BBVA | RESEARCH

Economic Watch

EAGLEs

Madrid, Feb 2014
Economic Analysis

Emerging Economies
Chief Economist
Alicia García-Herrero
alicia.garcia-herrero@bbva.com.hk

Cross-Country Emerging Markets Analysis
Chief Economist
Álvaro Ortiz Vidal-Abarca
alvaro.ortiz@bbva.com

Senior Economist
Alfonso Ugarte Ruiz
alfonso.ugarte@bbva.com

Excess Credit: Mind the Gap!... But which one?

- **Traditional measures of excess credit are too rough**

  The recent financial crisis has reminded us the need to develop early warning measures of excess credit and systemic risk in order to prevent stress episodes in the financial system. There are a number of traditional measures of excess credit, the best known being the growth of Credit-to-GDP. More recently, gaps of Credit-to-GDP derived from ad-hoc measures as stochastic trend (HP) or a linear trend have started to be used an early indicator of excess credit. Although easy to implement, these measures present some problems: First, as they are non-structural; in fact they do not focus on the reasons behind the excesses. Second, they face the "end-of-sample bias" rendering them less reliable when they are needed to take macro-prudential policy decisions Finally, some of these indicators tend to send false signals during tranquil times ("false positives").

- **We introduce a new measure of credit gap with promising results both in univariate and multivariate framework. Furthermore, we use our measure of excess credit to estimate the probability of a banking crisis with a very high prediction capability**

  We introduce a new structural credit gap indicator based on a panel data methodology. This methodology allowed us to estimate the credit gap for 68 countries on annual basis since the 1991 to now. Our credit gap presents several advantages with respect ad hoc measures: First, it takes into account the underlying conditions of each country. Second, it minimizes the so-called "*end-of-sample*" bias. Third, it displays a better forecasting accuracy than more traditional indicators, especially out-of-sample.

  The robustness of our results is reinforced by extending the analysis to a multivariate framework and finally estimating the probability of a banking crisis in a panel data model. To do this, we implement first a Bayesian Model Averaging (BMA) technique to confirm the superiority of our measure of excess credit when compared with several indicators of banking crisis. Secondly, we include our credit gap in a logit panel data model of probability of banking crisis for 68 countries. The results confirm the leading indicator properties of our structural gap in both developed and emerging economies. If estimated in 2007, our credit gap would have anticipated between 85% and 90% of the all the crises started afterwards with a low "*false positive*" – ratio.

- **Our model for the probability of banking Crisis allow us to estimate "dynamic" thresholds for the credit gap to be used an early warning tool**

  The estimated model allows us to estimate an early warning threshold of our structural credit gap which we fix at near 10% level. Values above these thresholds should be monitored cautiously. Besides, this threshold is a dynamic one, as it changes depending on the values of the rest of the variables so the monitoring should be done in a comprehensive way. We include several crisis examples (EU Periphery, The Baltics, the South East Asian Crisis and the Brazilian case) showing how the interaction of the credit gap with other key variables (such as interest rates, bank liquidity, current account deficit and US GDP growth) can lead to changes in the thresholds of our credit gap.

A new “structural” indicator of banking crises

The recent financial crisis has reminded us of the need of developing early warning measures of excess credit and systemic risk in order to prevent financial problems.

Empirical literature has developed several measures to estimate the probability of banking crisis focusing on **“excessive” credit growth indicators**, see for instance Lund-Jensen (2012), Borio and Lowe (2002), Borio and Lowe (2009).

Some of the IFIs (BIS, IMF and EU Commission)[1] **have included these excess credit measures as macro-prudential tools to monitor the health of the financial system**. They have relied on different indicators such as the **growth in the Credit-to-GDP** or **the Credit-to-GDP gap** derived from a stochastic Hodrick-Prescott or HP) or a linear trend. While the BIS has proposed a capital buffer framework conditional on the Credit-to-GDP gap derived from a HP filter, the IMF and the EU Commission use Credit-to-GDP-ratio changes to track the sustainability of Macroeconomic Imbalances (this procedure known as “six pack” renders the Credit to GDP gap threshold ratio to 15%). We envisage some problems within this approach:

- **Trends of Credit Gaps are derived from an “ad hoc” (i.e. non- structural) approach.** The use of filters or linear trends to estimate the gaps could be misleading especially in the more recent data due to the end-of-the-sample bias of the filters. This could be relevant if these measures are built to be used as macro-prudential tools.

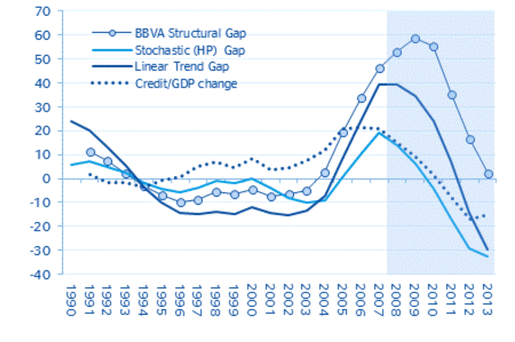
- **The use of changes in the credit-to-GDP ratio growth rates thresholds present also some problems**. First, it **does not account for structural differences as it assumes a common performance in all the countries independent of their idiosyncratic characteristics**. Second, the inclusion of variables that accumulate during expansions can introduce a bias in any “logit/probit” type of models because these variables usually depict high levels right before the crisis. Therefore, the influence of the credit could be the result of accumulative behavior during the boom rather than its predictive power[2].

In this Economic Watch we introduce our BBVA Structural Credit Gap, computed as a the deviation of the actual Credit to GDP ratio from the structural level estimated with our panel data model for the Private Credit-to-GDP ratio (see our previous EW Credit Deepening: The Healthy Path). Herein, we exploit the estimates of the long-term structural level of the private credit ratio, based on the long-term levels of several macroeconomic, regulatory and structural variables. The difference between the observed credit ratio and the estimated structural level is a measure of “excessive” leverage, and we will call it “Credit Gap” hereon. The estimation of the structural credit ratio includes around 20 explanatory variables that can be broadly classified into macroeconomic determinants, regulatory and institutional variables and structural determinants. As explained in the methodology section of the EW: “Credit Deepening, the healthy path”, the macroeconomic variables are decomposed into three time components (long, medium and short-term), but only the long-term component contribution is used to estimate the structural ratio together with the institutional, regulatory and macroeconomic variables. The difference between the Credit-to-GDP ratio and the estimated long term component will form our estimated credit gap.

The following charts (1 to 6) show the different patterns of the alternative credit gaps and Credit-to-GDP changes. Among them, we include our structural gap with two alternative gaps (linear and stochastic Hodrick Prescott filter) jointly with the simple annual change in the credit-to-GDP ratio[3].

