BANK OF ENGLAND

External MPC Unit

Discussion Paper No. 33

Banking crises and recessions: what can leading indicators tell us?

Matthew Corder and Martin Weale

September 2011

This document is written by the External MPC Unit of the Bank of England

BANK OF ENGLAND

External MPC Unit

Discussion Paper No. 33

Banking crises and recessions: what can leading indicators tell us?

Matthew Corder[(1)] and Martin Weale[(2)]

Abstract

It is widely suggested that there is some relationship between banking crises and recessions. We assess whether there is evidence for interdependency between recessions and banking crises using both non-parametric tests and unconditional bivariate probit models and find strong evidence for interdependence. We then consider whether leading indicators can help predict banking crises and recessions and if these variables can explain the previously observed interdependence. Inclusion of exogenous variables means that the observed interdependence between banking crises and recessions disappears — indicating that the observed interdependence is a result of easily observable common causes rather than unobserved links.

Key words: Crises, recessions, interdependency, bivariate probit analysis.

JEL classification: E37, G21.

(1) Bank of England. Email: matthew.corder@bankofengland.co.uk
(2) National Institute of Economic and Social Research and Bank of England. Email: martin.weale@bankofengland.co.uk

Information on the External MPC Unit Discussion Papers can be found at
www.bankofengland.co.uk/publications/externalmpcpapers/index.htm

External MPC Unit, Bank of England, Threadneedle Street, London, EC2R 8AH

ISSN 1748-6203 (on-line)

Contents

Summary

Policymakers need to be aware of the risk of extreme events such as recessions and banking crises. Given the future is uncertain, monetary policy makers need to consider the full range of risks around their forecasts when considering policy, including the probability of recession. Financial stability policy makers are by definition concerned with low probability events that could lead to banking crises.

The recent banking crisis and recession have raised the question of how closely the two events are linked, and whether a consideration of such interlinkages could have helped forecast these events better. There is a wide array of previous work that has considered ways to predict either banking crises or recessions; this paper considers whether estimating both together and allowing for interlinkages between the two events improves modelling performance.

Tests suggest there is evidence for correlation between banking crises and recessions, but this is the result of common observable causes rather than direct causal links. The probability that a country will be in recession is much higher in the year after a banking crisis than is normally the case. But, after accounting for external indicators, a bivariate probit model (which jointly estimates the probability of two correlated, separate discrete events) finds no evidence of correlation between banking crises and recessions. On the basis of previous work, this paper looks at an array of possible indicators of both banking crises and recession. Using a pooled sample of data since 1981 for fourteen advanced economies, including the United Kingdom, France, Germany and the United States, this paper finds that banking sector capital ratios, the deviation of the credit-to-GDP ratio from trend and the current account deficit are useful predictors of banking crises. The first principal component (a summary statistic of common trends in a selection of variables) derived from OECD leading indicators of GDP growth helps predict recessions, as do movements in real house and equity price inflation, and the current account deficit.

Like all forecasting models, the model presented here is prone to mistakes – tending to over-predict recessions and banking crises. In the period over which the model is estimated – 1981-2005 – up to three-quarters of the model's year-ahead predictions of recession were false alarms, while up to 90% of the predictions that a banking crises will occur were incorrect. The model's out-of-sample performance (post-2005) is substantially better, but this indicates how unusual the past few years are, and is consistent with tests which suggest some model coefficients changed after 2005 (possibly reflecting the fact that increased global financial and trade inter-linkages increased the risk of banking crises and recessions in all countries through contagion).

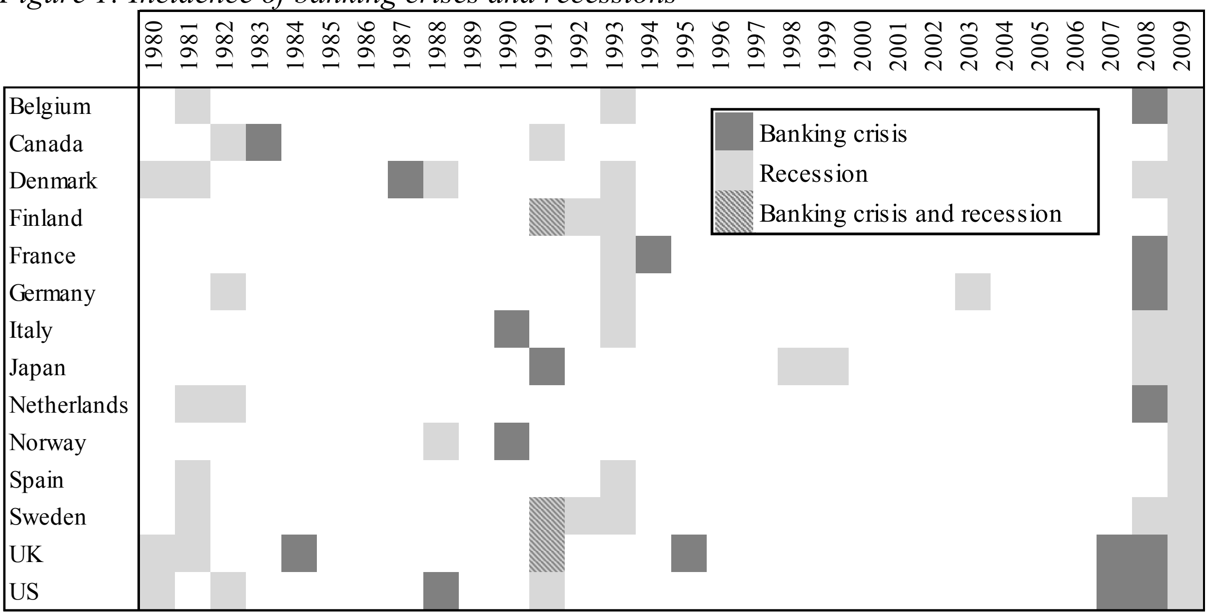
Even though the precise model predictions are imperfect, the model does still provide policymakers with useful information. For example, the model points to the increased risk of a banking crisis in the United Kingdom in the mid-to-late 2000s and that this risk was higher than for many other advanced economies. The model also corroborates the risk of a contraction in 2009 but probably identifies it no earlier than did forecasters using existing techniques. It may therefore be a useful tool to include in the suite of indicators used by policymakers.

1 Introduction

The banking crisis that gathered force after the failure of Northern Rock and reached its maximum intensity in the aftermath of the failure of Lehman Brothers was largely unanticipated. Equally it is fair to say that the recession that began in 2008 but did not appear in annual data until 2009 was largely unforeseen. Inevitably, after such an experience people wonder how closely banking crises and recessions are linked, and whether a consideration of such interlinkages could have helped forecast these events better.

In this paper we bring together two hitherto separate strands of work; the first looks at the ability of a range of indicators to predict banking crises, and the second looks at the role of leading indicators of recessions. It is widely suggested that there is some relationship between banking crises and recessions. A casual examination of the incidence of banking crises and recessions suggests they often cluster together; in some cases occurring in the same year, in other cases recession occur shortly after banking crises or *vice versa* (Figure 1).

Figure 1: Incidence of banking crises and recessions [(a)(b)]



(a) Recession defined as a fall in the level of real GDP relative to the previous calendar year.
(b) Incidents of banking crises taken from Barrell et al (2010).

If banking crises and recessions are interdependent, then one might expect that a joint analysis of the two could improve the capacity to predict both events. We first assess whether there is evidence for interdependency between recessions and banking crises using both non-parametric tests and unconditional bivariate probit models. We find strong evidence for interdependence. We then consider whether leading indicators can help predict banking crises and recessions and if these variables can explain the previously observed interdependence. We find that several leading indicators commonly used in the literature can help predict banking crises and recessions, and that one – the change in the current account balance – has predictive power for both banking crises and recessions. Once these exogenous variables are included we find no residual correlation between banking crises and recessions – indicating that the observed interdependence is a result of observable common causes rather than unobservable links.

