal model. Moreover once this optimal threshold is selected, an increase or decrease in the prevalence of crises will not affect the true positive or false negative rates. Thus the ranking of models based on ROC curves will vary depending on the chosen threshold range which in turn is a function of the policy maker's preferences.

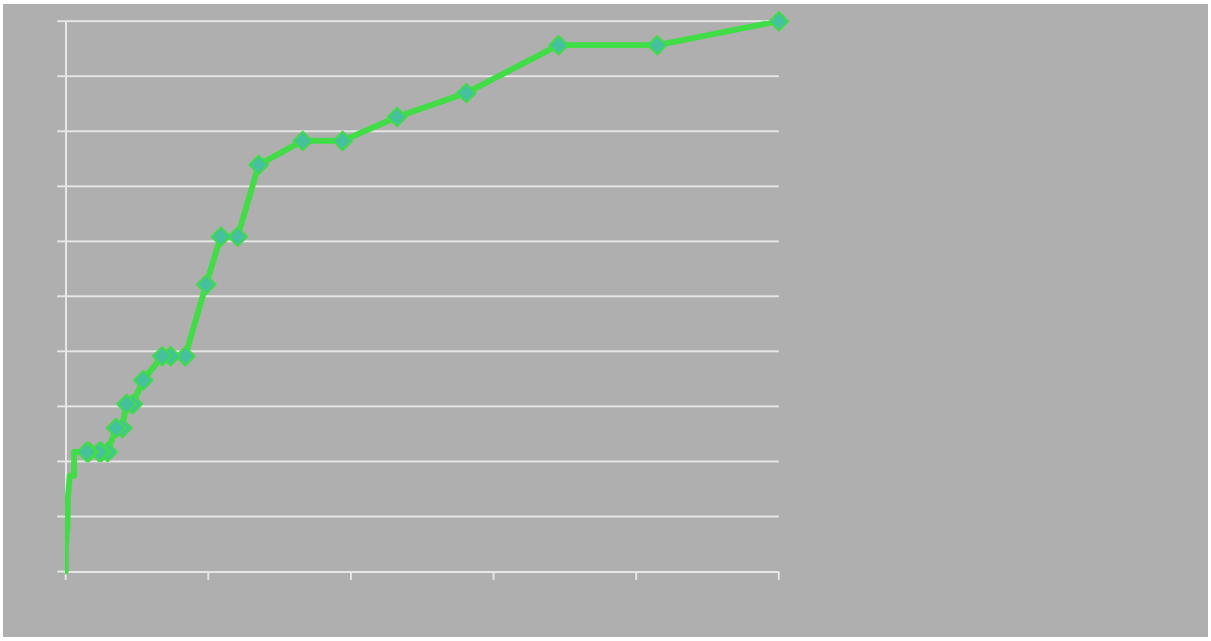
To separate out preferences from the decision making process, an alternative but related "global" measure of model skill can be used to select between competing models: the Area Under the Curve (AUC). If the true positive rate declines more slowly than the false positive rate when the threshold is raised then the AUC is above a half. The larger the difference between these two rates of decline the higher the AUC. This avoids evaluating or the ranking of models at particular thresholds. An AUC of 0.5 is equivalent to a "naïve" estimator that replicates a random coin toss (corresponding to the 45^0 line) so an AUC above 0.5 implies the model adds value in terms of the ability to call crises correctly with low false negative rates.

