BBVA | RESEARCH

Working Paper n° 15/02
Madrid, January 2015

# Introducing a New Early Warning System Indicator (EWSI) of banking crises

Álvaro Ortiz Vidal-Abarca
Alfonso Ugarte Ruiz

# Introducing a New Early Warning System Indicator (EWSI) of banking crises*

Álvaro Ortiz Vidal-Abarca and Alfonso Ugarte Ruiz

January 2015

## Abstract

We introduce a new Early Warning System Indicator (EWSI) of banking crises based on a non-linear (Gomperzt curve) panel data model of credit deepening. The capability of our estimated credit-gap measure relative to alternative credit gaps computed through ad hoc procedures (linear or Hodrick Prescott trends) is tested in the context of alternative univariate and multivariate Early Warning Systems (EWS). Our new EWSI proves to be an outperforming and reliable leading indicator of banking crises while overcoming most of the problems of linear and stochastic procedures. The estimated credit gap outperforms the rest of indicators in both in-sample and out–of-sample forecasting accuracy (e.g. AUROC statistics) in a univariate comparison. Furthermore, we also test the importance of our EWSI in a multivariate framework through a Bayesian Model Average (BMA) technique, confirming our initial positive results. Finally, we estimate an Early Warning System for 68 developed and emerging countries based on our credit gap which can be used to compute Banking Crises Probabilities and to estimate dynamic thresholds for the EWSI indicators.

Keywords: Credit gap, Early warning indicators, Bayesian model averaging, macro-prudential policies.

JEL: C33, E44, E51, E58, F47, E61, G01, G21.

(*)BBVA Research Department, Cross-Country Emerging Market Analysis Unit. We thank Alicia García-Herreros, Gonzalo de Cadenas Santiago, David Martínez Turegano and Tomasa Rodrigo for helpful comments and suggestions, the same as all the participants in our BBVA-Research internal seminars. The views expressed are those of the authors and do not necessarily reflect those of BBVA-Research or BBVA.

# 1 Introduction and motivation

The recent financial crisis has reminded us the need to develop early measures of credit excess and systemic risk in order to anticipate and prevent the surge of banking and financial crises. Different economic studies have used a variety of indicators and statistical models to estimate the systemic risk stemming from excessive credit imbalances (see for instance Lund-Jensen (2012), Borio and Lowe (2002), Borio and Lowe (2009)).

Policy makers (BIS, IMF and European Commission)[1] have actually proposed macro-prudential tools based on different leading indicators of credit excess such as the Credit-to-GDP gap derived from a stochastic trend (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%).

However, we envisage several problems with these mechanical approaches. For instance, the existing credit ratio gaps suffer from potential end-of-sample unreliability problems, given that ex-post revisions are sizable and sometimes of the same order of magnitude of the gap itself. This could generate difficulties for bank supervisors in implementing such macro-prudential policies (see Edge and Meizenzahl (2011)). Additionally, the trends and gaps derived from a HP-filter or a linear trend are derived from an “ad hoc” (i.e. non- structural) approach and they do not account for idiosyncratic factors of the individual crisis. Further, as Saurina and Repullo (2011) have shown “their introduction of mechanical application of the buffer would tend to reduce capital requirements when GDP growth is high and increase them when GDP growth is low, so it may end up exacerbating the inherent pro-cyclicality of risk-sensitive bank capital regulation”.

The use of changes in the credit-to-GDP ratio as a risk indicator also presents several limitations. First, it does not account for structural differences as it assumes a common performance in all the countries independently of their idiosyncratic characteristics. It is usually considered that if such ratio grows more than 5 points in a year, a country is going through a credit boom. But in reality it does not carry the same risk to grow 5 points in an emerging country with an initial ratio of 20 than in an advanced economy with an initial ratio of 200. Similarly, it is difficult to compare the risk of an annual change in the credit ratio in an emerging economy with a 5% GDP growth rate than in an advanced economy with a 2% GDP growth rate. Finally, a simple rule like a 5 pp. growth does not take into account the effect of other variables, such as a recent regulatory change that favors financial development.

Other critiques have also been made to the use of Credit-to-GDP ratio growth as a business cycle indicator. First, there is the empirical regularity that credit usually lags the business cycle (Giannone, Lenza, and Reichlin, 2010). In particular, credit-to-GDP ratio continues to be high in downturns. Second, the use of deviations of the credit-to-GDP ratio with respect to its trend exacerbates the problem, because it takes some time before the ratio crosses the trend (Saurina and Repullo (2011). Other authors (Gadea-Rivas and Perez-Quiros ,2012) claim that when a given macroeconomic variable has the property of accumulating during expansion periods, a potential

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”.

bias may arise when including them in a binary response model because these variables usually present high levels just before the turning points[2].

In this paper we introduce an Early Warning System Indicator (EWSI) of banking crises based on a credit gap measure estimated through a non-linear (Gompertz curve) panel-data model, with the Private Credit-to-GDP ratio as the dependent variable. Herein, we exploit the estimates of the long-term structural level of the private credit ratio, which are based on the long-term level of several macroeconomic, regulatory and structural variables. Thus rather than using mechanical procedures we estimate our credit gap by subtracting the observed Credit-to-GDP ratio from the estimated long-term component[3].

One of the main goals of our study is to compare the performance of our credit gap measure as EWSI relative to other ad-hoc measures of excessive private credit level in a large sample of developed and emerging economies. In doing so, we also show that our new indicator displays some of the ideal characteristics of a EWSI as timing, stability, relative performance, robustness, interpretability and causality (see Drehman and Juselius, 2013).

Since most of the recent studies of banking crises highlight the private Credit-to-GDP ratio (through different transformations) as possibly the most important EWI, our initial focus is to compare our credit gap with other alternative indicators derived from the Credit-to-GDP ratio, by analyzing both their performance and their characteristics as EWIs. Our second and most broad objective is to construct a general model that has the highest predictive performance among those that can be constructed with our new indicator and a complementary set of both global and idiosyncratic early warning indicators of banking crises. In order to fulfill these objectives we follow a three-step methodological approach:

- First, we compare the forecasting accuracy of our credit gap in a *univariate* framework. Thus, we compare the ability of our indicator against that of other measures of credit excess (leverage) used by international institutions, such as credit ratio deviations from both linear (LT) and stochastic (HP) trends and simple credit-to-GDP changes.
- Second, we test the probability of inclusion of our credit gap in a multivariate model of banking crises through the use of a Bayesian Model Average (BMA) technique[4], evaluating its Posterior Inclusion Probability (PIP) relative to other credit excess indicators and to all other EWIs.
- Finally, we used the variables selected through the BMA to build up an Early Warning System Model. The model will allow us to calculate probabilities of Banking Crisis but also to compute dynamic early warning thresholds for risk analysis.