We begin this paper by reviewing the literature on predicting crises and recessions using indicator variables. In section 3 we then discuss our data, while in section 4 we present unconditional analysis of the relationship between banking crises and recessions. Section 5 presents conditional indicator models of both events estimated jointly for the period from 1981 to 2005 across a range of OECD countries. In section 6 we review their performance out of sample as tools to predict the banking crises and recessions of 2008/9. We also examine the effect of using different thresholds to differentiate predicted outcomes (ie how strong a signal is required before a policymaker should anticipate a recession or banking crises).

2 Leading indicators of banking crises and recessions

Banking crises are discrete events and work on assessing the capacity of indicators to anticipate them is, not surprisingly, structured around logit or probit analysis. Recessions are similarly discrete events but there is a question whether they are better forecast in a similar probit framework or whether it is better to work from a linear or non-linear framework for forecasting output growth.

The first systematic set of current and leading indicators of the state of an economy was developed by Burns and Mitchell (1946) for the United States. Their focus was on identifying turning points in the economy cycle, possibly because at the time of their study national accounts were in their infancy and there were not, therefore, any generally agreed and understood high-frequency continuous variables, such as real GDP, indicating economic activity. However, in the decades since their work, many researchers have continued to focus on the question of identifying turning points in the economy cycle as distinct from the question of predicting movements in GDP. The OECD, for example, continues to produce leading indicators for its member states, which are not intended to predict movements in real GDP but are intended to identify turning points in the economic cycle.

Marcellino (2006) illustrates the ability of this index together with two other leading indicators for the United States, those produced by the Conference Board and the Economic Cycle Research Institute (ECRI) to anticipate the start and the end of recessions as identified by the National Bureau of Economic Research. He notes that the indicators anticipate the beginnings of recessions by about nine to ten months, with the Conference Board indicator showing a similar lead ahead of the ends of recessions but the OECD and ECRI indicators indicating the ends of recessions with leads of only three to four months. Marcellino notes, however, that all these indices are calculated with the benefit of hindsight; his analysis is therefore indicative of performance in-sample rather than a good indication of how well the indicators might operate in real time. Camba-Mendez, Kapetanios, Smith and Weale (2002) suggested that, for France, Germany, Italy and the United Kingdom, the OECD leading indicators in real time did not have the ability to contribute significantly to predictions of GDP growth; they did not explore, however, their ability to predict cyclical turning points.

The OECD indicator is produced by detrending each of its components, dividing each by its standard deviation and then taking the arithmetic mean. However an alternative approach, due to Rhodes (1937), involves the use of the first principal component as a summary a group of indicator variables. Stock and Watson (2002) set out the statistical properties of such an indicator when some of the variables in question are stationary and others include unit roots. Obviously the principal component can be used instead of the composite indicator produced by any of the above bodies as an explanatory variable in an equation designed to explain either the risk of a banking crisis or the risk of a recession.

A range of other authors have looked at the ability of leading indicators to anticipate either recessions or turning points using probit or logit models. Studies include Estrella and Mishkin (1997) and Birchenhall, Jensen, Osborn and Simpson (1999); little further work has, however, been published on the topic since the start of the current century.

Studies of the predictability of banking crises have used a similar structure (ie probit and logit models), but have used a range of indicator variables instead of giving a role to composite leading indicators. Bell and Pain (2000) provide a summary of much of the literature following the Asian banking crises and outline several theoretical explanations for banking crises.

As banks are likely to face financial difficulty when the value of their assets falls relative to their liabilities any shocks that affect the probability of default of a large number of borrowers or lead to dislocation in asset markets are likely to be associated with bank distress. If banks do not correctly measure their exposure to interest rate risk and hedge themselves accordingly then shocks to the yield curve can also lead to wide-scale banking sector problems (as happened in the case of Savings and Loans institutions in the United States in the 1980s). This suggests that measures of output, employment and interest rates may help predict banking crises, while banking sector capital ratios should indicate the general vulnerability of the banking system to shocks. High inflation has also been suggested as a factor behind banking sector failures as it has the potential to depress the real interest rate, increasing the demand for borrowing and potentially making it harder for banks to sift out riskier borrowers. High inflation can also be associated with increased economic uncertainty if accompanied by large relative price changes.

Lending booms can alter the short-term probability of success of some investments, as borrowers bid up the price of, for example, land and properties and the resulting increase in wealth raises aggregate demand. This process can make it harder for banks to assess which projects have a poor long-term chance of success and can leave banks more vulnerable to deterioration in the health of their borrowers. Credit growth and property prices are therefore often included in models of banking crises. For example, Schularick and Taylor (2009) use data for 14 advanced economies back to the 1870s and find that credit growth relative to GDP is a strong predictor of financial crises.

In addition to the risk of insolvency, banks are also exposed to liquidity risk. If depositors fear that other depositors will withdraw their savings and make a bank illiquid then they may also withdraw their savings leading to a bank run (Diamond and Dybvig (1983)). This coordination failure can be overcome through deposit insurance (which negates the risk that one depositor will lose money if the other depositors withdrawals make a bank illiquid), but if the premia charged to banks do not accurately reflect the riskiness of their portfolios then this can lead them to price risk incorrectly, increasing the vulnerability of the banking system. The risk of bank runs may be heightened if banks share common business models. If one bank fails as a result of an idiosyncratic shock, then depositors may withdraw funds from other banks they perceive as similar, if information asymmetries mean they are unable to tell whether the original shock was truly idiosyncratic. Contagion can also arise through direct exposures, either because banks suffer credit losses on interbank loans or because uncertainty about interbank exposures leads to an effective closure of interbank markets as banks try to reduce their exposure to suspect banks. This suggests that indicators of low banking sector liquidity, similarity in the structures of banks' loan books, and linkages between banks may be good indicators of increased vulnerability to crises.

Barrell et al (2010a) suggest that current account deficits may increase the risk of banking crises through their influence on some of the triggers discussed above. The monetary inflows that accompany deficits may enable banks to expand credit excessively, while potentially exposing them to volatile international wholesale markets. Such excess lending may lead to an overheating economy and unsustainable rises in asset prices.

Aside from the possible combination of vulnerabilities and triggers outlined above, some papers (eg Azariadis (1981), and Cass and Shell (1983)) suggest that some self-fulfilling bank runs can occur because of random trigger events. Such banking crises would be unpredictable and would suggest there is some residual probability of banking crises even after all identifiable vulnerabilities and trigger events are accounted for.

While there are papers that have previously considered the links between banking and currency crises (eg Falcetti and Tudela (2008)) we have not identified other work that considers the interdependence between bank crises and recessions.