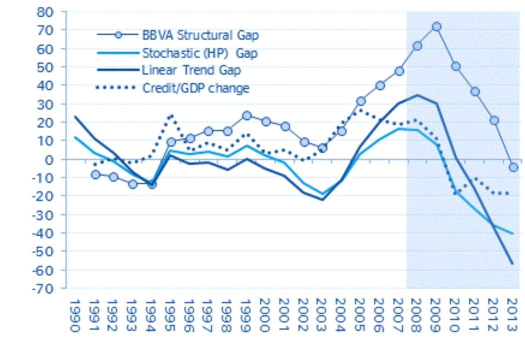
1: BIS Working Papers No. 317 (2010), BIS Countercyclical Capital Buffer Proposal (2010), Blancher et al (2013) WP/13/168 IMF “Systemic Risk Monitoring (“SysMo”) Toolkit, A User Guide”.
2: See Gadea-Rivas and Perez-Quiros (2012).
3: In all the cases, the Structural ratio, the HP-trend and the Linear-trend were estimated by estimating the panel data model, the HP filter and the linear trend using data only from 1990 up to 2006. Then, for
each year between 2007 and 2012 we re-estimate the panel model, the HP filter and the linear trend using data from 1990 up to the corresponding year and we add the resultant estimated value to the initial series obtained up to 2007 completing the series up to 2012. In order to observe how each of the methodologies would have performed before the onset of the international financial crisis in 2008, taking into account that both the HP-filter and the Linear-trend suffer from the “end-of-sample” problem and to compare them with our indicator under the same conditions

Chart 1
Spain: Credit-to-GDP and alternative credit gaps (Estimated, HP and linear trends)



Source: BBVA Research

Chart 2
Ireland: Credit-to-GDP and alternative credit gaps (Estimated, HP and linear trends)



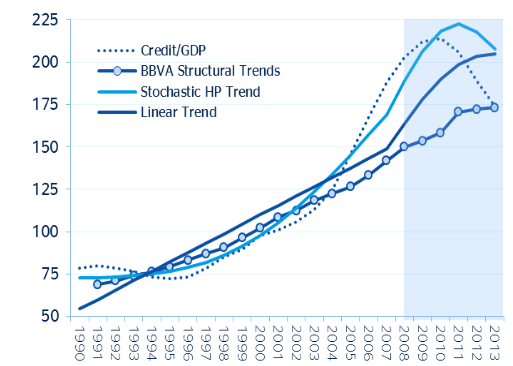
Source: BBVA Research

In Chart 1 and 2 **we can appreciate the alternative gaps and changes of two developed economies (Spain and Ireland)**. Some issues may be highlighted:

- **According to our Gap, the Credit-to-GDP ratio started to signal clear excesses during 2004-2005 in both Ireland and Spain.** However, while the Spanish ratio shows values below trend during the previous decade (1995-2005) the Irish gap pattern shows that some excesses where also apparent during this period, Thus, according to our model, part of credit deepening process which took place in Spain was justified by fundamentals..
- **In both cases, The HP-filter (using the standard lambda) generates a trend which is closer (than the linear trend or our structural ratio) to the actual ratio**. Thus, the corresponding gap is lower.
- **Our Structural Credit Gap minimizes the end of sample problem**. As can be observed from the trends of Spain and Ireland (Charts 3 & 4) our measure does not change radically after 2008 while the linear trend and specially the HP-filter trend change radically their slope after 2008 due to the so-called "end-of-sample" problem[4]. This is a key issue as Central Banks and Supervisory agencies need to assess the credit situation in real time
- **Finally, the structural credit gap maintains the leading indicator properties** (as it worsens before the other measures). Besides, it **signals the magnitude of the problem as the gap is wider than the alternative measures, especially those which use filtering approaches.**

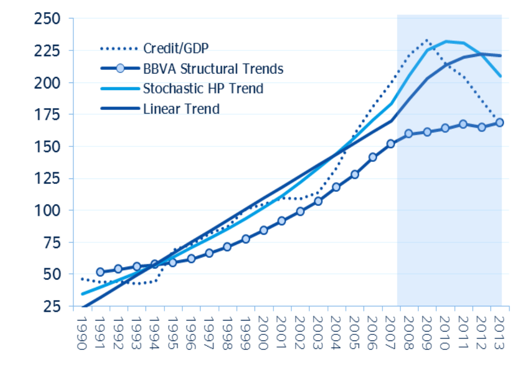
4: While simple trend extraction methods (linear trends and filters) are more convenient and easy to implement, the economic interpretation of their results may pose problems. This is mainly because it is not possible to adjust the filter to properties of the time series to be filtered. These ad-hoc procedures may also give rise to "spurious cycles" which reflect more the properties of the filter used rather than those of the time series. An additional problem concerns the instability of trend estimations at the end of the data sample. The trend values of the last sample periods can change significantly when the sample is extended with the arrival of new data.

Chart 3
Spain: Credit-to-GDP and alternative trend estimations



Source: BBVA Research

Chart 4
Ireland: Credit-to-GDP and alternative trend estimations

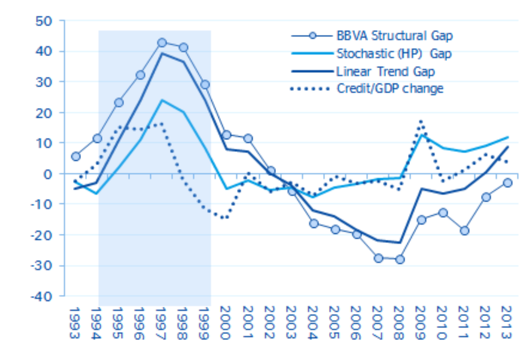


Source: BBVA Research

The extension of the analysis to the Emerging Markets maintains some similar characteristics with the Developed economies. The cases of Malaysia & Latvia (Chart 5 and 6) confirm the following about the estimated gap**:**

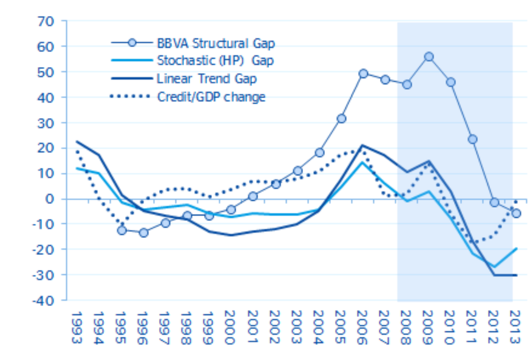
- **It keeps the leading indicator properties. In both countries the gap deterioration became apparent long before the crisis. In the Malaysian case the Credi-to-GDP ratio started to become risky in 1994 while the warning signal in Latvia was triggered long before the 2008 crisis.**
- It provides a clear signal of the magnitude of the problem, In the case of Malaysia our gap reached an excess of 40% while some of the alternatives ranged between 10%-20%. In the Latvian example, the excess credit surpassed the 30% level three years before the crisis while the rest of showed mild excesses.

Chart 5
Malaysia: Credit-to-GDP and alternative credit gaps (Estimated, HP and linear trends)



Source: BBVA Research

Chart 6
Latvia: Credit-to-GDP and alternative credit gaps (Estimated, HP and linear trends)



Source: BBVA Research

The Credit Gap as a leading indicator of banking crisis

So far we have shown that the Structural Credit Gap can be a good candidate as a leading indicator of financial and banking crisis in both developed and emerging markets. **In this section we extend the analysis to a multivariate framework to reinforce the robustness of our analysis and to show the suitability of our indicator as Early Warning instrument for financial crisis**.