The results of this three step strategy proved to be satisfactory. The univariate tests show that our credit gap has the highest significance, the higher explanatory power and the highest in sample forecasting accuracy of the five alternative predictors in the whole estimation period (1990-2011) but also the highest out-of-sample performance in the period 2008-2011. Importantly, the superior performance of our new indicator holds after considering a wide set of statistical evaluation criteria and an extensive range of methodological variations.

2: For instance, they create an artificial random variable that grows during an expansionary phase but that is by construction unrelated to the onset of a future recession. However, they show that this artificial variable would appear to be a good predictor of a crisis or recession. Therefore, the apparent influence of credit could be due just to its build-up behavior in the expansion phase and not because it has actual predictive power on the business cycle.
3: 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. 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.
4: Babecky et al (2012) introduce the BMA analysis for their study of bank, currency and debt crisis of OECD economies.

Moreover, besides showing that our credit gap outperforms other measures of credit excess, we formally show the importance of some shortcomings of these indicators: i) we show how in a large sample of countries the posterior revisions to the gaps estimated by a HP-filter or a linear trend are many times larger than the estimated risk thresholds for these gaps; ii) the risk signal of the change in the credit-to-GDP ratio seems to be unstable since its correlation with the occurrence of a crisis is sometimes negative and other times positive depending on the number of lags this variable is introduced. Thus we can use our credit gap as a reliable EWSI while overcoming some of the main problems of the ad hoc mechanical procedures.

The multivariate tests based on the BMA analysis confirms that excessive credit measures are indeed good leading indicators of banking crisis as explored in previous empirical works. Besides, our credit gap appears at the top of all the measures of excess credit, confirming the univariate results in a multivariate framework. Besides our credit gap, the BMA analysis confirms the importance of the global factors (global interest rates and the US Global factors) as robust indicators of banking crisis not only in the case of OECD economies (as ) but also in the emerging economies. Additionally, domestic variables as current account deficits and banking liquidity (measured by the credit-to-deposits ratio) proved to be reliable indicators.

Thus, our multivariate results remind us of the need of monitoring different indicators at the same time. Since the effect of different variables could be compounded and interact with each other it is important to estimate “dynamic” thresholds for the credit gap if we want to use it as an early warning tool.

The rest of the paper is organized as follows: Section 2 reviews the main and most recent empirical literature related to our study. Section 3 analyzes some of the main shortcomings related to the existing credit excess indicators. Section 4 describes our methodological strategy. Section 5 shows and discusses the main findings. Section 6 describes the potential uses of our EWI and EWS as Risk Analysis Tools. Section 7 concludes.

# 2 Literature review

Most of the literature related to the prediction of banking crises and the development of Early Warning Systems (EWS) have flourished in cycles, according to the occurrence of periodic international banking crises such as the crises of the Nordic European countries at the beginning of the 90s, the Emerging crises in Mexico (1995), Asia and Latin America at the end of the last century, and more recently, after the large financial crisis in 2008.

Some of the earlier studies with similar objectives, geographical scope and methodological approach to ours are those of Eichengreen and Rose (1998), Hardy and Pazarbasioglu (1998), Demirgüç-Kunt and Detragiache (2000), and Demirgüç-Kunt and Detragiache (2005). Although these models also included developed countries in their samples, their results were mainly driven by the emerging-countries banking crises, since before 2008 the number of events in developed countries was relatively scarce. Most of these early studies explicitly acknowledged the lack of success in the out-of-sample evaluation of their models. Besides, most of them focused in Currency Crises[5] rather than in Banking Crises as they perceived banking problems as a result of the currency crises arising from the combination of sharp currency devaluations in fixed exchange regimes combined with liability dollarization problems. Thus, the devaluations triggered a rapid increase in non-performing loans and bankruptcies or bank runs which end up in banking crisis.

Eichengreen and Rose (1998) analyzed banking crises using a panel of macroeconomic and financial data for more than one hundred developing countries from 1975 through 1992. They found that banking crises in emerging markets were strongly associated with adverse external conditions. In particular, high interest rates in developed economies were strongly correlated with the onset of banking crises in developing countries, even after taking into account a host of internal macroeconomic factors.

Hardy and Pazarbasioglu (1998) examined different episodes of banking crises in a large sample of countries, both developed and developing. They identified several domestic macroeconomic and financial variables as useful leading indicators. However, they found that the main macroeconomic indicators were of limited value in predicting the Asian crisis; the best warning signs were proxies for the vulnerability of the banking and corporate sector.

Demirgüç-Kunt and Detragiache (2000) developed a multivariate logit model of the probability of a banking crisis to monitor banking sector fragility using a wide set of macroeconomic indicators. Their results pointed to a robust in-sample correlation and performance but a weak out-of-sample performance. Later on, Demirgüç-Kunt and Detragiache (2005) reviewed the two basic methodologies adopted in cross-country empirical studies, the signals approach and the multivariate probability model, and their application to study the determinants of banking crises. They found that at the time, empirical models have been more useful in identifying factors associated with the occurrence of banking crises than at predicting the occurrence of crises out-of-sample, reflecting the fact that, most empirical models were not conceived as forecasting tools.

The studies more closely related to ours are Borio and Lowe (2002), Borio and Drehman (2009) and Lund-Jensen (2012). Borio and Lowe (2002) presented empirical evidence that sustained rapid credit growth combined with large increases in asset prices appears to increase the probability of an episode of financial instability. They found the credit-to-GDP gap from an HP-filter

5: A summary of the Currency crisis in emerging markets can be found in Kaminsky, Lizondo and Reinhart 1998.

as the main indicator of excessive credit growth. Their analysis suggests that it is possible to construct simple composite indicators of banking crises that can be useful in assessing the risk of future financial distress with a reasonable degree of confidence. However, their analysis did not take into account the out-of-sample performance of the indicators, which could be the Achilles heel of the HP-credit-gap.