3 Data and Methodology

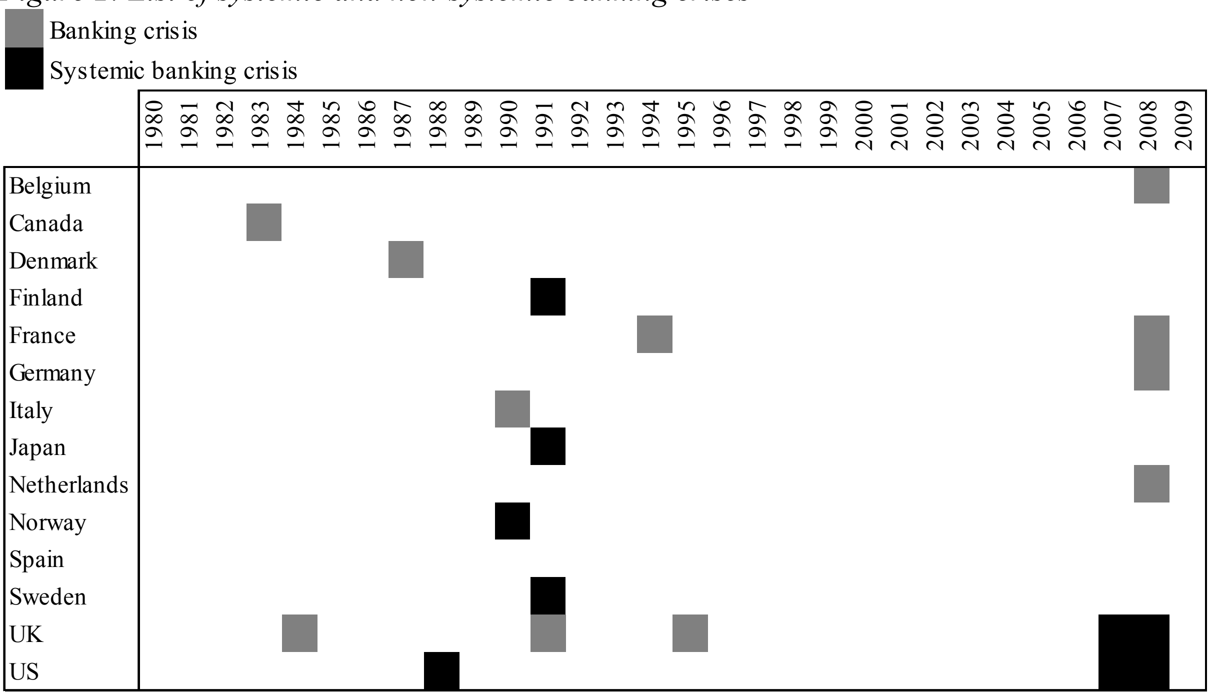
Data on banking crises are taken from Barrell et al (2010a), which are themselves mainly sourced from the World Bank database of banking crises and the IMF Financial Crisis Episode database. The full dataset of crises used in this study is shown in Figure 2. We define recessions as years in which real GDP falls relative to the previous year. This definition was chosen to match our use of annual data in the model, at the cost of omitting some periods of quarterly recession from the sample(1).

We restrict our analysis to fourteen OECD countries for which there are suitable data. Previous work has suggested that banking crises in developed economies have determinants distinct from those of emerging and developing economies (eg Barrell et al (2009) and (2010b)) and we wished to focus on the policy implications for developed economies.

Between 1981 and 2009 there were 20 banking crises (including nine classified as systemic) and 47 cases of year-on-year falls in output.

(1) Of course two successive years with falling output could be regarded as a single recession.

Figure 2: List of systemic and non-systemic banking crises



Source: Barrell et al (2010a).

We use a bivariate probit model, a generalisation of the standard probit model that allows for correlation between outcomes to be jointly estimated. We set $y^*_{it} = 1$ if there is a banking crisis and $y^*_{it} = 0$ otherwise; and $z^*_{it} = 1$ if output falls in country *i* in year *t* and $z^*_{it} = 0$ if it does not. The general specification of our bivariate model is:

$$y_{it} = a_y + b_y x_{it} + \sum_k^k c_{yk} y^*_{it-k} + \sum_1^k d_{yk} z^*_{it-k} + \varepsilon_{yit} \,, \qquad y^*_{it} = 1 \text{ if } y_{it} > 0, 0 \text{ otherwise}$$

$$z_{it} = a_z + b_z x_{it} + \sum_1^k c_{zk} y^*_{it-k} + \sum_1^k d_{zk} z^*_{it-k} + \varepsilon_{zit} \,, \qquad z^*_{it} = 1 \text{ if } z_{it} > 0, 0 \text{ otherwise}$$

$$\text{where } \begin{pmatrix} \varepsilon_{yit} \\ \varepsilon_{zit} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$$

Based on the literature discussed above we included macroeconomic and banking sector variables as potential exogenous indicators of both banking crises and recessions. We include banking sector capital and liquidity ratios, house and equity price growth deflated by growth in the private consumption deflator, a measure of credit relative to GDP, the current account to GDP ratio and the first principal component indicator taken from an array of predictors used in the OECD's composite leading indicator.

Banks' capital adequacy (the ratio of capital and reserves for all banks over the end of year total of the balance sheet) was taken from the OECD Bank Income and Balance Sheet database. Liquidity data were sourced from the IMF's International Financial Statistics Database; liquid assets were defined as a sum of banks' claims on general government and the central bank, while total assets comprise foreign assets, claims on general government, central bank and private sector. Data on house prices, equity prices and credit were taken from Goodhart and Hofmann (2008).[2]

Following Borio and Drehmann (2009) we use deviations in the credit-to-GDP ratio from a trend measured using a Hodrick-Prescott filter. The series was calculated for each year recursively using the data up to that year in order to provide a measure of the deviation of the credit-to-GDP ratio from the underlying trend.

The principal component index was calculated using the first principal component of the leading indicators used in the OECD composite leading indicator index. The indicator was restricted to the selection of variables included in the OECD's indices in 2005 (to avoid introducing bias into the out-of-sample test conducted on the post-2005 data), although the factor loadings were estimated on data up to 2008. As not all the component series were available back to 1980 (the start of the estimation period) the first principal component was estimated using a modified methodology that could incorporate an unbalanced dataset. An EM algorithm was used to achieve this. The algorithm used the assumed factor structure to estimate factors and missing observations iteratively until both factor and missing observation estimates converged over the iterations.[3]

4 Unconditional analysis of the relationship between banking crises and recessions

Basic non-parametric analysis suggests that the broad conclusion taken from Figure 1 – that banking crises and recessions are correlated – is true. Table 1 shows that the probability of a recession if a banking crisis occurred in the previous year is 25%, three times higher than the unconditional probability. The difference is even greater for systemic banking crises (Table 2). Fisher exact tests indicate that there is an association between banking crises and recessions at the 10% significance level both contemporaneously and with recessions following banking crises with a one year lag.

Table 1: Occurrences of recessions in year after banking crises (1980-2005)

		Banking crisis in previous year		
		No	Yes	Total
Recession	No	311	9	320
	Yes	27	3	30
	Total	338	12	350
Probability of recession		8.0%	25.0%	8.6%

Table 2: Occurrences of recessions in the year after systemic banking crises (1980-2005)

		Systemic banking crisis in previous year		
		No	Yes	Total
Recession	No	317	3	320
	Yes	28	2	30
	Total	345	5	350
Probability of banking crisis		8.1%	40.0%	8.6%

In addition to these non-parametric tests, we ran simple bivariate probit regressions for all banking crises and recessions estimated with only lagged incidences of banking crises and

(2) The authors would like to thank Charles Goodhart and Boris Hofmann for permission to use the data and to thank Adam Posen, Neil Meads and Tomas Hellebrandt who provided data to extend the dataset to 2009.
(3) The authors would like to thank George Kapetanios for his help in estimating the first principal component indices.

recessions (ie without any of the exogenous variable used subsequently). The equation was estimated only up to 2005 to exclude the effects of the recent crises and recessions. If there is a causal link between recessions and banking crises the coefficients on lagged values of the dependent variables and/or the covariance term (ρ) should be non-zero.