In order to obtain the best possible model to assess the risk of a systemic banking crisis **we follow a three steps strategy** (These three steps are explained in more detail in the methodological box):

- **First, we compare the forecasting accuracy of banking crisis of the alternative credit indicators (gaps and changes) on individual basis.** For this we bi-variate panel data logit models of banking crisis[5] (as defined by Reinhart & Rogoff, 2010) including the alternative gaps as explanatory variable-
- Second, **we evaluate the robustness of the credit gap as a leading indicator in a multivariate framework through a Bayesian Modeling Average (BMA) technique.** In this sense we estimate the posterior probability of being an explanatory variable of banking crisis in a wider set of banking crisis indicators[6]
- Finally, and using the information provided with the BMA analysis, we chose among a large amount of possible model specifications to **specify a logit model of the Probability of Banking Crisis according to in-sample and out-of-sample forecasting performance.**

To check the forecasting accuracy of our **"structural credit gap" and the rest of the candidates we run several bivariate Logit panel data regressions of Banking Crisis with all these candidate variables as individual predictors[7]** . We use lagged values of the explanatory variables as our goal is to test the Early Warning capacity of the indicators.[8]

After each regression, **we compute different statistics to measure the statistical significance and the forecasting ability (in-sample and out-of-sample[9]) of the alternative leading indicators for each possibility**. In Table 1 we show the average of each statistic across the 12 methodological variations for each one of the possible predictors, in the in-sample and out-of-sample cases. We highlight in blue cells with the best statistic among the five variables (see methodological box for a more comprehensive explanation). The main results are the following:

5: The banking crisis binary dependent variable takes 1 for the crisis years and 0 otherwise. We have built up these variables by using the Reinhart Rogoff (2010) and Laeven and Valencia (2012)
6: See Babecky et al (2012). Banking, Debt and Currency Crises early warning indicators for developed countries. ECB wp 1485
7: The candidate variables are the Credit-to-GDP "gaps" derived from a HP-filter or a linear filter, the annual change in the Credit-to-GDP ratio, and a binary indicator which takes the value 1 when the Credit-to-GDP ratio is above a predetermined threshold (>5%) and 0 otherwise)
8: We test also for several possibilities including Unconditional vs. Conditional Logit models (Random vs. Fixed Effects), using alternative samples for the Banking Crisis variable (including all years of a crisis vs. sample including only the first-year of a crisis) and accounting for different lags of the leading indicators
9: The values of the Credit-Gap, the HP-Gap and the Linear-trend Gap (LT-Gap) from the years 2008 to 2011 are estimated separately as explained in the previous footnote in order to make the out-of-sample exercise as close as possible as if it had been if estimated in the year 2007.

Table 1
Summary of statistical and prediction performance statistics for five different individual predictors of a banking crisis. The statistics are the average of 12 different regressions

	Credit Gap	HP- Gap	Credit/GDP change	Credit/GDP change>5	LT-Gap
	Total in-sample 1990-2011				
z-stat	9.19**	8.74**	3.17**	5.89**	8.86**
ps-R2	0.2	0.17	0.04	0.07	0.17
NSR*	0.46	0.51	0.71	0.69	0.50
Loss*	0.60	0.63	0.77	0.76	0.63
	In-sample 1990-2007				
z-stat	5.92**	6.23**	0.03	2.43**	6.30**
ps-R2	0.12	0.13	0.02	0.02	0.13
NSR*	0.53	0.59	0.80	0.79	0.60
Loss*	0.66	0.68	0.85	0.86	0.68
	Out-sample 2008-2011, All crises				
NSR*	0.27	0.40	0.86	0.72	0.38
Loss*	0.52	0.73	0.90	0.77	0.73
	Out-sample 2008-2011, New crises				
NSR*	0.33	0.39	0.80	0.43	0.37
Loss*	0.56	0.74	0.86	0.61	0.73

1. z-stat: The z-statistic.**Significant at 1% level
2. Ps-R2: The pseudo-R-squared
3. NSR: The Noise-to-Signal Ratio
4. Loss*: The so-called loss function which is the percentage of missed crises (% type I error) plus the percentage of false signals (% type II error).
Source: BBVA Research

- **The structural credit gap is highly significant a (1% level), showing the higher R2 and the best forecasting accuracy according to the Noise-to-Signal ratio (NSR) and Loss Function.** This is maintained in both the whole sample (1990-2011) and the restricted sample (1990-2007) on in-sample basis
- **The out-of-sample forecasting analysis** (estimating until 2007 and keeping constant the parameters throughout the 2008-2011 period) **maintains the positive results in terms of forecast accuracy (lowest NSR and loss function).**

In the second step of our analysis we check the robustness of our individual results by testing the usefulness of our credit gap in a multivariate framework, when controlling for several other possible explanatory variables. Thus, **we identify the most useful early warning indicators of banking crisis by means of Bayesian model averaging (BMA).**[10] We include also the alternative gaps (HP, Linear) and Credit-to-GDP in the analysis in order to test that validity of our indicator in combination with the control variables (robustness in a multivariate framework).

The robustness of the potential indicators in explaining the banking crisis can be expressed by the probability that a given variable is included in the regression. **The posterior inclusion probability (PIP) captures the extent to which we can assess how robustly a potential explanatory variable is associated with the dependent variable.** Variables with a high PIP can

be considered robust determinants of the dependent variable, while variables with a low PIP are deemed not robustly related to the dependent variable.

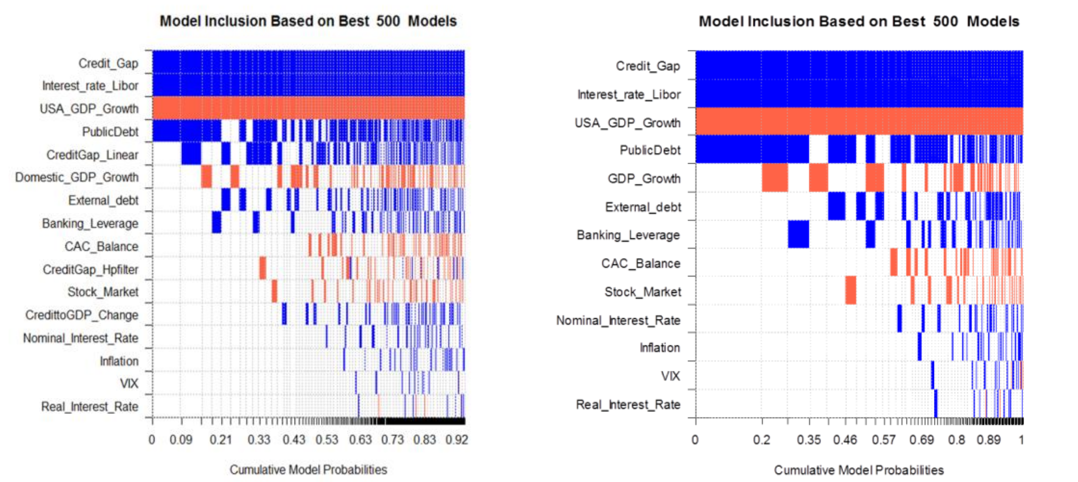
10: As signaled by Babecky et al (2012) this technique takes into account model uncertainty by considering various model combinations10 and thus has the advantage of minimizing the author's subjective judgment in determining the optimal set of early warning indicators.