Figure 1: Receiver Operating Characteristic Curves

Table 6: Area Under the Curve (AUC) and model skill

AUC = 0.5	No discrimination (equivalent to coin toss)
$0.7 \leq AUC < 0.8$	Acceptable discrimination
$0.8 \leq AUC < 0.9$	Excellent discrimination
$AUC \geq 0.9$	Outstanding discrimination (not possible in logit frameworks)

Table 6 indicates discrimination performance in terms of the AUC. Hosmer and Lemeshow (2000) indicate that an AUC ≥ 0.9 is highly improbable for logit models since this level of discrimination would require complete separation of the crisis and non-crisis event and the logit coefficients could not be estimated. Hence for our EWS approach we would accept models with AUCs ≥ 0.7. The AUCs for our competing models are given in Table 7 whilst



We plot ROC curves for each of the three models for the OECD countries and then for the Emerging Markets. In each case we plot a ROC curve for each step in the elimination. If one ROC is inside another then it is clearly dominated by the one further out, which has a lower generalised Signal to Noise Ratio. However, it is not always the case that it is clear which model dominates. At low cut off thresholds the 'final' model is not as good as the others, but as the threshold rises its relative performance improves. A more complicated method of judging overall signal to noise ratios is clearly required, and this is reported in table 6 where we have the sequence of AUCs for the three OECD models. As they all end up with the same model they have the same AUC at an acceptable level of 72, and in each case the addition of the penultimate variable raises the AUC to 74. In two cases this is the growth of GDP, and in the third it is the growth of credit to GDP (which is related). The AUC is unchanged when we remove the credit gap in the first experiment and the ratio of credit to GDP in the second and hence their generalised signal to noise ratio is low. Only in the case of the growth of credit to GDP is there an impact on the AUC, and a case may be made for monitoring it in the OECD. However, Barrell et al (2012) show that either the increase in off balance sheet activity or the growth of house prices (lagged 3 periods) is superior to this variable.

The ROC curves for Latin America and East Asia look generally much flatter than do those for the OECD countries, and the models are disparate in their end points. In general within each experiment no model is dominant at all thresholds, with ROC curves crossing, and the final parsimonious model does not look in any way worse than the steps toward it. As we can see from Table 7 the model with the credit to GDP gap in it at the start has AUC values that are very low, so the generalised signal to noise ratio is low with this variable. At the point where the gap drops out the AUC is 59 and falls to 58. If we include the ratio of credit to GDP this stays the model, and the AUC hovers around 63, which is lower than really acceptable, and not greatly better than coin toss. The model with the growth of credit to GDP

is noticeably better, with an AUC of around 70, and this is at a relative maximum when we reach the most parsimonious model.

Figure 3: ROC Curves for the Latin America and Asia Models

Table 7: Area Under the Curve (AUC) for general to specific estimations

Regression Number	1	2	3	4	5	6	7	8
OECD	0.76	0.76	0.76	0.75	0.75	0.74	0.74	0.72
OECD	0.76	0.76	0.75	0.75	0.75	0.74	0.74	0.72
OECD	0.76	0.76	0.76	0.75	0.75	0.74	0.74	0.72
LA and EA	0.63	0.63	0.60	0.60	0.59	0.58	0.59	NA
LA and EA	0.64	0.64	0.62	0.62	0.63	0.63	0.63	NA
LA and EA	0.69	0.70	0.70	0.69	0.69	0.70	NA	NA