Equation 1: Bivariate probit model of recessions and all banking crises

	Coefficient	Standard error	z	P>|z|
All banking crises				
Banking crisis$_{t-1}$	-4.52	2.17×10^3	0.00	1.00
Recession$_{t-1}$	0.40	0.37	1.09	0.28
Constant	-1.85	0.14	-13.29	0.00
Recessions				
Banking crisis$_{t-1}$	0.59	0.43	1.39	0.16
Recession$_{t-1}$	0.75	0.26	2.83	0.01
Constant	-1.51	0.11	-13.79	0.00
ρ	0.35	0.18	0	0.065
Number of observations:		350	Log likelihood:	-147.04
Sample period:		1981-2005		

The result (Equation 1) suggests that the covariance in the error terms of the two equations is significant at the 10% level, and that past occurrences of recessions may help predict recessions in the current year. The standard errors of the coefficient suggest that that recessions can help predict recessions, but systemic banking crises do not help predict further crises or recessions.

Similar results hold if attention is restricted to systemic banking crises (Equation 2). The standard errors of the coefficient suggest that that recessions can help predict recessions, but systemic banking crises do not help predict further crises or recessions. Nevertheless, there is once again a marked element of simultaneity as shown by the value of ρ.

Equation 2: Bivariate probit model of recessions and systemic banking crises

	Coefficient	Standard error	z	P>|z|
Systemic banking crises				
Banking crisis$_{t-1}$	-4.21	1.83×10^5	0.00	1.00
Recession$_{t-1}$	-4.79	1.95×10^4	0.00	1.00
Constant	-2.15	0.18	-12.15	0.00
Recessions				
Banking crisis$_{t-1}$	0.92	0.62	1.47	0.14
Recession$_{t-1}$	0.72	0.27	2.68	0.01
Constant	-1.49	0.11	-13.86	0.00
ρ	0.50	0.21		0.03
Number of observations:		350	Log likelihood:	-120.59
Sample period:		1981-2005		

The highly negative (if statistically insignificant) coefficients on some of the lagged dependent variables in the banking crises and systemic banking crises equations imply that there is negligible chance for there to be banking crises in two successive years. This reflects the fact that (as Figure 2 shows) there were no banking crises in a year immediately following a crisis in our sample of countries pre 2006. But the banking crises in the United Kingdom and United States of America in 2008 and 2009 highlight that this conclusion is faulty. This, of course, reflects the general problem of drawing inferences from relatively small samples of data.

5 Indicator models of banking crises and recessions

While these results suggest there is a causal link from banking crises to recessions, this could be a consequence of identifiable underlying causes rather than unobservable variables. As discussed in the methodology section, alongside lagged occurrences of banking crises and recessions we also tested for the effect of real house and equity price inflation, current account deficits as a proportion of GDP, the first principal component of OECD leading indicators of GDP growth, banking sector liquidity and banking sector capital ratios. Using annual data between 1980 and 2005, we estimated a bivariate probit model with the lags of all of these variables. There was evidence that lags of up to three years were statistically significant. But the economic rationale for such lags was unclear and could itself be evidence of overfitting. We therefore restricted our model to use only one-year lags of the available variables. Variables were then eliminated one at a time; removing the least significant macroeconomic indicators first and then the least significant banking sector indicators, until only those variables significant at the 10% level remained.[(4)]

The resulting model (Equation 3) indicates that a low bank capital ratio and a high credit-to-GDP ratio relative to trend both increase the probability of a banking crisis, as does a high current account deficit. However, house and equity prices and cyclical indicators do not appear to be significant predictors of banking crises.

Equation 3: Bivariate probit model of banking crises and recessions

	Coefficient	Standard error	z	P>|z|
Banking crises				
Capital_{t-1}	-0.10	0.06	-1.64	0.10
Credit to GDP ratio gap_{t-1}	0.13	0.04	3.20	0.00
Current account as % GDP_{t-1}	-0.13	0.06	-2.06	0.04
Constant	-1.49	0.33	-4.57	0.00
Recession				
First principal component_{t-1}	-0.23	0.10	-2.44	0.02
Real house price inflation_{t-1}	-0.06	0.02	-2.53	0.01
Real equity price inflation_{t-1}	-0.03	0.01	-3.37	0.00
Current account as % GDP_{t-1}	-0.13	0.05	-2.70	0.01
Constant	-1.82	0.19	-9.43	0.00
ρ	-0.020	0.313		0.95
Number of observations:		347	Log likelihood:	-100.76
Sample period:		1981-2005		

(4) The first run of the equation with all variables is shown in Appendix 1.

The leading indicators of GDP growth can help predict a recession now even though the indicators are taken from the previous year. Low or negative real house and equity price inflation both foreshadow recessions, as does a high current account deficit.

In both equations the constant term was statistically significant at each stage of the process of estimation. This contrasts with the findings of Barrell et al (2010b); they imposed the restriction that the constant term in their banking crises equation was zero so that the mean risk of a crisis was explained by the mean values of the explanatory variables.

The results also show that the covariance between error terms in the estimates for recessions and banking crises disappears when these other, previously omitted, variables are included. This indicates that the observed correlation between banking crises and recessions shown by Equation 1 is the result of identifiable underlying causes.

To estimate the bivariate model for systemic banking crises and recessions, we took the model for all banking crises and re-estimated for the sub-sample of systemic crises. The initial result for the sub-group of crises suggested that the capital ratio and current account were not statistically significant predictors of systemic banking crises: only the credit-to-GDP is statistically significant alongside the constant term (Equation 4). In addition the constant term is more negative than for all banking crises, possibly reflecting the lower probability of systemic banking crises.

Equation 4: Bivariate probit model of systemic banking crises and recessions

	Coefficient	Standard error	z	P>|z|
Banking crises				
Credit to GDP ratio gap$_{t-1}$	0.10	0.05	2.07	0.04
Constant	-2.33	0.23	-10.32	0.00
Recession				
First principal component$_{t-1}$	-0.25	0.10	-2.61	0.01
Real house price inflation$_{t-1}$	-0.06	0.02	-2.57	0.01
Real equity price inflation$_{t-1}$	-0.03	0.01	-3.32	0.00
Current account as % GDP$_{t-1}$	-0.13	0.05	-2.50	0.01
Constant	-1.80	0.18	-9.77	0.00
ρ	0.69	0.48		0.17
Number of observations:		347	Log likelihood:	-80.32
Sample period:		1981-2005		

Equation 3 includes a measure of the leverage of the banking sector, but not liquidity. Equation 4 includes neither measure. In both equations, the deviation of the credit-to-GDP ratio from its long-run trend is significant. The marginal significance of the capital term in the equation for all banking crises is consistent with recent work by the Basel Committee on Banking Supervision (2010), which suggests that that leverage ratios of individual banks are poor indicators of impending severe bank stress or failure.

5.1 *Diagnostic checks*

In order to conduct diagnostic tests on our preferred regressions we calculated the generalised residuals of the equations given by Gourieroux et al (1987). The residuals for each equation were calculated separately (given the covariance between the error terms in the two equations is not statistically significantly different from zero).

We use the test for cross-section dependence developed by Hsiao, Pesaran and Pick (2007) for non-linear panel models. The CD-test statistic uses the pair-wise correlation coefficients of the residuals in the regression equations over the period 1981-2005 for the i^{th} and j^{th} countries, ρ_{ij}, and is asymptotically normally distributed.

$$\mathrm{CD} = \sqrt{\frac{2\mathrm{T}}{\mathrm{N}(\mathrm{N}-1)}\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\tilde{\rho}_{ij}\right)}$$

The CD-test statistic for the banking crises equation is 0.87 and for systemic banking crises is 1.5 – neither are statistically significantly different from zero. The test statistic for the recession equation is 12.6 (12.8 when estimated alongside the systemic banking crisis equation) – significantly different from zero, suggesting that there is evidence of cross-sectional dependence and implying an omitted variable. Attempts to remove this using country fixed effects and time dummies were unsuccessful. Adding the unweighted yearly averages of all the exogenous variables to the regression also failed to correct for the cross-sectional correlation. However, adding the proportion of countries in recession as an explanatory variable to the recession equation, as proposed by Mitchell, Smith and Weale (2011) following on from Pesaran (2006), did eliminate the cross-sectional dependence – reducing the recession equation test statistic to just 0.73.[(5)] This model is presented in Equation 5.