In Chart 9 we represent the posterior inclusion probability (PIP) of every leading indicator in a model of Banking Crisis (based on the best 500 models out of the potential 65500 combinations).In panel (a) we include also the alternative gaps in order to recheck the validity of the estimated gap in a multivariate framework. In panel (b) we include only our estimated gap in a set of alternative variables. The different colors stand for the different sign of the explanatory variables (blue=positive sign, red=negative sign) The BMA analysis confirms several interesting features:

- In line with the existing literature, **excessive credit measures are good leading indicator of banking crisis**. Besides, **our credit gap is at the top** among the measures of excess credit, confirming its validity in a multivariate framework. Only the linear gap looks to include some extra information not included in our estimated gap
- **There are three variables which are key potential candidates to explain banking crisis** with **two years in advance** (presents PIP= 1): **our estimated credit gap, the libor interest rate and the US global Growth.** This reflects the importance of the **global factors as robust indicators of banking crisis. In this sense, the advantage of international economic coordination is straightforward, and the process of interest rates normalization by the Federal Reserves should be coordinated and carefully monitored.**
- **There are other domestic variables (excluding credit gaps and the variables including in their estimation) which can complement the analysis although their importance is lower than the credit gaps**. The stock market, the current account deficit and bank liquidity (credit to deposits ratio) could be complementary indicators.
- **Global Market indicators (VIX) can be suitable coincident indicators or triggers of banking crisis but could be less useful leading indicators** (lagged two years). I

Chart 9
Posterior Inclusion Probability in a model of Banking Crisis
(explanatory variables introduced with a 2 year lag, blue=positive sign, red=negative sign)

(a) Including Alternative Gaps **(b) Estimated Gap only**



Source: BBVA Research

Finally, we chose a final model among a large amount of possible model specifications, including different transformations of the explanatory variables. **The final model has been selected in terms of forecasting accuracy** (the one with the highest in-sample and especially out-of-sample performance) **among those that provide the earliest possible warning. The model includes 68 countries and the estimation sample includes annual data since 1991 to 2011. In this logit panel data model, the probability of Banking crisis two years ahead is expressed as a function of our previously estimated credit gap, the libor interest rate, the**

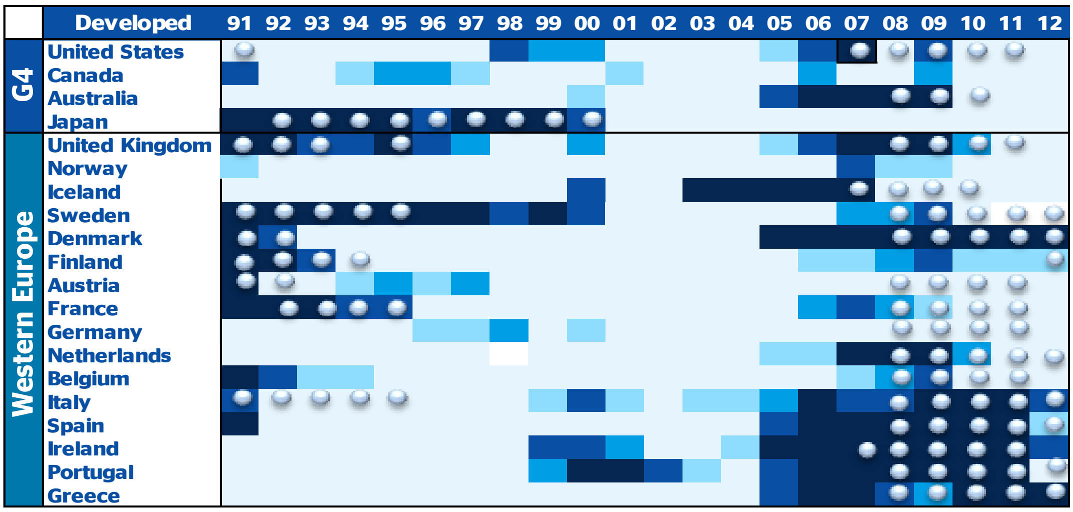
US GDP growth, liquidity (credit-to-deposit ratio) and current account deficit The final model's chosen specification is the following[11].:

$$Prob.Crisis_{i,t+2} = f(CreditGap_{i,t}\ , Libor_{i,t}\ , US\ GDP\ growth\ rate_{i,t}, Liquidity_{i,t}, (\frac{CA}{GDP})_{i,t}) \quad (1)$$

In Chart 10 and Chart 11 we can observe the two years ahead probability of a crisis estimated with our preferred model in both Developed and Emerging Markets, The color of the cell denotes the probability of a crisis, with a darker color indicating a higher probability. The light blue dots denote actual crises.

In Chart 10 we present the result for Developed Markets. As can be observed, the model would have provided a correct early warning signal in most of the last financial crisis (2007-2008) and in many cases with several years of anticipation (Australia, UK, Iceland, Denmark). The only exceptions in the developed economies were the crises in Austria and Germany whose banking problems originated in their foreign borrowing. It also identifies well the banking crises of the early 90's of Northern European countries and other events in UK, France and Italy.

Chart 10
Probability of Banking Crisis (2 years ahead) vs Actual Crises. Developed Economies

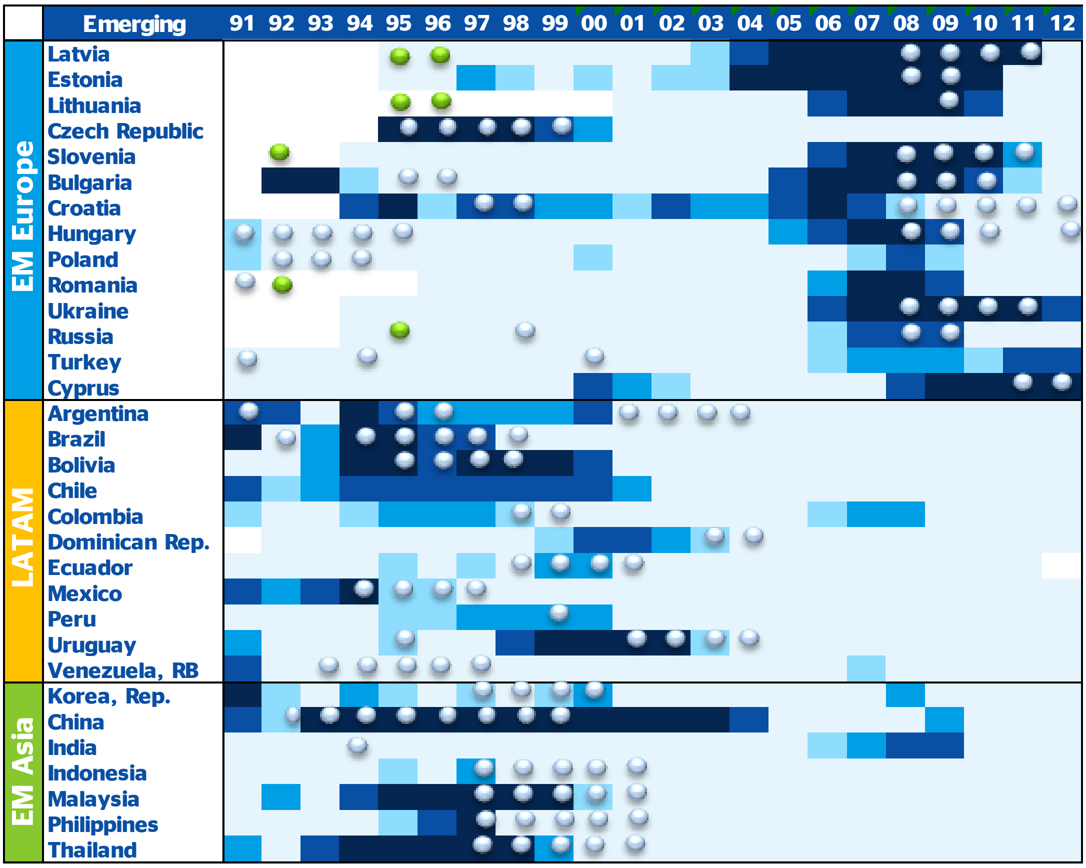


Source: BBVA Research. The light blue dots denote actual crises and the color of the cell denotes a higher predicted probability.