Davis and Karim, (2008) assessed the properties of a logit-model EWS and a signal-extraction EWS for banking crises on an extensive dataset of around 105 countries from 1979 to 2003. Their results suggest that the logit-model is the most appropriate approach for predicting global crises and the signal extraction methodology for country-specific crisis. They found that growth and terms of trade are robust leading indicators of banking crisis in their sample. However, the results for other variables were mixed in terms of robustness and varied according to the dataset used and the banking crisis definition adopted.

Borio and Drehman (2009) tried to correct the drawback of not analyzing the out-of-sample performance in the previous BIS document and so they focused on the out-of-sample performance of the same leading indicators of banking system distress evaluated before, extending them to incorporate property prices. They claimed that they were fairly successful in providing a signal for several recent banking systems crises, including that of the United States. Once again, they remarked the importance of the credit-to-GDP gap based on the HP-filter.

Lund-Jensen (2012) used a dynamic binary response model with country fixed effects to model the time varying conditional probability of a systemic banking crisis. He found that the level of systemic risk depends significantly on several risk factors: banking sector leverage, credit-to-GDP growth, changes in banks' lending premium, equity price growth, increasing interconnectedness in the financial sector and real effective exchange rate appreciation. He showed how to translate the systemic risk estimates into crisis signals and that this method provides accurate crisis signals in terms of type I and type II errors. He also argued against the signal extraction method, because of its constant thresholds: a crisis signal is issued if a variable is above a specific threshold, independent of the value of other systemic risk indicators. On the contrary, in a more realistic environment the appropriate threshold depends on the state of the economy as measured by a combination of other risk factors.

Babecky et al (2012) analyze the forecasting accuracy of several EWSI by introducing the BMA analysis for banking, currency, and Debt crisis in developed countries. Among the main results they confirm the ability of credit indicators but also the role of global variables. Our work follows a similar strategy although our credit gap is the result of estimations and our work extends to emerging markets.

The IMF reviewed the toolkit used at the IMF for evaluating and monitoring systemic ri in their Systemic Risk Monitoring "SysMo" project in 2013. They explained how IMF rely on different banking crisis prediction models, including binary response models *à la* Lund-Jensen (2012) or other threshold models. In the latter, they estimate a threshold value for each indicator minimizing its noise-to-signal ratio, and a weight is then assigned to each indicator based on its predictive power.

Drehman and Juselius (2013) evaluated the relative performance of different Early Warning Indicators (EWI) from the perspective of a macro-prudential policy maker. From this point of view, they established and described the ideal characteristics that an EWI should have and translated many of these requirements into statistical evaluation criteria. They found the credit-to-GDP gap

(HP filter) to be the best indicator at longer horizons, whereas the Debt Service ratio (DSR) dominates at short horizons

Finally, Drehman and Tsatsaronis (2014) review the main practical and conceptual criticisms of the credit-to-GDP gap as a guide to setting countercyclical capital buffers under Basel III. They claim that, historically, for a large cross section of countries and crisis episodes, the credit-to-GDP gap is a robust single indicator for the build-up of financial vulnerabilities. They also confirm that many criticisms have merit and consequently, its role must be to inform, rather than dictate, supervisors' judgmental decisions regarding the appropriate level of the countercyclical buffer.

Our study differs from all the previous studies in the sense that we first estimate an non- linear panel data model for the credit-to-GDP ratio from which we obtain the credit gap measure overcoming the usual ad hoc measures of credit gaps. We use the BMA procedure in the selection of variable as in Babecky et al (2012) but we extend our analysis to a wider set of countries including not only OECD countries but also Emerging Countries during the recent financial crisis. Additionally, and following our objective to prove that our indicator is in fact a better one, we explore a much wider set of methodological alternatives than other previous studies that also rely on binary response models.

# 3 Estimating the credit gap

Rather than relying on ad hoc filters to estimate the excess of credit measures we propose a methodology based on the idea that the long-term relationship between the Private Credit-to-GDP ratio and income per capita follows a non-linear relationship with a saturation level at the highest levels of income, i.e. a Gompertz-curve type of relationship[6]. This specification has the advantage of providing a smooth path for financial development with an explicit upper limit (saturation) and to obtain an estimation of the possible levels of economic development where the relationship between the variables changes[7].

Thus we first assume the following relationship between the credit ratio and income per capita:

$$\frac{C}{Y} = \alpha \cdot \exp(\gamma \cdot \exp(\beta \mathrm{Ypc}) \qquad (1)$$

Where $\alpha$ is the constant "maximum" saturation level. If there were no other variables in place, this is the level that a country will approach as long-term per capita income tends to infinity. $\gamma$ is the parameter that defines the curvature of the Gompertz curve and $\beta$ defines the sensitivity to income per capita,

We estimate the regression using non-linear maximum likelihood techniques. One caveat of this methodology is that there are no specific techniques designed to deal with panel data. However, we rely on the fact that using robust standard errors clustered by countries we are able to account for the possible serial-correlation within country-panels, and thus we are applying a similar approach to that of the GLS random-effects estimator. Thus, our methodological strategy is similar to using OLS with robust-cluster standard errors[8].

Another advantage of using a non-linear estimation technique is that we are able to account for an additional characteristic of the credit ratio that has usually been ignored by most of previous empirical studies, which is the fact that the ratio cannot take negative values. Considering the different methodological alternatives to deal with this problem[9], we follow Nichols (2010), Wooldridge (2010) and others and we assume a Poisson-like distribution of the dependent variable. Thus, we assume the following specification:

$$C/Y = \mathrm{e}[\alpha \cdot \exp(\gamma \cdot \exp(\beta \mathrm{Ypc})]] \qquad (2)$$

However, we do not specifically use Poisson regression, since the applied methodology already allows us to deal with non-linearities in the specification.