[(5)] Predictions of recessions reported later in the paper are based on this equation, but the specification of the recession equation estimated alongside systemic banking crises is reported in Appendix 2 for information.

Equation 5: Bivariate probit model of banking crises and recessions including proportion of sample countries in recession

	Coefficient	Standard error	Z	P>|z|
Banking crises				
Capital$_{t-1}$	-0.10	0.06	-1.64	0.10
Credit to GDP ratio gap$_{t-1}$	0.13	0.04	3.15	0.00
Current account as % GDP$_{t-1}$	-0.13	0.06	-2.06	0.04
Constant	-1.49	0.33	-4.56	0.00
Recession				
First principal component$_{t-1}$	-0.17	0.11	-1.46	0.15
Real house price inflation$_{t-1}$	-0.05	0.03	-1.96	0.05
Real equity price inflation$_{t-1}$	-0.04	0.01	-2.86	0.00
Current account as % GDP$_{t-1}$	-0.11	0.06	-1.83	0.07
Proportion of sample in recession$_t$	3.81	0.75	5.08	0.00
Constant	-2.53	0.31	-8.13	0.00
ρ	-0.024	0.414		0.95
Number of observations:		347	Log likelihood:	-85.70
Sample period:		1981-2005		

Lagged values of the dependent variables were included at the first stage of the regression estimation process, together with country fixed effects, but both were dropped from the regression as part to the process of sequentially dropping variables that were not statistically significant. As the final result of any process of elimination may depend on the path used, we ran some robustness checks – adding country fixed effects and lagged dependent variables back into the model. Table 3 indicates confirms that these coefficients are not significant collectively.

Table 3: Likelihood ratio tests on restrictions on removal of country fixed effects and lagged dependent variables

		χ^2-statistic	Probability> χ^2
All banking crises	Country fixed effects	20.20	(0.78)
	Lagged dependent variables	5.32	(0.26)
Systemic banking crises	Country fixed effects	26.43	(0.44)
	Lagged dependent variables	2.02	(0.73)

As a further robustness check we looked for evidence of a structural break in the model after 2005, by extending the estimation period up to 2008 and using a dummy variable interacting with the other model variables to test for changes in coefficients. Table 4 indicates there is strong evidence of a structural break after 2005.

Table 4: Likelihood ratio tests for structural break after 2005

		χ^2-statistic	Probability > χ^2
All banking crises	Structural break in either recession or banking crisis equation	40.52	(0.00)
	Structural break in banking crisis equation only	22.71	(0.00)
	Structural break in recession equation only	18.51	(0.01)
Systemic banking crises	Structural break in either recession or banking crisis equation	22.57	(0.00)
	Structural break in banking crisis equation only	6.63	(0.04)
	Structural break in recession crisis equation only	16.07	(0.01)

An examination of the re-estimated models suggests that the structural break test for all banking crises is picking up a change in the coefficient of the deviation from the trend credit-to-GDP ratio. The break in the equation for systemic banking crises probably reflects a change in the constant – perhaps reflecting the higher incidence of systemic banking crises in the three years after 2005. The break in the equation of all banking crisis probably reflects a change in the coefficient on the credit to GDP gap term. This may reflect the fact that a country's own credit-to-GDP ratio was less important than cross-border financial linkages to other stricken banking systems in determining the probability of financial crises. The break in the recession equation reflects a change in the coefficient on the current account surplus. This may reflect that trade linkages meant that countries with significant trade surpluses also experienced recessions as demand in deficit countries fell sharply.

6 Predictive power

When assessing the usefulness of a predictive model of banking crises and recessions, an important factor is the choice of threshold used to determine a positive prediction for an event. Table 5 to Table 7 show the in-sample and out-of-sample performance of our preferred models of recessions and banking crises.

All forecasting models are prone to error and ours are no exception: depending on the threshold used our models can give a substantial number of false positive predictions. Using the proportion of crises or recessions in the period up to 2005 as the threshold, the model accurately predicts whether an economy will be in recession or not over 80% of the time (Table 5). But the model is prone to over-predict recessions, with an in-sample probability incorrectly predicting a recession of around 75%. Raising the threshold to twice the sample average or 50% prevented false predictions in-sample, but at the cost of failing to foresee a greater proportion of recessions (Table 6 and Table 7).

Table 5: Performance of model – sample average threshold

			Recession		Banking crisis		Systemic banking crisis	
	Predicted	Actual	In sample (1981-2005)	Out of sample (2006-2009)	In sample (1981-2005)	Out of sample (2006-2009)	In sample (1981-2005)	Out of sample (2006-2009)
i)	0	0	231	38	228	28	250	19
ii)	1	0	87	0	107	20	92	33
iii)	0	1	3	4	2	2	1	0
iv)	1	1	26	14	10	6	4	4
Event correctly predicted[(a)]			23%	100%	9%	23%	4%	11%
Absence of event correctly predicted[(b)]			99%	90%	99%	93%	100%	100%
Kuipers Score			0.62	0.78	0.51	0.33	0.53	0.37
Matthews correlation coefficient			0.37	0.84	0.20	0.23	0.14	0.20

(a) Probability of event correctly predicted given event is forecast = iv ÷ (ii + iv)
(b) Probability of absence of event correctly predicted given absence of event is forecast = i ÷ (i + iii)

Table 6: Performance of model – twice sample average threshold

			Recession		Banking crisis		Systemic banking crisis[(a)]	
	Predicted	Actual	In sample (1981-2005)	Out of sample (2006-2009)	In sample (1981-2005)	Out of sample (2006-2009)	In sample (1981-2005)	Out of sample (2006-2009)
i)	0	0	256	38	302	36	322	33
ii)	1	0	62	0	33	12	20	19
iii)	0	1	6	4	6	5	4	4
iv)	1	1	23	14	6	3	1	0
Event correctly predicted[(a)]			27%	100%	15%	20%	4%	0%
Absence of event correctly predicted[(b)]			98%	90%	98%	88%	99%	89%
Kuipers Score			0.60	0.78	0.40	0.13	0.14	-0.35
Matthews correlation coefficient			0.38	0.84	0.23	0.10	0.07	-0.19

(a) Probability of event correctly predicted given event is forecast = iv ÷ (ii + iv)
(b) Probability of absence of event correctly predicted given absence of event is forecast = i ÷ (i + iii)

Table 7: Performance of model – 50% threshold

			Recession		Banking crisis		Systemic banking crisis	
	Predicted	Actual	In sample (1981-2005)	Out of sample (2006-2009)	In sample (1981-2005)	Out of sample (2006-2009)	In sample (1981-2005)	Out of sample (2006-2009)
i)	0	0	310	38	334	45	342	52
ii)	1	0	8	0	1	3	0	0
iii)	0	1	19	15	12	8	5	4
iv)	1	1	10	3	0	0	0	0
Event correctly predicted[(a)]			56%	100%	0%	0%	0%	0%
Absence of event correctly predicted[(b)]			94%	72%	97%	85%	99%	93%
Kuipers Score			0.32	0.17	0.00	-0.06	0.00	0.00
Matthews correlation coefficient			0.40	0.35	-0.01	-0.10	N/A	N/A

(a) Probability of event correctly predicted given event is forecast = iv ÷ (ii + iv)
(b) Probability of absence of event correctly predicted given absence of event is forecast = i ÷ (i + iii)

The recession equation includes the contemporaneous proportion of the sample in recession, so predictions are made using an iterative process. An initial prediction of the proportion of the sample in recession is made using Equation 3. This first iteration is then fed into Equation 5 alongside the lagged exogenous data to provide a new estimate of the countries in recession. This process is repeated until the prediction of the countries in recession that is fed into the model is the same as the prediction coming from the model. This process can be thought of as analogous to the dynamic simulation of a time-series model.