In Emerging Markets the model seems to perform quite well in predicting the crises of the mid-90s in Latin America and Asia and the crises in Eastern Europe after 2008. In the case of the Asian crisis of 1997 the model anticipated the financial problems in South Korea, Malaysia, Indonesia, Thailand and Philippines The model also anticipates well many isolated episodes such as different crises in some Latin American countries such as the Tequila crisis in México (well in advance) and the episodes in the early 00s (Argentina, Ecuador, Peru, Uruguay, Dominican Republic) and other crises in Eastern Europe in the early and mid-90s (Hungary, Poland, Czech Republic, Bulgaria). It is interesting to notice that the case of Cyprus that appears in Chart 11 is a strict out-of-sample case, since it is not included in the logit-regression. However, we can see that the model anticipates correctly the crisis started in 2011.

11: In the methodological box we show the prediction power of this specific model.

Chart 11
Probability of a Crisis (two years ahead) vs Actual Crises. Emerging Economies



Source: BBVA Research
The light blue dots denote actual crises and the color of the cell denotes a higher predicted probability. The green dots denote years of crises but that cannot be predicted by the model due to the lack of data

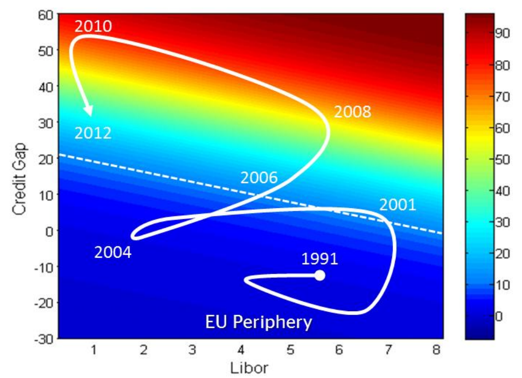
The model warns also about the existence of disequilibria in several cases not only before an actual banking crisis occurs, but also several years before and after the event, potentially signaling the state of de-leveraging process.

The estimated model allows us to analyze the early warning thresholds of our structural credit gap and how this threshold can change depending on the values of the rest of the variables (conditional marginal probabilities). In the following graphs (12 to 15) we observe the different regions of probability corresponding to the combinations of values of the Credit Gap (in the vertical axis) and each of the explanatory variables (horizontal axis). The color pattern oscillates between blue dark (meaning zero or low probability of crisis as can be observed in the right bar), the lighter blue colors around the risk threshold[12] (the dash line representing the level of the probability which triggers a warning signal) and the yellow and red area (where the probability of crisis is clearly above the risk threshold).

As observed in the graphics **the optimal threshold is dynamic in the sense that is changing with different values of the risk indicators**[13].The following examples show how the credit threshold can be affected by changes in the rest of variables:

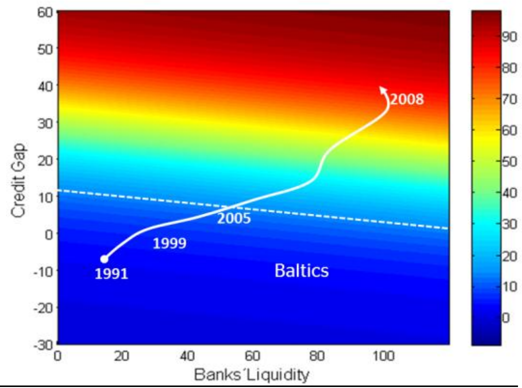
12: This optimal threshold probability is estimated by maximizing the difference between the percentage of correct signals and the percentage of false signals, (see methodological box)
13: A similar approach is used by Lund-Jensen (2012).Monitoring systemic risk based on Dynamic Thresholds. IMF Working Paper 12/159

Chart 12
Banking Crisis Probability surface (Credit Gap and the Libor interest rate): EU Periphery



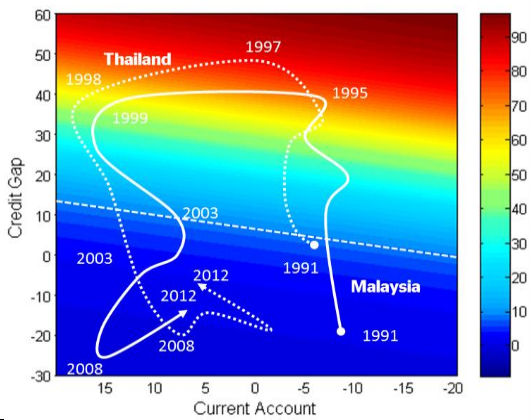
Source: BBVA Research

Chart 13
Banking Crisis Probability surface (Credit Gap and Credit-to-Deposits ratio): The Baltics



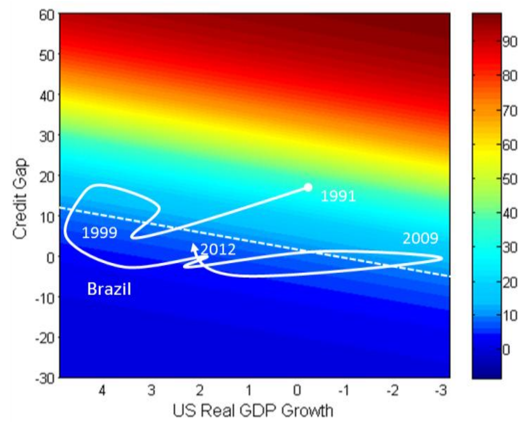
Source: BBVA Research

Chart 14
Banking Crisis Probability surface (Credit Gap and Current Account Balance): South East Asia



Source: BBVA Research

Chart 15
Banking Crisis Probability surface (Credit Gap and US GDP growth): Brazil



Source: BBVA Research

- **The first chart (12) shows the dynamic trajectory of both the credit gap and the global interest rate (libor) and the credit gap threshold for the EU Periphery countries.** As can be checked, when Global interest rates decreased **in 2003-04 the structural credit gap was nil (i.e credit was near equilibrium). Thereafter, the private credit acceleration triggered a rapid increase of the structural credit gap which surpassed the warning threshold during 2006, anticipating potential problems** just two years before the financial crisis erupted. **One important additional result is that the credit thresholds are not constant and they can change depending of the level of Libor interest rate** (from near 20% when monetary policies remain loose and interest rates stand at 1% to a more prudent 10% threshold when interest rates normalize at 4%-5%)
- **Bank liquidity (measured as Credit-to-Deposit ratio) can also affect the threshold ratio** although to a lesser extent than the global interest rate (lower slope). The second graph (chart 14) shows how **the Baltics move rapidly right-upwards with both the Structural credit gap and bank´s liquidity increasing faster since 2005. It can be observed that moving to a Credit-to-Deposit ratio of near 100% reduces also the size of structural credit gaps which triggers the warning signal.**
- **The relationship of the structural credit gap and the current account balance in the South East Asian countries (Malaysia and Thailand)** is represented in chart 14. **The initial combination of strong CAC deficits and the accelerating credit gaps was at the root of the Asian Crisis of 1997-1998** (although the model was anticipating the problems years

before). **Later**, after the crisis, **both the current account deficit and excess credit experienced a sharp reversal which was followed by a long lasting de-leveraging process.**

- Finally, **the Brazilian example (chart 15) shows us how the US GDP growth can affect the early warning threshold and probability of crisis even when the credit gap keeps constant.**

In sum, the model shows that credit gaps above 10% should be monitored. Besides, this should be done jointly with the rest of the key variables as, our analysis shows, thresholds are dynamic. Thus, the same credit gap could be indicating different levels of risk depending of the situation of the rest of variables.