In the model we also allow to a different elasticities of the credit ratio to income per capita and to other explanatory variables could differ in the long run versus the medium or the short run. Thus we allow for different elasticities based on the notion that the structural relationship between credit depth and income per capita originates in a long-term process of development, whereas the actual credit ratio could diverge from such structural relation in the medium-term and in the short-run. We will compute our Credit Gap as the difference between the observed level of the credit ratio and the estimated "structural" level allows us to

6: A complete description of the model can be found in Ugarte (2014), BBVA Working Paper.
7: By contrast a linear regression would imply that the credit ratio could increase indefinitely as long as income per capita keeps on growing, since the elasticity $\beta$ is constant. Other alternatives as quadratic forms, would imply the unrealistic feature that credit disappears when countries attain very high-income levels. Another possibility for dealing with non-linearities is to perform split-sample regressions, but in this case we would need to know ex-ante the threshold points of income per capita where the sensitivity changes.
8: Although this may look initially as a caveat, it could actually be an advantage, since we are interested in estimating long-term structural coefficients that depend on the cross-country (between) variation that is swept away when using a Within-estimator. Or, in the case of using a GLS random-effects, it is known that the cross-country information contained could be ignored if the "between" variance is much higher than the "within" variance because in such case the GLS random-effect estimator tends to the "within" estimator. That would clearly be the case of the Private Credit-to-GDP ratio, since the cross-country variance accounts for nearly 85% of the total variance.
9: Log-regression, tobit-regression, negative-binomial regression and Poisson regression.

evaluate how far/close an economy is from a long-term equilibrium according to the stage of development. Therefore, we extend the specification shown in (2) and include different sensitivities to income per capita:

$$\frac{C}{Y} = \exp[\alpha \cdot \exp(\gamma \cdot \exp(\beta_{LT}\overline{Ypc}_{it} + \beta_{MT}\widetilde{Ypc}_{it} + \beta_{ST}\widehat{Ypc}_{it})] \quad (3)$$

In (3) $\overline{Ypc}_{it}$ represents the long-term (15 years) moving average of GDP per capita, $\widetilde{Ypc}_{it}$ represents the medium-term deviation of income per capita with respect to its long-term level, i.e. $\widetilde{Ypc}_{it} = (\overline{Ypc}_{it}^{5yr} - \overline{Ypc}_{it}^{15yr})$, and $\widehat{Ypc}_{it}$ represents the short-term deviation of the observed income per capita with respect to its medium-term (5-years) moving average, i.e. $\widehat{Ypc}_{it} = (Ypc_{it} - \overline{Ypc}_{it}^{5yr})$. Therefore, $\beta_{LT}$ , $\beta_{MT}$ and $\beta_{ST}$ represent the long, medium and short-term sensitivities to per-capita inco(3)me respectively.

In addition to the different sensitivities of the credit ratio to income per capita, we also estimate different sensitivities to other macroeconomic variables according to the time-horizon components. Moreover, the saturation level and the shape of the relationship between financial deepening and income should depend on institutional and regulatory determinants such as creditors' protection, information sharing, banking structure, the long-term evolution of interest rates and so on. Therefore, within the Gompertz-curve framework, we allow each country to have a different saturation level that depends on the long-term level of institutional and structural variables:

$$\frac{C}{Y} = \exp[\{\alpha\} \cdot \exp(\gamma \cdot \exp(\beta_{LT}\overline{Ypc}_{it}^{15yr} + \beta_{MT}\widetilde{Ypc}_{it} + \beta_{ST}\widehat{Ypc}_{it}) + \emptyset_{LT}\bar{X}_{it}^{15yr} + \emptyset_{MT}\tilde{X}_{it} + \emptyset_{ST}\hat{X}_{it} + \eta_i + \upsilon_{it}]]$$

We define the credit gap as the difference of the current Credit-to-GDP ratio with the "structural" part of the income per capita, i.e. to its long-term level:

$$CreditGap_{i,t} = \frac{C}{Y} - [\exp[\alpha \cdot \exp(\gamma \cdot \exp(\beta_{LT}\overline{Ypc}_{it}) + \emptyset_{LT}\bar{X}_{it}^{15yr}] \quad (4)$$

# 4 Examples of shortcomings of existing indicators

In this section we briefly introduce some of the problems arising in alternative measures of credit gaps, particularly those based in mechanical ad hoc filters or credit to GDP growth rates and how our credit gap is able to deal with them.

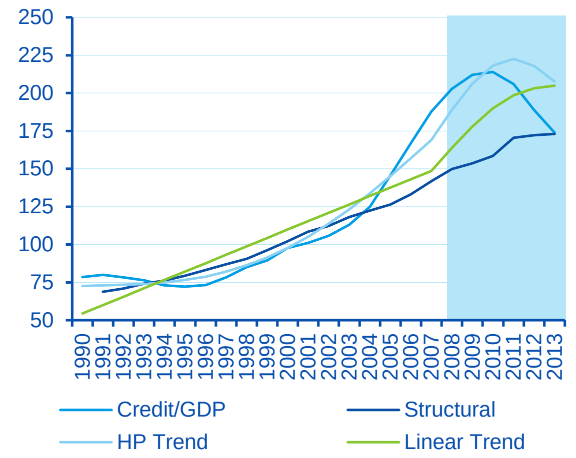
The following graphs (Chart 1 and Chart 2) shows an example of how the available indicators of excessive credit relate to the onset of banking crises in the case of Spain, and how do they compare with our new proposed indicator.

In the first example in Chart 1 we can see the Credit-to-GDP ratio of Spain together with our "credit gap" ratio estimated through our panel-data methodology, and the estimated trends by a Hoddrick-Prescott filter (HP-trend) and a linear trend. In all the cases, the credit gap, the HP-trend and the linear-trend were estimated in real time. To do this, first, we estimate 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 data underlying panel-data model, the HP-filter and the linear-trend in real time. Therefore, we use data from 1990 up to the corresponding year and we add the resultant estimated value to the initial series obtained up to 2006 completing the series up to 2012.