If instead we take the proportion of countries with recessions as exogenous to the model – analogous to the static simulation of a time-series model – the resulting in-sample fit of the regression equation is much improved. For example, if we use the sample average as a threshold, the probability of correct prediction rises from 23% to 41% in sample but falls from 100% to 93.3% out of sample.

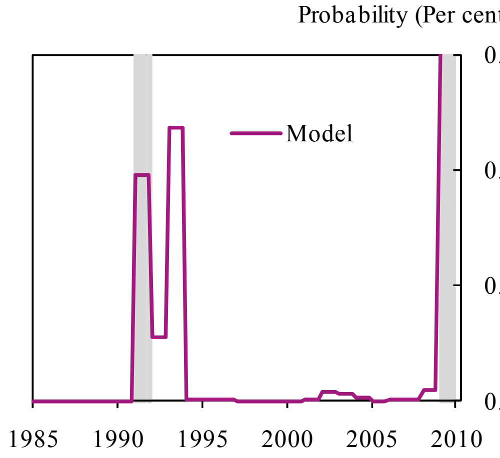
The tables also include the Kuipers Score and the Matthews Correlation Coefficient. As Cipollini and Kapetanios (2009) note, both these measures combine all the information contained in the upper parts of the table into a single value. The Kuipers score is defined as the difference between the proportion of events that were correctly forecasted (ie row iv divided by the sum of rows iv and iii) and the proportion of non-events that were incorrectly forecasted (ie row ii divided by rows i and ii). The Matthews correlation coefficient is the correlation coefficient between the forecast and actual outcomes.[(6)] For both measures a positive value indicates an improvement relative to a naive predictor such as the sample average occurrence. The results show that our models outperform a naive predictor, unless a very high threshold is used to indicate a banking crisis (Table 7).

[(6)] Mathematically this can be written $MC = \frac{(iv \times i)-(ii \times iii)}{\sqrt{(iv+ii)\times(iv+iii)\times(i+ii)\times(i+iii)}}$.

Out of sample, the model made over-predictions. At the lowest threshold used (the unconditional sample probability of recession), the model failed to predict only one recession over the sample period of 1981-2005. But the model failed to predict over a fifth of the recessions between 2006 and 2008. This probably reflects the degree to which the past few years were unusual, with problems in banking systems in some countries leading to a synchronised recession across all the countries in the sample.

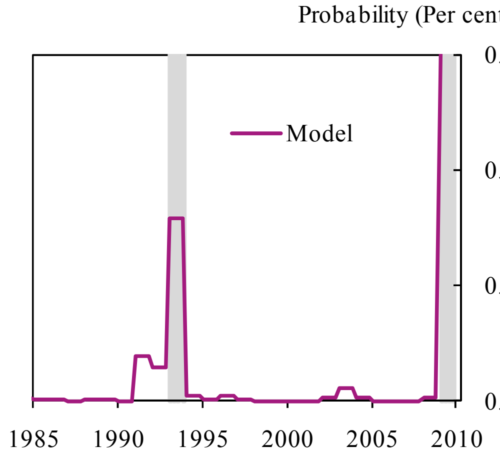
An alternative assessment of the model is offered by exploring its ability to predict the recent banking crisis and recession in the some of the larger countries in the sample: the United Kingdom, France, Germany and the United States. Chart 1 to Chart 4 shows the implied year-ahead probabilities of an annual fall of GDP in these countries. In all the countries the model correctly shows a sharp pick-up in the probability of recession in years when recessions occurred – particularly in 2009. But the charts also show that there are occasions when the model gives a strong indication of a recession even when none occurs. For example, in 1993 there were recessions in eight of the fourteen countries in the sample. This high proportion of recessions generates a high estimated probability of recession in other countries, such as the United Kingdom and United States.

Chart 1: Year-ahead predictions of recession in the United Kingdom



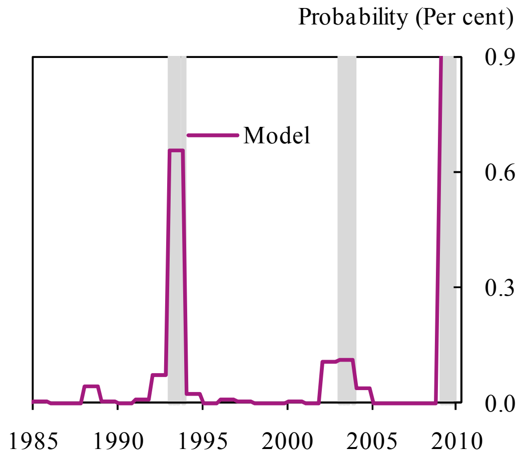
(a) Shaded area denotes actual recession.

Chart 2: Year-ahead predictions of recession in France



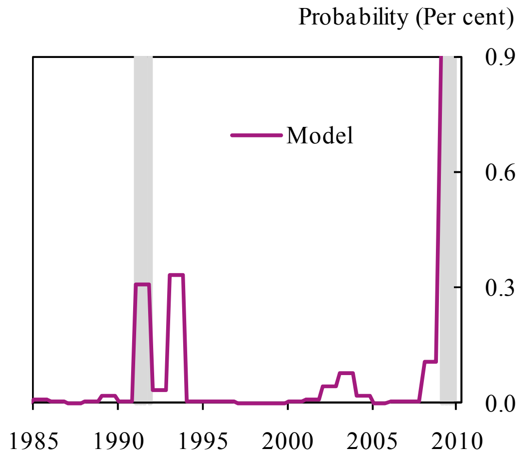
(a) Shaded area denotes actual recession.

Chart 3: Year-ahead predictions of recession in Germany



(a) Shaded area denotes actual recession.

Chart 4: Year-ahead predictions of recession in the United States



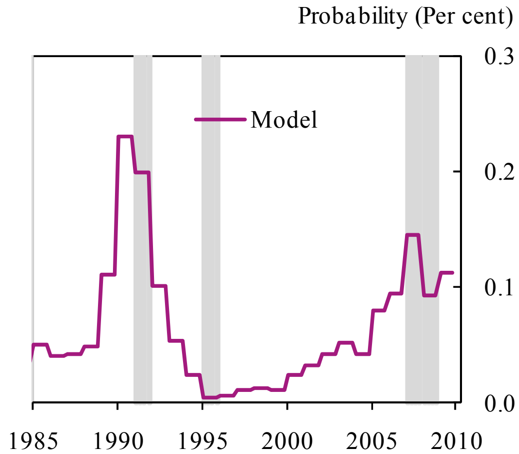
(a) Shaded area denotes actual recession.

The fit of the equations for banking crises are significantly worse than that for recessions. Relative to the recession equation, the equations for banking crises had a lower probability of being correct overall (at around 70% when measured in sample with the sample proportion as a threshold), as the model regularly predicted banking crises when none occurred. The probability of incorrectly forecasting a crisis when none occurred was over 90% when the unconditional sample probability was used as a threshold (Table 5). Raising the threshold to twice the sample proportion reduced the probability of over-predicting crises, but the performance was still poor with over 85% of predictions of crises turning out to be inaccurate (Table 6). If the threshold is raised to 50% then the model correctly predicts the outcome in over 90% of cases (ie the vast majority of years when there is no crises) but fails to predict any of the banking crises either in sample or out of sample (Table 7).