References

- Blancher, N., Mitra, S., Morsy, H., Otani, A., Severo, T. and Valderrama, L. (2013) WP/13/168 IMF "Systemic Risk Monitoring ("SysMo") Toolkit, A User Guide.
- Borio, C. and Drehmann, M. (2009) "Assessing the risk of banking crises revisited", BIS Quarterly Review, March (2009)
- Borio, C. and Lowe, P. (2002) "Asset prices, financial and monetary stability: exploring the nexus", BIS Working Papers No. 114
- Davis, E.P. and D. Karim, (2008), "Comparing early warning systems for banking crises," Journal of Financial Stability, 4 (2), 89–120.
- Demirgüç-Kunt, A. and E. Detragiache, 1998a, "The determinants of banking crises in developing and developed countries," International Monetary Fund Staff Paper, 45 (1), 81–109.
- Demirgüç-Kunt, A., and E. Detragiache, 2000, "Monitoring banking sector fragility: a multivariate logit approach." World Bank Economic Review, 14 (2), 287–307.
- Drehmann, M., Borio, C., Gambacorta, G.J., Trucharte, C. (2010), "Countercyclical capital buffers: exploring options", BIS Working Paper No 317
- Edge, R. and Meisenzahl, R. (2011), "The unreliability of credit-to-GDP ratio gaps in real-time and the implications for countercyclical capital buffers", International Journal of Central Banking, Volume 7, Issue 4, December 2011, Pages: 261-298.
- Gadea-Rivas, M., and Perez-Quiros, G., (2012), "The failure to predict the great recession, the failure of academic economics? A view focusing on the role of credit", Working Paper, Bank of Spain No. 1240.
- Laeven, L and F. Valencia, 2008, "Systemic banking crises: a new database," IMF Working Paper, No. WP/08/224, International Monetary Fund.
- Laeven, L and F. Valencia, 2012, "Systemic Banking Crises Database: An Update," IMF Working Paper, No. WP/12/163, International Monetary Fund.
- Lund-Jensen, K. (2012), "Monitoring Systemic Risk Based on Dynamic Thresholds", IMF Working Paper WP/12/159
- Reinhart, C. M. and K. S. Rogoff, 2010, "From Financial Crash to Debt Crisis," NBER Working Paper, No. 15795

Methodological Box: Panel Logit Regressions

The empirical analysis is based on the assumption that the probability of a systemic banking crisis follows a logistic distribution that depends on k systemic risk factors, $X_{i,t\text{-}j}$, such that the probability of a systemic banking crisis, in country i, can be written as.

$$\text{Prob}\ \left(Y_t = 1 \mid X_{i,t-j}, \beta\right) = \frac{\exp(\alpha + X_{i,t-j}{}'\beta + \varepsilon)}{(1 + \exp(\alpha + X_{i,t-j}{}'\beta + \varepsilon))} \qquad (1)$$

The dependent variable is a binary variable equal to one if there is a crisis in the period t, and the systemic risk factors, $X_{i,t\text{-}j}$ are known j periods in advance, in order to be able to provide an early warning signal.

The purpose of the analysis is not to replicate or challenge previous results in the literature, but to show that our estimated credit gap is a more precise and robust predictor of banking crises than the options currently available. Thus, we initially follow closely the studies by Lund-Jensen (2012) and Borio et al (2002) and (2009), but we diverge from them in the sense that we explore and test a much wider set of methodological possibilities.

The empirical analysis is based on an unbalanced annual panel of 68 advanced and emerging economies over the time period 1990-2011. We restrict the analysis to this period because it is the period for which we have previously estimated the underlying panel data model for the Credit-to-GDP ratio.

For the dependent variable we follow the definition of a crisis from Lund-Jensen (2012) and Reinhart and Rogoff (2010), although in the cases in which we do not have information from those sources we follow Laeven and Valencia (2012).

Initially, we test the individual performance of 5 different indicators of an excessive credit level and we compare their performance in each methodological variation based on 5 different statistics. We finally compare the performance of each one of the 5 variables across all estimated regressions.

When we are able to show the superior performance of the credit gap as an individual predictor, we then examine its robustness when considering a large set of control variables and we look for the best possible model that we can have when combining the credit gap with other indicators of systemic risk.

Comparing its performance against most commonly used leading indicators of excessive credit

Research economists and policy makers have recently proposed and tested several alternative indicators of excessive credit growth, using different transformations of the Credit-to-GDP ratio.

1. The annual change in the Credit-to-GDP ratio
2. The Credit-to-GDP "gap" derived from a HP-filter.
3. The Credit-to-GDP "gap" from a linear trend.
4. The annualized change in the Credit-to-GDP ratio when such ratio is higher than a certain threshold (like 5 points) and zero when is lower than the threshold.

In order to compare how well these different indicators anticipate the onset of financial crises versus our new "credit gap" we run several regressions with all these variables as explanatory variables (using their lagged values), comparing their performance with the performance of our "credit gap". Our empirical exercise comprises several possible methodological possibilities. These possibilities include:

- **Unconditional vs. Conditional Logit models (Random vs. Fixed Effects):** In the type of models we are dealing with, this option seems to be quite important, since each one has both advantages and shortcomings. For instance, although a "Fixed-Effects" type of model should help us to avoid the "omitted variables" problem, it presents important shortcomings when dealing with a binary (crisis/no crisis) variable. A fixed-effect model is actually a "conditional" probability model, in the sense that the probability is conditional to have had a crisis during the sample period. Thus, it eliminates all the countries in the sample that have never had a crisis. Besides losing valuable degrees of freedom, we also lose important information that such countries could be giving us regarding why a country has never had a crisis.

- **Sample including all years of a crisis vs. sample including only the first-year of a crisis:** Some recent similar studies have restricted the analysis to a sample that only includes the first years of a crisis, eliminating from the sample all the subsequent years in which a crisis is still ongoing, despite of whether such crisis goes on for a year or for several years. The rationale for this restriction is that we can claim that it is more important to anticipate a crisis than predicting its duration, which is true. However, this approach has also a shortcoming. The start of a crisis is very difficult to date at an exact point in time and some crises are even controversial in the sense that some researchers classify them as crises and others don't. Other times, different sources differ even in the actual year in which a crisis started. Many times, a crisis can last several years, but the start of it could be rather mild, while the worst impact could emerge later on. Additionally, there are obvious advantages of knowing the duration of a crisis, for instance, knowing ex-ante whether we will have to deal with a short crisis or a long and deep one.

- **Different lags of the leading indicators:** As we are trying to anticipate and predict a future crisis, we obviously want to rely on the most readily available information that can give us an early signal. For instance it would be more useful to predict a crisis two years in advance rather than only one year in advance. Thus, we run the regressions and compare the indicators introducing them with one or two year lags. We also estimate the option of including the average of the previous two lags of all the possible indicators.

Thus, we run a logit regression for each variable under each one of all the possible combinations of these methodological variations. After doing so, we compute five statistics that measures the statistical significance and the predictive power of the corresponding leading indicator in each methodological variation. Those statistics are:

1. The z-statistic.
2. The pseudo-R-squared
3. The Noise-to-Signal Ratio
4. The Signal-to-Noise Ratio

5. The so-called loss function which is the percentage of missed crises (% type I error) plus the percentage of false signals (% type II error).

The statistics 3 to 5 are computed at the optimal "cut-off" point that is estimated for each one of the regressions. This is, for each logit regression we estimate the cut-off probability that maximizes the percentage of correct signals versus the percentage of false signals. We compute both the Noise-to-Signal ratio and the Signal-to-Noise ratio since it can be shown that these ratios could lead to different conclusions. Maximizing the difference between the percentage of correct signals (true positives) minus the percentage of incorrect signals (type II error) or "false positives" is exactly equivalent to minimizing the "loss function" which is the sum of the percentage of type I and type II errors. The use of the loss function can help us avoiding the contradictions between the NSR and the SNR.