The real time analysis is useful to evaluate the magnitude of the end-of-sample problem of the alternative methodologies. In Chart 2 we can observe the estimated gaps between the observed Credit-to-GDP ratio and each one of the estimated trends. We can also observe the annual change in the Credit-to-GDP ratio. We can point out the following observations according to the illustrative example of the Spanish case:

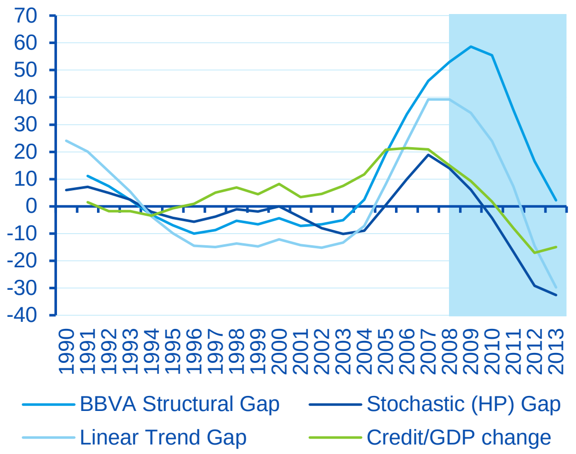
- The HP-filter generates a trend that is closer to the actual ratio than the linear trend or our estimated structural ratio and thus, the corresponding gap is lower.
- The Linear-trend, and specially the HP-filter change radically their slope after 2008, as a consequence of the previously mentioned "end-of-sample" problem. Our estimated structural ratio does not change much after 2008 even though each point after 2007 is the result of a different regression estimation.
- We can also observe one of the clear shortcomings of using the annual change in the Credit-to-GDP ratio as a leading indicator: The Credit-to-GDP ratio grew more than 5 pp of GDP several times and long before 2008. For instance, in 1998 the ratio grew 7 pp and in 2000 it grew 8.2 pp. However, in these years our estimated credit gap was negative or very close to zero, and therefore the likelihood of a crisis was small despite the rapid growth in the credit ratio.
- The estimated credit gap anticipates well the worst part of the crisis, since it reaches its highest level in 2009 and 2010, meanwhile all the other indicators were already decreasing just after 2008.

Figure 1
**Spain: Credit-to-GDP and alternative trend estimations**




Note: E&N exCh = EAGLEs and Nest countries excluding China
Source: BBVA Research

Figure 2
**Spain: Credit-to-GDP and alternative trend estimations**




Source: BBVA Research

These revisions can be sizeable as we can observe from figures in Table 1. We can see that on average (across the 83 countries), the gap estimated for 2006 would have been revised -4.4 pp when re-estimated in 2012 with the information available until 2012. Similarly the 2007 gap would have been revised -3.0 pp when estimated in 2012 versus the estimated value in 2007. Contrary, the revisions on our estimated credit gap are significantly lower, with an average revision of -0.7% pp of the value estimated in 2006 relative to the final re-estimations in 2012. Similarly the value estimated in 2007 would have been revised -0.9 pp when re-estimated in 2012.

Table 1
**Average revisions in alternative credit-to-GDP ratio gaps after re-estimating them with newer information**

| | | Average Revision | | | Average Revision |
|---|---|---|---|---|---|
| Credit Gap 06 | Revision 2012 vs. 2006 | -0.7 | Credit Gap 07 | Revision 2012 vs. 2007 | -0.9 |
| | Change in pp of threshold (10) | -7% | | Change in pp of threshold (10) | -9% |
| HP Gap 06 | Revision 2012 vs. 2006 | -4.4 | HP Gap 07 | Revision 2012 vs. 2007 | -3.0 |
| | Change in pp of threshold (3) | -157% | | Change in pp of threshold (3) | -108% |

Source: BBVA Research

In order to understand the magnitude of these revisions we can compare the value of the revisions to the level of the risk thresholds that would trigger a crisis signal. This level has been estimated by different studies for the HP credit gap. For instance, Borio et al (2002) estimate a risk threshold of 2.3 pp and Lund-Jensen (2012) estimate a 4.4 pp as the optimal individual risk-threshold. Following a methodology we estimate an optimal risk threshold of 2 pp. In all the cases, the average revisions for the year 2006 are at least the same size of the risk thresholds. What this means is that, on average for the 83 countries, an alert signal given by the HP credit gap would have not been emitted if we had had the entire posterior information from the subsequent years.

On the contrary, if we estimate an optimal risk threshold for our credit gap we get that it would be approximately 10 pp. Therefore, the average revisions of the 2006 and 2007 gaps would not only be much smaller than the HP ones, but they would only represent 7% and 9% of the risk threshold. This means that the signals given in 2006 and 2007 would not change after including all the new information coming in the posterior years.

Last, but not least, our measure also reduces the problem of pro-cyclicality of risk-sensitive bank capital regulation signaled by Saurina and Repullo (2011). The following table shows the cross correlation coefficients of the real GDP growth with the different measures of credit-gaps for all the countries, including individual coefficients for both developed and emerging markets and both crisis and non-crisis periods. As can be observed the pro-cyclicality problem is lower for our estimated Gap than in the rest of gap measures and in line with the credit gap change.

Table 2
**Correlation of alternative credit Gap measures with Real GDP Growth (DEV: Developed Economies, EM: Emerging Economies)**

| | All Total | All Crisis | All Non Crisis | DEV Total | DEV Crisis | DEV Non Crisis | EM Total | EM Crisis | EM Non Crisis |
|---|---|---|---|---|---|---|---|---|---|
| Estimated Credit Gap | -0.11 | -0.22 | 0.07 | -0.29 | -0.37 | 0.00 | -0.05 | -0.18 | 0.08 |
| Credit Gap (HP Filter) | -0.21 | -0.28 | -0.11 | -0.3 | -0.39 | -0.19 | -0.20 | -0.26 | -0.11 |
| Credit Gap (Linear Trend) | 0.22 | -0.29 | -0.12 | -0.35 | -0.40 | -0.20 | -0.20 | -0.27 | -0.10 |
| Credit GDP Change | 0.03 | -0.18 | 0.07 | 0.03 | -0.32 | -0.12 | 0.09 | -0.14 | 0.13 |

Source: BBVA Research

# 5 Methodological strategy

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, $X_{i,t-j}$, such that the probability of a systemic banking crisis, in country i, can be written as.

$$\mathrm{Prob}\left(Y_t = 1 \middle| X_{i,t-j}, \beta\right) = \frac{\exp(\alpha + X_{i,t-j}{}'\beta + \varepsilon)}{(1 + \exp(\alpha + X_{i,t-j}{}'\beta + \varepsilon))} \quad (5)$$

The dependent variable is a binary variable equal to one if there is a crisis in the period t, and the systemic risk factors, $X_{i,t-j}$ are known *j* periods in advance, in order to be able to provide an early warning signal. For the dependent variable we follow the definition and of a crisis and the dates established by Laeven and Valencia (2012), although in the cases in which we do not have information from them, we use the crises as defined in Lund-Jensen (2012) and Reinhart and Rogoff (2010).