The out out-of-sample performance of the model is substantially better, although the model is still more likely than not to make a false positive prediction of a banking crisis. While the model failed to predict some banking crises, the model correctly spotted all four systemic banking crises out of sample when the sample average was used as a threshold. The improved ability of the model to predict banking crises correctly after 2005 again highlights how unusual recent years have been and is consistent with our stability tests which suggested that some coefficients had changed after 2005.

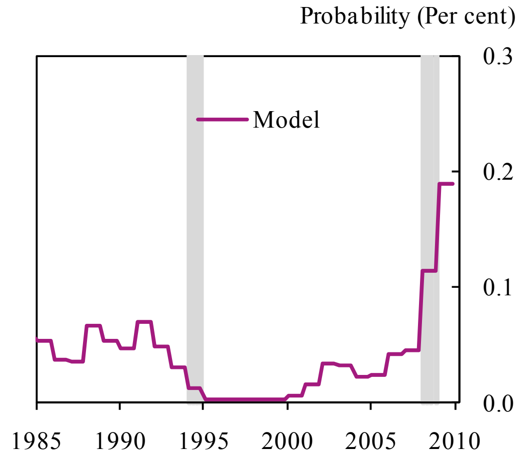
Given the difficulty in accurately predicting banking crises, the test of the model should not, perhaps, be whether it accurately predicts all outcomes, but whether it is a useful tool to assist policymakers interpret the data. Chart 5 to Chart 8 show the year-ahead predictions of banking crises in the United Kingdom, France, Germany and the United States from our model. Using the unconditional sample probability of a banking crisis across all countries would lead the model to indicate that a banking crisis was likely in the United Kingdom for the majority of the sample period. This headline result suggests the model would have been of little use to policymakers except as offering a general and rather useful indication that the United Kingdom was a high-risk country. But Chart 5 shows that the model did give a clear indication of the small banks crisis in the early 1990s and was pointing to a steadily rising risk of a banking crisis in the mid-to-late 2000s. The model also pointed to some slow pick-up in the risk of banking crises in France and the United States over the 2000s. The model does, however, fail to point to any increase in risk in the German banking system in recent years. Although the dataset does include the assets of Landesbanks, the failure to predict any of the recent problems in these banks may reflect the fact that problems in these banks generally arose from their exposure to cross-border investments rather than the domestic credit risks captured in our model. This is a reminder that models of this type are a complement to rather than a substitute for a careful examination of the risks associated with individual banks.

Chart 5: Year-ahead predictions of banking crises in the United Kingdom[(a)]



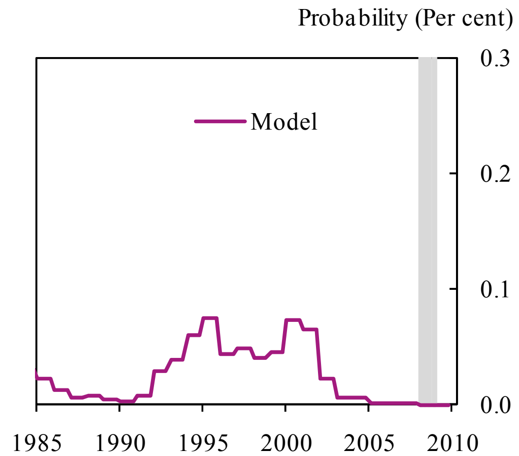
(a) Shaded area denotes actual banking crisis.

Chart 6: Year-ahead predictions of banking crises in France[(a)]



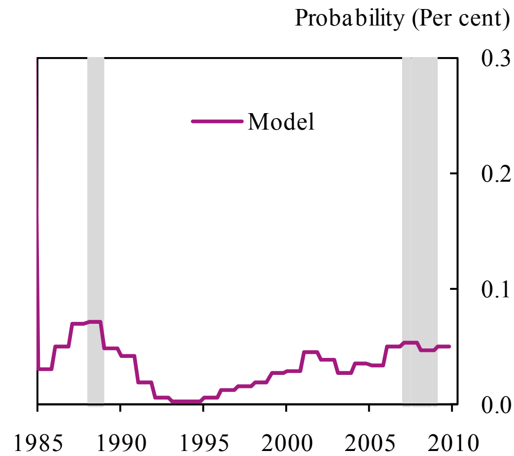
(a) Shaded area denotes actual banking crisis.

Chart 7: Year-ahead predictions of banking crises in Germany[(a)]



(a) Shaded area denotes actual banking crisis.

Chart 8: Year-ahead predictions of banking crises in the United States[(a)]



(a) Shaded area denotes actual banking crisis.

These results again highlight that the choice of threshold used to determine a positive prediction for an event is important. If the costs of failing to predict a recession or banking crisis are high, then the threshold for should be set low in order to avoid missing a crisis. However, if the threshold is set too low the model may regularly predict a crisis or recession. If the costs to over-predicting recessions or crises is high (perhaps because the public ignores warnings of impending crises from a policymaker who has over-predicted the probability of a crisis in the past), then the threshold for calling a crisis should be set very high. Thus the appropriate choice of these depends on the loss function of the policymaker.

7 Conclusions

We have found strong evidence for interdependency between recessions and banking crises using both non-parametric tests and unconditional bivariate probit models. Using data for a selection of OECD countries since 1981, we find that banking sector capital ratios, the deviation of the credit-to-GDP ratio from trend and the current account deficit are useful predictors of banking crises. We also find that the first principal component derived from OECD leading indicators of GDP growth helps predict recessions, as do movements in real house and equity price inflation, and the current account deficit. Once these exogenous variables are included we find that the observed interdependence between banking crises and recessions disappears – indicating that the observed interdependence is a result of common causes rather than direct causal links.

Our preferred models are, like all forecasting models, prone to mistakes. For commonly used thresholds to determine a positive prediction they tend to over-predict recessions and banking crises. Their out-of-sample performance (post-2005) is substantially better, but this indicates how unusual the past few years are, and is consistent with stability tests which suggest some model coefficients changed after 2005 (possibly reflecting the fact that increased global financial inter-linkages increased the risk of banking crises in all countries through contagion). Even though the precise model predictions are imperfect, the models do still provide policymakers with useful information. For example, our model points to the increased risk of a banking crisis in the United Kingdom in the mid-to-late 2000s, and it also corroborates the risk of a contraction in 2009 but probably identifies it no earlier than did forecasters using existing techniques.

Appendix 1: Initial bivariate probit model of banking crises and recessions with all lags of exogenous variables.