Also it is important to highlight that we are stricter than other studies when evaluating the performance of the indicators, since we define a signal to be correct only if in that same year a crisis occurs, and to be false if a crisis does not occur, independently of what happens in the previous or the next year. Other studies consider a signal to be correct if a crisis occurs in a certain window of two or three years before and/or after the crisis.

In Sample Analysis

In Chart 17 we can see the results of the one of the 12 possible methodological variations considered. In this case we show the results of the least restrictive regression for each one of the possible leading indicator. The least restrictive case is a "random-effects" logit regression, including all the years of a crisis and with each variable introduced with a two year lag, i.e. explanatory variables are introduced in (t-2). In Chart 17 we can see that the Credit Gap is the variable with the highest z-statistic, the highest pseudo R-square, the highest Signal-to-Noise Ratio and the highest difference between the percentages of true positives vs. false positive, in this particular specification.

As explained before, we have run 12 regressions for each indicator, combining all the possible variations of lags, fixed vs. random effects and different definition of the dependent variable. This means that we actually have 11 alternative results to the ones shown in Chart 17. Later on, we summarize and compare the results of the 12 different regressions.

Chart 17
In-sample statistical results: Unconditional Logit, variables included with a two year lag (t-2), and the sample includes all years of a crisis. Period 1991-2011

	Credit Gap	HP- Gap	Credit/GDP change	Credit/GDP change>5	LT-Gap
z-stat	11.84	11.43	3.52	7.76	11.50
ps-R2	0.22	0.17	0.01	0.06	0.17
NSR*	0.50	0.38	0.71	0.73	0.41
SNR*	3.50	2.59	2.30	2.75	2.96
Loss*	0.58	0.55	0.76	0.76	0.55
True Positive %	59%	73%	42%	37%	68%
False Positive %	17%	28%	18%	14%	23%

Source: BBVA Research
NSR=(%Type II)/(1-%Type I); SNR=(%Type I)/(1-%Type II); %TP-FP=(1-%Type I) -(%Type II). Loss*=(%Type II)+(%Type I)

Out of Sample Analysis

The results shown in Chart 17 are "in-sample" results using the period from 1991 to 2011. In order to see how well the different indicators would have performed if used to predict the probability of a crisis before 2008, we have run an out-of-sample exercise in which we run the same regression but using only information up until 2007. Using the in-sample results of the period 1991-2007 we estimate two kind of out-sample predictions:

- The onset of new crises between 2008 and 2011, i.e. we only consider how well the model anticipates the first year of a crisis, without taking into account if it predicts well the second and further years.
- The first year and all the subsequent years of a crisis between 2008 and 2011

In Chart 18 we can see the same "in-sample" results as in Chart 9, but for the more restricted period of 1991-2007. In this example we can see that the in-sample performance of the gap calculated based on a linear trend (LT-Gap) is actually better than our credit gap's performance. Another important result to highlight is that the estimated sign for the change in the Credit-to-GDP ratio is negative, and thus it would have actually suggested that the higher the change in this ratio the lower the probability of a crisis. Additionally if we had considered only the effect of the change in the Credit-to-GDP ratio when it is higher than 5 points, we would have found that this variable had a very poor prediction performance. Hence, before 2008 we might have disregarded the change in the credit ratio as a useful predictor of a crisis, although ex-post it appears to be a good one

Chart 18
In-sample statistical results: Unconditional Logit, variables included with a two year lag (t-2), sample including all years of a crisis. Period 1991-2007

	Credit Gap	HP- Gap	Credit/GDP change	Credit/GDP change>5	LT-Gap
z-stat	8.07	8.06	-1.46	2.86	8.12
ps-R2	0.13	0.13	0.00	0.01	0.13
NSR*	0.57	0.46	0.80	0.91	0.45
SNR*	2.34	2.79	1.82	1.92	2.84
Loss*	0.67	0.59	0.84	0.92	0.58
True Positive %	57%	65%	36%	17%	65%
False Positive %	24%	23%	20%	9%	23%

Source: BBVA Research
NSR=(%Type II)/(1-%Type I); SNR=(%Type I)/(1-%Type II); %TP-FP=(1-%Type I) -(%Type II). Loss*=(%Type II)+(%Type I)

In Chart 19 we can see the results of the out-of-sample prediction of the onset of new crises between 2008 and 2011 and in Chart 20 we can see the results of the out-of-sample prediction that in the next year there will be a crisis, independently of whether the crisis has already started or not.

In Chart 19 we can see that the Credit Gap is by far the best predictor of new crises, since it displays the lowest NSR, the second highest SNR and the lowest "loss" value. In Chart 20 we can see that it is also the best predictor of all the years of a crisis, since it displays the lowest NSR, the second highest SNR and the lowest "loss" value.

Chart 19
Out-of-sample statistical results: Unconditional Logit, variables lagged two periods (t-2). Prediction of "new crises" started between 2008-2011

	Credit Gap	HP- Gap	Credit/GDP change	Credit/GDP change>5	LT-Gap
NSR*	0.27	0.36	1.19	0.58	0.38
SNR*	3.11	1.45	0.00	3.66	1.41
Loss*	0.46	0.73	1.16	0.64	0.75
True Positive %	80%	85%	0%	50%	85%
False Positive %	26%	58%	16%	14%	60%

Source: BBVA Research
NSR=(%Type II)/(1-%Type I); SNR=(%Type I)/(1-%Type II); %TP-FP=(1-%Type I) - (%Type II). Loss*=(%Type II)+(%Type I)

Chart 20
Out-of-sample statistical results: Unconditional Logit, variables lagged two periods (t-2). Prediction of all crisis years between 2008 and 2011

	Credit Gap	HP- Gap	Credit/GDP change	Credit/GDP change>5	LT-Gap
NSR*	0.27	0.28	1.15	0.52	0.29
SNR*	3.11	1.51	0.22	4.05	1.47
Loss*	0.46	0.70	1.12	0.58	0.72
True Positive %	80%	88%	4%	55%	88%
False Positive %	26%	58%	16%	14%	60%

Source: BBVA Research
NSR=(%Type II)/(1-%Type I); SNR=(%Type I)/(1-%Type II); %TP-FP=(1-%Type I) - (%Type II). Loss*=(%Type II)+(%Type I)

As explained in the "in-sample" analysis, we have repeated the exercise shown in Charts 17 to 20, a total of 12 times, one for each possible methodological variation discussed previously.

In order to have a complete picture of which indicator is the best one considering the large amount of possible methodologies and the discrepancies between different statistics, we have computed the total number of times that each indicator is the best one out of the 5 indicators, in all the in-sample and out-of-sample exercises, and regarding each possible statistic, i.e. z-stat, pseudo R-square, NSR, SNR, and the loss value (i.e. the sum of missed crises and false signals).

For instance in Chart 17, our Credit Gap is the best indicator according to 3 statistics, (z-statistic, pseudo R-squared, and Loss value, the "LT-Gap" is the best indicator according to one statistic (SNR) and the "HP-Gap" is the best indicator according to one statistic (NSR).

The summary of the number of times that each indicator is the best one according to each statistic is shown in Chart 21. There we can see our Credit Gap is the best indicator of a crisis a total of 50 times according to in-sample statistics. The Linear-trend Gap is the second best with a total of 38 times. The Credit Gap is clearly the best indicator in the period 1991-2011; meanwhile the LT-Gap is the best one in the period before the onset of the Great Recession, i.e. before 2008.