The purpose of the analysis is not to replicate or to 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. Additionally, we also follow Drehman and Juselius (2013) and incorporate the AUROC as the key measure of predictive performance.

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.

In order to accomplish our objectives we follow a three-step methodological approach:

- First, we check the forecasting accuracy of our "Credit Gap" against that of other measures of credit excess (leverage) used by international institutions, such as credit ratio deviations from both linear (LT) and stochastic (HP) trends and simple credit-to-GDP changes.
- Second, we examine the robustness of the individual results in a multivariate framework by testing the posterior inclusion probability of inclusion of our credit gap in a multivariate model of Banking Crises through the use of a Bayesian Model Average (BMA) technique.[10]
- Finally, and based on the results of the BMA analysis, we run and compare several multivariate models combining all the different possible EWIs, looking for the one with the best overall performance and the most desirable characteristics.

## 5.1 Comparing Individual Performance

Research economists and macro-prudential policy makers have tested several alternative indicators of excessive credit growth as leading indicators of banking crises, using different transformations of the Credit-to-GDP ratio. Among them we can consider the annual change in the Credit-to-GDP ratio or Credit-to-GDP gaps derived from extracting the trend to the Credit-to-GDP series by alternative procedures as linear or stochastic (i.e Hodrick Prescott) trends. Alternative one could consider the annual change in the Credit-to-GDP ratio, but considering only the values higher than a certain threshold, for instance the values higher than 5 points.

10: See Babecki et al (2012) for a similar approach.

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 these alternative variables as explanatory variables (introducing their lagged values), comparing their performance with the performance of our credit gap. Our empirical exercise comprises several possible methodological possibilities:

- 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 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 the crisis is still in place. The rationale for this restriction is that the goal of the analysis can be just to anticipate a crisis and not its duration. However, we also test the possibility of including all the crisis years for several reasons. First, It could be difficult to date precisely the exact time of crises. Besides, some crisis can last several years, but the start of it could be rather mild, while the worst impact could emerge later on. As a consequence, the difficulty of dating the precise moment of time when a crisis starts is one of the most important criticisms to the use of parametric estimations such as binary response models.
- Different lags of the leading indicators: As we are testing the ability of our credit gap measure as a Early Warning System we will 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, two or three year lags. We also estimate the option of including the average of the previous two lags of all the possible indicators.

To take into account all of the possibilities described above, 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 under each methodological variation: 1) the z-statistic, 2) the pseudo-R-squared, 3) the Noise-to-Signal Ratio, 4) The loss function or the percentage of missed crises (% type I error) plus the percentage of false signals (% type II error) and 5) the Area Under Receiving Operating Characteristis, (AUROC).

The Noise-to-Signal Ratio and the Loss Function 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 minimizes the loss function, i.e. the sum of the percentage of false positives plus the percentage of false negatives[11].

In summary, we run 15 regressions for each indicator, combining all the possible variations of lags, fixed vs. random effects and different definitions of the dependent variable. However, considering the large amount of possible methodologies there could also be a large amount of discrepancies about which indicator performs better according to each statistic and according to different methodologies.

11_ In an early version of this paper we also computed both the Signal-to-Noise ratio and the Noise-to-Signal ratio since it can be shown that these ratios could lead to very different conclusions.

Therefore, we summarize and compare the results of the 15 different regressions through an overall “score” that assigns a higher value to the indicator that obtains a higher rank among the five indicators considered. More specifically, the score is calculated in the following way: the best indicator in every regression according to each statistic receives a score of 4, the second a score of 3, the third a score of 2, the fourth a score of 1 and the fifth a score of 0. This score is then summed across all the possible statistical measures and methodologies.

Additionally, we repeat the exercise several times in order to be able to compare the performance of the five indicators in both the in-sample and out-of-sample estimations. We first estimate 15 different regressions with each indicator as an individual explanatory variable for the whole sample, i.e. with data from 1991 to 2011. Next, we estimate the same 15 regressions but restricting the sample to the years before the recent global financial crisis of 2008, i.e. using only data from 1991 to 2007. After each one of the latter regressions we estimate the corresponding in-sample performance statistics and calculate the corresponding score.

Then, we use the estimates from the restricted regressions (1991-2007) to predict the probability of a crisis in the period 2008-2011. With these out-of-sample predictions we compute the same statistical measures of performance than in the in-sample analysis and we also compute their corresponding out-of-sample score.

With the whole set of results from the in-sample and out-of-sample exercises we calculate a final score that gives us a clear and precise measure of which one is the best indicator out of the five considered indicators.

In this stage 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. Most studies consider a signal to be correct if a crisis occurs in a certain window of one to three years before and/or after the crisis. In this individual comparison stage we do not want to consider a window of more than one year since we are actually comparing the performance of each indicator under different lags specifications and such approach would alter the results.

## 5.2 Multivariate Test: Bayesian Model Averaging (BMA)

It is well-known that any univariate regression results may vary if we take into account a wider set of control variables. Thus, in a second step we check the robustness of the first stage results by testing the usefulness of our credit gap in a multivariate framework, when controlling for several other possible explanatory variables.

In this step we also want to identify the most useful early warning indicators of banking crisis by means of a Bayesian Model Averaging (BMA) technique. 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.

Bayesian Model Averaging (BMA) technique takes into account model uncertainty by considering various model combinations and thus has the advantage of minimizing the author’s subjective judgment in determining the optimal set of early warning indicators. Among them we include all the Credit Gaps measures and domestic variables as GDP growth, inflation, interest rates, liquidity (Credit-to-Deposit ratio), public debt to GDP and the current account balance. Besides, we also include external variables as the USA GDP growth rate, Libor Rate and the global stock volatility index (VIX)

The results of the BMA analysis for these indicators will be summarized in the Posterior Inclusion Probability (PIP). This will help us to identify the ability of the any indicator as explanatory variable, it will provide us with a ranking of the importance of all the indicators in explaining the dependent variable and will identify the best possible combinations of the available indicators.