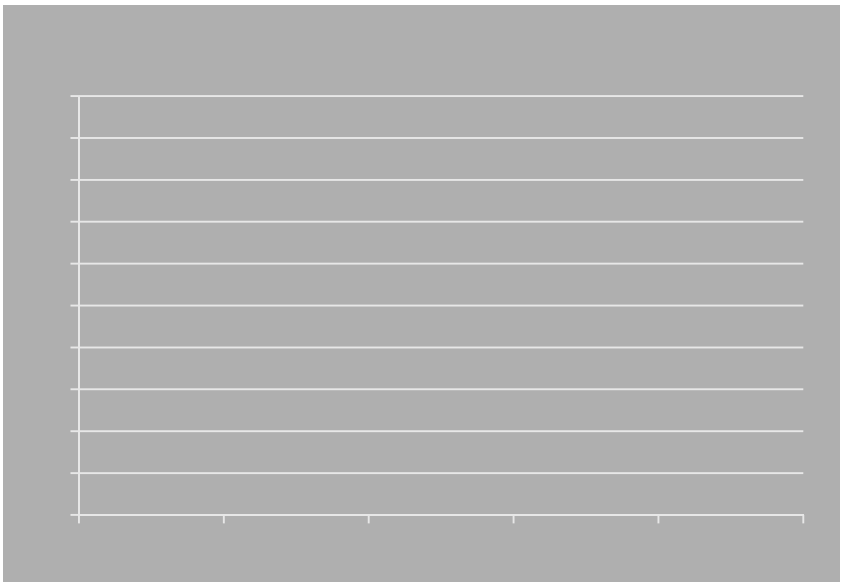
Note: NA indicates not applicable

If we wished to immunise the financial system against crises we would select the variables in the final models as candidates for further investigation. This suggests that credit based countercyclical buffers may, subject to further investigation, have a regulatory role in Latin

imbalances to reduce the risk of financial crises. This would mean raising capital in a calibrated way in response to an increased current account deficit, and releasing the capital slowly some years after the deficit has disappeared. We know from Barrell et al 2012 that responding to property price booms and changes in off balance sheet activity would also be good countercyclical buffer triggers. There might also be a case for responding to the growth of credit to GDP, but this has to be seen as a second order response. It is far more important in emerging markets where there are credit constraints, where the risks of sharp currency movements that might flow from inadequate reserves should also indicate that the ratio of M2 to reserves should act as a trigger. In no case can we see a role for the credit to GDP gap, despite its prominence in BIS work and its role in current legislation.

Forecast Evaluation

Although ROC curves allow us to select amongst competing hypotheses for crises determinants, the chosen model should be checked for out of sample robustness. We therefore conduct forecast tests for 2009-12 using data from 2007 to assess our model in terms of its ability to identify crises and false alarms. We use Laeven and Valencia (2012) to date crises in the out of sample period, and they note crises in Denmark (2009), Germany (2009), and Spain (2011).



Panel 4 presents the forecast ROC for the OECD and the corresponding AUC. The model performs well out of sample and all three crises are called at the in-sample threshold; reduction in performance arises due to false alarms in Denmark (2008) and Spain (2008). There were no crises in Latin America and Asia in the forecast period which eliminates one dimension (sensitivity) of the ROC curve. Hence we examine the predicted crisis probabilities for the forecast period to check for false alarms, and they occur only around 2008-9 in Chile, Panama and South Korea if we use the in-sample threshold. This additional robustness check reiterates our original conclusions: there appears to be no significant role for credit aggregates, including the credit to GDP gap, in OECD crises over the past three

decades. For Latin American and Asian crises credit to GDP ratios may be a determinant but this should not warrant countercyclical buffer provisions[22].

Conclusion

We have constructed early warning systems for the OECD and emerging markets using variables that can be directly influenced by policy makers in the latter for the first time. We then test for the crisis inducing role of credit in both regions. In contrast to previous work, we include all variations of credit that have been cited in the literature to comprehensively

effect of credit growth is relegated below other determinants. Such measures were already in place under Basel II and have been strengthened under Basel III.

It is thus possible that conditioning bank capital on credit growth alone may not avert future crises in financially liberalised economies especially when these are driven by property prices in an otherwise benign environment. Under these circumstances, countercyclical buffers may not accumulate because business lending continues in line with GDP growth but risky lending may continue in the housing markets or commercial property markets. It is also clear from our work above, that we should provision against the current account, and that the triggers for building capital buffers should include different variables in liberalised and un-liberalised financial markets. In particular there may be a role for the credit related buffer and for monitoring foreign exchange reserves in emerging market economies whilst they remain unliberalised. Judging the tools is the same as judging the economy.