The most remarkable difference comes regarding out-of-sample results. In this case or Credit Gap is the best indicator a total of 40 times out of a total of 73 out-of-sample computed statistics. The second best indicator is the change in the Credit-to-GDP when it is larger than 5 pp. However this same variable display a very lousy statistical performance in-sample and it is even non-significant several times in the period 1991-2007.

Chart 21
Summary of the number of times each leading indicator is the best performer

	Credit Gap	HP- Gap	Credit/GDP change	Credit/GDP change>5	LT-Gap	Total
In sample 1990-2011	32	12	2	2	12	60
In sample 1990-2007	18	12	0	4	26	60
Out of sample	40	3	6	21	3	73
All indicators	90	27	8	27	41	193

Source: BBVA Research

Although we do not show explicitly all the possible results, we are able to claim that the credit gap was in general the best performer in almost any of the cases considered, for instance, when using fixed effects or when restricting the sample to only the first years of the crises. Hence, we can safely conclude that our credit gap is by far the best leading indicator among the options considered herein and that this result is robust to several methodological possibilities.

Testing its robustness when using it together with other leading indicators

In the previous section we have established that our credit gap is the best individual performance among those derived from the Credit-to-GDP ratio when we run a binary regression model with only one explanatory variable. However we need to explore if it is still the best predictor when controlling for several other possible explanatory variables. Since the different combinations of explanatory variables are huge, we rely on a methodology that helps us account for the uncertainty in the model selection process.

Thus we run a Bayesian Model Averaging (BMA) with the dependent variable being the dummy for banking crisis and as possible explanatory variables the following ones:

- Our Credit Gap
- The Credit-to-GDP "gap" derived from a HP-filter.
- The Credit-to-GDP "gap" from a linear trend.
- The annualized change in the Credit-to-GDP ratio when such ratio is higher than a certain threshold (like 5 points) and zero when is lower than the threshold.
- The Credit-to-Deposits ratio (Liquidity) in country i
- The current account balance as percentage of GDP in country i
- The short-term interest rate in country i
- The inflation rate
- The real GDP growth rate in country i
- The annual growth rate of the stock market
- The public debt to GDP ratio
- The Libor interest rate
- US GDP growth rate
- The S&P volatility index (VIX)

Bayesian Model Averaging is a technique designed to help account for the uncertainty inherent in the model selection process, something which traditional statistical analysis often neglects. By averaging over many different competing models, BMA incorporates model uncertainty into conclusions about parameters and prediction.

Choosing the best model for crises' prediction

Once we have established that our new estimated Credit Gap is by far the best predictor of crises among the most commonly used indicators of excessive credit, we also want to find out which is the best possible model that we can obtain using our Credit Gap and other indicators like interest rates, external vulnerabilities and banking liquidity.

We also take advantage of the results of the Bayesian Average Modeling performed in the previous section, which suggest that the best leading indicators of a banking systemic crisis are, besides our credit-gap, the Libor interest rate, the GDP growth rate of the USA and the current account balance. Since the possible combinations are plentiful, we have run a similar exercise than in section 2, estimating a large number of models, estimating and comparing for each one of them a set of statistics measuring its performance as a crisis predictor, considering both in-sample and out-of-sample performance. We combine different lags of the credit-gap and the following control variables:

- Libor interest rate,
- GDP growth rate in the US,
- Credit-to-Deposits ratio (banking liquidity)
- Current account balance as percentage of GDP

Additionally, we try different models in which we include only one of these control-variables or a combination of them. We end up with a total of 32 different models with different control variables and different lags of those control variables.

However, different from comparing the performance of different individual indicators, in this case is not straightforward to compare two different models since they could have different appealing characteristics. For instance, in general we find that the models in which we introduce our credit-gap lagged one period is more accurate than the models in which it enters with two periods lag. However, the latter ones could be more desirable in the sense that they provide an earlier warning and thus allow more time for policy makers to react before the risk of a crisis.

Therefore, after comparing those 32 different models we can obtain the following conclusions:

- The model that has the best out-of-sample performance i.e. that would have better predicted the new crises starting after 2008 is a model that only includes our credit-gap and the Libor interest rate, with both variables included with a one year lag.
- The models that better predict all the years of a crisis are models that do not include interest rates and only include our credit-gap and.
- The model that has an overall best performance (both in-sample and out-of-sample) is a model that includes the credit-gap with a one year lag, and all other control variables with a two year lag.
- The second best overall model is a model that includes the credit-gap and all other control variables with a two year lag. This model has also the most desirable characteristic of relying on earlier data, and is therefore, our preferred early warning model for banking crises.

In Table 2 we show the prediction performance of the best three models:

Table 2
Prediction performance of the best possible models.

	Model 1	Model 2	Model 3
Total in-sample 1990-2011			
ps-R2	0.20	0.23	0.18
NSR*	0.35	0.36	0.50
SNR*	2.61	3.12	3.01
Loss*	0.54	0.51	0.60
True Positive %	75%	72%	60%
False Positive %	29%	23%	20%
In-sample 1990-2007			
ps-R2	0.11	0.14	0.11
NSR*	0.51	0.45	0.60
SNR*	2.39	2.59	2.28
Loss*	0.64	0.59	0.69
True Positive %	62%	67%	55%
False Positive %	26%	26%	24%
Out-sample 2008-2011, All crises			
NSR*	0.26	0.29	0.44
SNR*	4.01	2.80	7.32
Loss*	0.41	0.49	0.48
True Positive %	79%	79%	60%
False Positive %	20%	28%	8%
Out-sample 2008-2011, New crises			
NSR*	0.13	0.21	0.16
SNR*	4.55	3.00	10.37
Loss*	0.30	0.43	0.23
True Positive %	89%	85%	85%
False Positive %	20%	28%	8%

Source: BBVA Research
NSR=(%Type II)/(1-%Type I); SNR=(%Type I)/(1-%Type II); %TP-FP=(1-%Type I) - (%Type II). Loss*=(%Type II)+(%Type I)

Model 1 is the model that displays the best overall performance ("in" and "out" of sample) in terms of NSR, SNR and Loss value. It has the following specification:

$$\mathrm{Prob}_{it} = \mathrm{f}(\mathrm{CreditGap}_{i,t-1}, \mathrm{Libor}_{t-2}, \mathrm{US\ GDP\ growth\ rate}_{t-2}, \mathrm{Liquidity}_{i,t-2}, (\frac{\mathrm{CA}}{\mathrm{GDP}})_{i,t-2})$$

However, this model includes the Credit Gap with only a one year lag and the rest of the variables with a two year lag.

Model 3 is the model that has the highest prediction performance regarding out-of-sample prediction of new crises, although again, it is based on information lagged one year. It has the following specification:

$$\mathrm{Prob}_{it} = \mathrm{f}(\mathrm{CreditGap}_{i,t-1}, \mathrm{Libor}_{t-1})$$

The model that is based only on variables with a two year lag and has the best prediction performance is Model 2, which has the following specification:

$$\mathrm{Prob}_{it} = \mathrm{f}(\mathrm{CreditGap}_{i,t-2}, \mathrm{Libor}_{t-2}, \mathrm{US\ GDP\ growth\ rate}_{t-2}, \mathrm{Liquidity}_{i,t-2}, (\frac{\mathrm{CA}}{\mathrm{GDP}})_{i,t-2})$$

Thus Model 2 is our preferred model for assessing the risk of a banking crisis.

DISCLAIMER