## 5.3 Choosing a final model as an Early Warning System (EWS)

Once we select the best candidate variables for a multivariate EWS through the BMA analysis, we proceed to select the best possible model among different combinations of those final candidates. Since the possible combinations are plentiful, we run a similar exercise than in the first stage. Thus, we estimate a large number of models, computing and comparing for each one of them a number of statistics measuring their performance as crisis' predictors. Once again, we consider both in-sample and out-of-sample performance. We combine different lags of our credit gap and of the following control variables: Libor interest rate, GDP growth rate in the US, Credit-to-Deposits ratio (banking liquidity and the current account balance as percentage of GDP

Finally, we end up with a total of 32 different models with different control variables and different lags of those control variables. Checking the performance of different models is not straightforward as in the individual analysis procedure. For instance, in general we find that the models in which we introduce our credit gap with a one period lag are more accurate than the models in which it enters with a two periods lag. However, the latter type of models could be more desirable in the sense that they provide an earlier warning and thus allow more time to policy makers to react before the risk of a crisis.

Therefore, in the information taken into account to choose a final model we also consider some "subjective" criteria that could be important from a macro-prudential perspective.

.

# 6 Summary of results and empirical findings

In this section we explain in detail the results of each one of the three steps described in the methodological section

## 6.1 Individual Comparison Results

### In Sample Analysis

In Table 3 we can see the results of the one of the 15 possible methodological variations considered in our empirical exercise. In this case we show the results of the least restrictive specification for each one of the five indicators[12]. The least restrictive specification 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 this particular specification we can see that our Credit Gap is the variable with the highest significance (as measured by the z-statistic), the highest pseudo R-square and the highest AUROC. In this specific example the HP-Gap is the indicator with the lowest NSR and Loss function.

Table 3
**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 |
| Loss* | 58% | 55% | 76% | 76% | 55% |
| AUROC | 0.770 | 0.76 | 0.59 | 0.63 | 0.766 |
| True Positive % | 59% | 73% | 42% | 37% | 68% |
| False Positive % | 17% | 28% | 18% | 14% | 23% |

NSR=(%Type II)/(1-%Type I);. Loss*=(%Type II)+(%Type I). The NSR, Loss, True Positive and False Positive values are estimated at the optimal cut-off probability that minimizes the loss function.
Source: BBVA Research

### Out of Sample Analysis

The results shown in Table 3 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 first run the same regressions but using only information up until 2007. Afterwards we use the in-sample coefficients estimated in the period 1991-2007 to compute 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

12: As explained before, we have run 15 regressions for each indicator, combining all the possible variations of lags, fixed vs. random effects and different definitions of the dependent variable. This means that we actually have 14 alternative results to the ones shown in Table 3. Later on, we summarize and compare the results of the 15 different regressions.

In order to make the out-of-sample exercise as realistic as possible, the structural credit ratios used to estimate the credit gaps (structural panel, linear-trend and HP-trend) from the years 2008 to 2011 are all calculated using only information available up to the corresponding year.

In Table 4 we can see the same in-sample results as the ones shown in Table 3, but for the more restricted period of 1991-2007. For this sub-sample period the in-sample performance of the gap calculated based on a linear trend (LT-Gap) is slightly better than our credit gap's performance. The use of credit to GDP changes would have resulted in disappointing results. 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 and in this particular example, we might have disregarded the change in the credit ratio as a useful predictor of a crisis, although ex-post it would appear to be a useful one.

Table 4
**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 |
| Loss* | 67% | 59% | 84% | 92% | 58% |
| AUROC | 0.72 | 0.74 | 0.55 | 0.54 | 0.74 |
| True Positive % | 57% | 65% | 36% | 17% | 65% |
| False Positive % | 24% | 23% | 20% | 9% | 23% |

NSR=(%Type II)/(1-%Type I);. Loss*=(%Type II)+(%Type I). The NSR, Loss, True Positive and False Positive values are estimated at the optimal cut-off probability that minimizes the loss function.
Source: BBVA Research

In Table 5 we can see the results of the out-of-sample prediction of the onset of new crises between 2008 and 2011. In Table 6 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, i.e. all years of a crisis. The results in Table 5 show that our credit gap is by far the best predictor of new crises, since it displays the lowest NSR, the highest AUROC and the lowest loss value. In Table 6 we can see that it is also the best predictor of all the years of a crisis, since it displays the lowest NSR, the highest AUROC and the lowest loss value.[13]

13: As explained in the in-sample analysis, we have repeated the whole exercise shown in Table 3 to Table 6 a total of 15 times, one for each possible methodological variation discussed previously.

Table 5

**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* | 0.25 | 0.31 | 4.58 | 0.38 | 0.32 |
| Loss* | 40% | 66% | 104% | 52% | 69% |
| AUROC | 0.81 | 0.66 | 0.20 | 0.76 | 0.66 |
| True Positive % | 80% | 85% | 95% | 70% | 85% |
| False Positive % | 20% | 51% | 99% | 22% | 54% |

NSR=(%Type II)/(1-%Type I);. Loss*=(%Type II)+(%Type I). The NSR, Loss, True Positive and False Positive values are estimated at the optimal cut-off probability that minimizes the loss function.
Source: BBVA Research

Table 6

**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 |
| Loss* | 46% | 70% | 112% | 58% | 72% |
| AUROC | 0.85 | 0.76 | 0.22 | 0.76 | 0.76 |
| True Positive % | 80% | 88% | 4% | 55% | 88% |
| False Positive % | 26% | 58% | 16% | 14% | 60% |

NSR=(%Type II)/(1-%Type I);. Loss*=(%Type II)+(%Type I). The NSR, Loss, True Positive and False Positive values are estimated at the optimal cut-off probability that minimizes the loss function.
Source: BBVA Research

## Computing and comparing overall prediction performance

In order to increase robustness we have perform two different exercises. First, in Table 7 we show the average of each statistic across the 15 methodological variations for each one of the possible predictors, in all the in-sample and out-of-sample cases. We highlight in blue cells the best average statistic among the five variables. The main results are the following:

- Our credit gap has the highest significance, the higher explanatory power and the best prediction performance of the five possible predictors if we consider the whole estimation period, 1990-2011.
- Our credit gap has the best prediction performance in the pre-crisis period of 1990-2007, i.e. best AUROC, NSR and Loss function. However, the HP-gap and the Linear-trend gap show better statistical properties in this in-sample restricted period, higher R-squared and higher z-statistics. Meanwhile, the change in the Credit-to-GDP ratio has the worst statistical and prediction performance in this period.
- The credit gap has the best overall out-of-sample prediction performance in the period 2008-2011. It predicts better all the years in which a crisis occurs and anticipates better the onset of new ones according to all the estimated statistics. Importantly, our credit gap displays the higher average AUROC across all the methodological variations considered, in the overall, in the in-sample and out-of-sample cases.