	Coefficient	Standard error	z	P>|z|
Banking crises				
Banking crisis$_{t-1}$	-7.02	8.32x10^{4}	0.00	1.00
Recession$_{t-1}$	0.60	0.62	0.96	0.34
First principal component$_{t-1}$	0.15	0.15	1.03	0.30
Liquidity$_{t-1}$	-2.09	5.10	-0.41	0.68
Capital$_{t-1}$	-0.05	0.11	-0.43	0.67
Real house price inflation$_{t-1}$	-0.04	0.03	-1.63	0.10
Real equity price inflation$_{t-1}$	0.00	0.01	-0.09	0.93
Credit to GDP ratio gap$_{t-1}$	0.21	0.07	3.19	0.00
Current account as % GDP$_{t-1}$	-0.18	0.10	-1.79	0.07
Country dummies				
Canada	2.99	93.01	0.03	0.97
Denmark	3.03	93.02	0.03	0.97
Finland	3.76	93.02	0.04	0.97
France	4.28	93.01	0.05	0.96
Germany	-1.29	1.20 x10^{4}	0.00	1.00
Italy	4.48	93.01	0.05	0.96
Japan	4.82	93.01	0.05	0.96
Netherlands	-2.74	5.16 x10^{6}	0.00	1.00
Norway	4.19	93.01	0.05	0.96
Spain	-2.31	5.38 x10^{4}	0.00	1.00
Sweden	3.82	93.01	0.04	0.97
UK	4.53	93.01	0.05	0.96
US	3.83	93.01	0.04	0.97
Constant	-5.56	93.02	-0.06	0.95
Recession				
Banking crisis$_{t-1}$	-0.10	0.78	-0.13	0.90
Recession$_{t-1}$	-1.02	0.46	-2.24	0.03
First principal component$_{t-1}$	-0.43	0.14	-2.99	0.00
Liquidity$_{t-1}$	-5.36	3.72	-1.44	0.15
Capital$_{t-1}$	0.10	0.14	0.75	0.46
Real house price inflation$_{t-1}$	-0.08	0.03	-2.64	0.01
Real equity price inflation$_{t-1}$	-0.03	0.01	-2.27	0.02
Credit to GDP ratio gap$_{t-1}$	0.07	0.05	1.22	0.22
Current account as % GDP$_{t-1}$	-0.19	0.06	-2.96	0.00
Country dummies				
Canada	-2.63	1.15	-2.28	0.02
Denmark	-1.29	1.04	-1.23	0.22
Finland	-1.86	1.14	-1.64	0.10
France	-1.31	0.88	-1.50	0.13
Germany	-0.23	0.68	-0.34	0.73
Italy	-0.89	0.88	-1.01	0.31
Japan	0.27	0.73	0.38	0.71
Netherlands	-0.84	0.91	-0.93	0.36
Norway	-0.76	0.80	-0.94	0.35
Spain	-1.52	1.06	-1.43	0.15
Sweden	-0.34	0.82	-0.42	0.68
UK	-1.32	0.96	-1.38	0.17
US	-1.11	0.94	-1.18	0.24
Constant	-0.78	1.06	-0.74	0.46
ρ	-1.00	0.00		0.31
Number of observations:		347	Log likelihood:	-83.795369

(7)

(7) Note that problems with near-collinearity of some variables (notably lagged dependent variables and country dummies) meant that some variables had to be excluded from early stages of the estimation process. These were added into the model after an initial process of elimination for the exogenous variables and the process of

Appendix 2: Bivariate probit model of systemic banking crises and recessions including proportion of sample countries in recession

	Coefficient	Standard error	z	P>|z|
Banking crises				
Credit to GDP ratio gap$_{t-1}$	0.11	0.05	2.30	0.02
Constant	-2.35	0.23	-10.43	0.00
Recession				
First principal component$_{t-1}$	-0.18	0.11	-1.56	0.12
Real house price inflation$_{t-1}$	-0.06	0.03	-2.00	0.05
Real equity price inflation$_{t-1}$	-0.03	0.01	-2.77	0.01
Current account as % GDP$_{t-1}$	-0.10	0.06	-1.76	0.08
Proportion of sample in recession$_t$	3.72	0.76	4.91	0.00
Constant	-2.52	0.31	-8.17	0.00
ρ	-0.353	0.491		0.83
Number of observations:		347	Log likelihood:	-65.97

elimination was restarted (adding in previously eliminated exogenous variables one at time to check they were still insignificant).

References

Azariadis, C (1981), 'The Economics of Self-Fulfilling Prophecies', *Journal of Economic Theory*, Vol. 25, pages 380-96.

Barrell, R, Davis, E P, Fic, T, Holland, D, Kirby, S and Liadze, I (2009), 'Optimal regulation of bank capital and liquidity: how to calibrate new international standards', *FSA Occasional Paper*, No. 38.

Barrell R, Davis, E P, Karim, D and Liadze, I (2010a), 'Was the subprime crisis unique? An analysis of the factors that help predict banking crises in OECD countries?', *NIESR Working Paper*, No 364.

Barrell R, Davis E P, Karim D and Liadze I (2010b), 'Bank regulation, property prices and early warning systems for banking crises in OECD countries', *NIESR Working Paper*, No 331.

Basel Committee on Banking Supervision (2010), *Calibrating regulatory minimum capital requirements and capital buffers: a top-down approach*, Bank for International Settlements, http://www.bis.org/publ/bcbs180.pdf

Bell, J and Pain, D (2000), 'Leading indicator models of banking crises – a critical review', *Bank of England Financial Stability Review*, December, pages 113-29.

Birchenhall, C R, Jessen, H, Osborn, D R and Simpson, P (1999), 'Predicting US Business Cycle Regimes', *Journal of Business and Economic Statistics*, Vol. 17, Pages 313-23.

Boyd, J, Gomis, P, Kwak, S and Smith, B (2000), 'A user's guide to banking crises', April mimeo, University of Minnesota

Burns, A F and Mitchell, W C (1946), 'Measuring business cycles', *NBER Studies in Business Cycles*, New York.

Camba-Mendez, G, Kapetanios, G, Smith R J, and Weale, M (2002), 'The Forecasting Performance of the OECD Composite Leading Indicators for France, Germany, Italy and the UK', in Clements M P and Hendry D F (eds.), *A Companion to Economic Forecasting*, Blackwell Publishers.

Cipollini, A and Kapetanios, G (2009), 'Forecasting financial crises and contagion in Asia using dynamic factor analysis', *Journal of Empirical Finance*, Vol. 16, pages 188-200.

Cass, D and Shell, K (1983), 'Do Sunspots Matter?' *Journal of Political Economy*, Vol. 91, pages 193-227.

Diamond, D and Dybvig, P (1983),'Bank runs, deposit insurance, and liquidity', *Journal of Political Economy*, Vol. 91, pages 401-19.

Estrella, A and Mishkin, F S (1997), 'The Predictive Power of the Term structure of Interest Rates in Europe and the United States: Implications for the European Central Bank', *European Economic Review*, Vol. 41, pages 1375-401.

Falcetti, E and Tudela, M (2008), 'What do twins share? A joint probit estimation of banking and currency crises', *Economica*, Vol. 75, pages 199-221.

Goodhart, C and Hofmann, B (2008), 'House prices, money, credit and the macroeconomy', *Oxford Review of Economic Policy*, Vol. 24, No. 1, pages 180-205.

Gourieroux, C, Monfort, A, Renault, E and Trognon, A (1987), 'Generalized residuals', *Journal of Econometrics*, Vol. 34, pages 5-32.

Hsiao, C, Pesaran, H and Pick, A (2007), 'Diagnostic tests of cross section independence for non-linear panel data models', *Cambridge Working Papers in Economics*, No. 0716.

Lui, S, Mitchell, J, and Weale, M (2011), 'Qualitative business surveys: signal or noise?', *Journal of the Royal Statistical Society: Series A*, Vol. 174, pages 327-48.

Marcellino, M (2006), 'Leading indicators', Chapter 16, in Elliott, G, Granger, C and Timmermann, A (eds), *Handbook of Economic Forecasting*, Vol. 1, North Holland.

Pesaran, M H (2006), 'Estimation and inference in large heterogeneous panels with a multifactor error structure', *Econometrica* Vol. 74, No. 4, pages 967–1012.

Rhodes, E C (1937), 'The construction of an index of business activity', *Journal of the Royal Statistical Society: Series C*, Vol. 100, No. 1, pages 18-66.

Schularick, M and Taylor, A (2009), 'Credit Booms Gone Bust: Monetary Policy, Leverage Cycles and Financial Crises, 1870–2008', *NBER Working Paper,* No. 15512.

Stock, J H and Watson, M W (1989), 'New indexes of coincident and leading economic indicators' in Blanchard, O and Fischer, S (eds), *NBER Macroeconomics Annual*, MIT Press, Cambridge MA.

Stock, J H and Watson, M W (2002), 'Forecasting using principal components from a large number of predictors', *Journal of the American Statistical Association*, Vol. 97, pages 1167-79.