Therefore, more work is required on the links between credit cycles and property prices in both regions since as we have already shown (Barrell et. al., 2010) residential property price growth outperforms credit as a crisis determinant on the OECD. Strong house price appreciation is currently a concern in many emerging market economies and some of these now have fully liberalised, globalised financial systems. Unless the dynamics of property prices and their relation to credit growth are properly examined, the latest generation of banking reforms may not be sufficient to ensure future financial stability.

References

Abaid, A., E. Detragiache and T.Tressel (2008), “A New Database of Financial Reforms”, IMF Working Paper No. 08/266, Washington, DC.

Aiyar, S, Calomiris, C and Wieladek, T (

Brunnermeier, Markus K., Crockett Andrew, Goodhart ,Charles, Persaud, Avi, and Shin, Hyun (2009) “ ”, Geneva London,

Caprio, Gerard, Klingebiel, Daniela, Laeven, Luc and Noguera, Guillermo (2003). “Banking Crisis Database” in Honohan, Patric and Laevan, Luc (eds), “Systemic Financial Crises: Containment and Resolution”, The World Bank, pp 307-340.

Cook, Nancy (2007). “Use and Misuse of Receiver Operating Characteristic Curve in Risk Prediction”, Journal of the American Heart Association, 115:928-935.

Davis, E. P. and Karim, D. (2008). “Comparing Early Warning Systems for Banking Crises”, , Vol. 4, pp. 89 – 120.

Davis, E. P., Karim, D. and Liadze, I., (2011). “Should Multivariate Early Warning Systems For Banking Crises Pool Across Regions?”. , Vol. 147, No. 4, pp. 693-716.

Demirgüç-Kunt A. and Detragiache E., 2005. Cross-country empirical studies of systemic bank distress: a survey. 05/96.]

Demirguc-Kunt, A. and Detragiache, E., 1998. The determinants of banking crises in developed and developing countries. , Vol. 45, no.1.

Eichberger, Jürgen and Summer, Martin (2005). "Bank Capital, Liquidity, and Systemic Risk," , MIT Press, vol. 3(2-3), pages 547-555.

Francis, W and Osborne, M (2012), ‘Capital requirements and bank behaviour in the UK: are there lessons for international capital standards?’, , Vol. 36, No. 3, pages 803–16.

Freixas, Xavier and Rochet, Jean-Charles (2008). " ," MIT Press Books, The MIT Press.

Giese, Julia, Andersen, Henrik, Bush, Oliver, Castro, Christian, Farag, Marc, and Kapadia, Sujit (2013), “The credit-to-GDP gap and complementary indicators for macroprudential policy: Evidence from the UK”. Bank of England Mimeo.

Hardy, D. C; Pazarbasioglu, C (1999), “Leading Indicators of Banking Crises: Was Asia Different?”, , 46, 247-258

Hosmer, D.W. and Lemeshow, S. (2000). "Applied Logistic Regression". 2nd ed. John Wiley and Sons . 156-164

Jonung, L., (2008) ‘Lessons from Financial Liberalisation in Scandinavia’ , Vol. 50, Issue 4, pp. 564-598, 2008

Laeven, L., and Valencia, F., (2007). Systemic banking crises: A new database. No. WP/08/224

Laeven, L., and Valencia, F., (2012). Resolution of banking crises. No. WP/12/146

Laeven, L., and Valencia, F.,(2010), Resolution of Banking Crises: The Good, the Bad, and the Ugly, 10/146.

Mendoza, E. G. and Terrones, M. E., (2008). “An anatomy of credit booms: evidence from macro aggregates and micro data”. , no.14049.

Metz, Charles, E. (2006). “Receiver Operating Characteristic Analysis: A Tool for the Quantitative Evaluation of Observer Performance and Imaging Systems”. , Volume 3, Issue 6, June 2006, Pages 413-422

Morrison, A., and L. White (2005). “Crises and Capital Requirements in Banking,” 95, 1548-1572.

.Reinhart, M.C., and. Rogoff. S. K., (2009) “ Re, and LSarien, EJ, (2012).