Table 7

**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 | 8.93 | 8.18 | 3.83 | 5.93 | 8.30 |
| ps-R2 | 0.19 | 0.15 | 0.05 | 0.07 | 0.15 |
| AUROC | 0.76 | 0.72 | 0.62 | 0.64 | 0.72 |
| NSR* | 0.44 | 0.49 | 0.67 | 0.69 | 0.49 |
| Loss* | 0.60 | 0.63 | 0.75 | 0.75 | 0.63 |
| **In-sample 1990-2007** | | | | | |
| z-stat | 5.84 | 5.90 | 0.67 | 2.31 | 5.97 |
| ps-R2 | 0.12 | 0.12 | 0.02 | 0.02 | 0.12 |
| AUROC | 0.71 | 0.70 | 0.59 | 0.56 | 0.70 |
| NSR* | 0.50 | 0.56 | 0.75 | 0.79 | 0.55 |
| Loss* | 0.65 | 0.67 | 0.81 | 0.86 | 0.66 |
| **Out-sample 2008-2011, All crises** | | | | | |
| AUROC | 0.83 | 0.72 | 0.56 | 0.68 | 0.73 |
| NSR* | 0.29 | 0.37 | 0.70 | 0.67 | 0.36 |
| Loss* | 0.53 | 0.73 | 0.83 | 0.75 | 0.73 |
| **Out-sample 2008-2011, New crises** | | | | | |
| AUROC | 0.77 | 0.66 | 0.56 | 0.77 | 0.65 |
| NSR* | 0.39 | 0.49 | 0.66 | 0.52 | 0.50 |
| Loss* | 0.60 | 0.80 | 0.80 | 0.67 | 0.80 |

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

Secondly, we have computed a score that summarizes all the statistics obtained for each indicator and allows us ranking them according to their performance across all the methodological variations considered. It is calculated in the following way: the best indicator in every regression according to each statistic receives a score of 4, the second best a score of 3, the third one a score of 2, the fourth a score of 1 and the fifth a score of 0. The statistics considered are the z-stat, pseudo R-square, NSR, the loss value and AUROC.

The total scores corresponding to the example of Table 3 can be observed in Table 8. The overall score based on all estimated regressions is shown in Table 9. Finally, in Table 10 we show the score based only on the AUROC of each regression. The latter is important since in theory, the AUROC is the measure with the most desirable characteristics.

In Table 8 to Table 10 we can see that our credit gap obtains the highest score when we calculate the overall score (based on both the in-sample and out-of-sample statistics). However, in the specific example shown in Table 8 our credit gap shares the first position with the LT-gap. Nevertheless, it gets the highest possible score in the out-of-sample exercises and more importantly it beats all the other indicators if we measure the performance based only on the AUROC.

The most important results are the ones in Table 9 and Table 10. In both cases we can see that our credit gap has the highest total score and the highest score in the in-sample cases and the out-of-sample cases. Thus, after considering several methodological variations such as different lags of the indicators, the use of random or fixed effects and restricting the dependent variable or not, our credit gap outperforms all the other indicators in the sample from 1991 to 2011, in the sample from 1991 to 2007, and it outperforms the other indicators in the out-of-sample exercises, in both predicting the start of new crises and in predicting all the years in which a crisis is ongoing.

Table 8
**Score of Each Indicator in the case of unconditional Logit, two-year lag (t-2), and including all years of a crisis. Period 1991-2011**

| | Credit Gap | HP- Gap | Credit/GDP change | Credit/GDP change>5 | LT-Gap | Max |
|---|---|---|---|---|---|---|
| In sample total | 16 | 13 | 1 | 4 | 16 | 50 |
| In sample 1990-2007 | 11 | 14 | 3 | 2 | 20 | 50 |
| Out of sample (All) | 12 | 6 | 0 | 6 | 6 | 30 |
| Out of sample (First) | 12 | 6 | 0 | 7 | 5 | 30 |
| All indicators | 51 | 39 | 4 | 19 | 47 | 160 |

Source: BBVA Research

Table 9
**Final performance score based on all statistics across all 15-methodological variations**

| | Credit Gap | HP- Gap | Credit/GDP change | Credit/GDP change>5 | LT-Gap |
|---|---|---|---|---|---|
| TOTAL | 834 | 581 | 260 | 298 | 610 |
| In sample | 276 | 178 | 48 | 79 | 219 |
| In sample 2007 | 239 | 218 | 55 | 53 | 232 |
| Outsample (All) | 182 | 89 | 60 | 51 | 96 |
| Outsample (First) | 137 | 96 | 97 | 115 | 63 |

Source: BBVA Research

Table 10
**Final performance score based on AUROC across all 15-methodological variations**

| | Credit Gap | HP- Gap | Credit/GDP change | Credit/GDP change>5 | LT-Gap |
|---|---|---|---|---|---|
| TOTAL | 225 | 118 | 82 | 84 | 130 |
| In sample | 61 | 31 | 14 | 17 | 37 |
| In sample 2007 | 49 | 45 | 16 | 4 | 45 |
| Outsample (All) | 64 | 27 | 19 | 16 | 34 |
| Outsample (First) | 51 | 15 | 33 | 47 | 14 |

Source: BBVA Research

Additionally, our credit gap displays other interesting characteristics. As can be observed ( Table 13 to Table 18 in Appendix A1), the relative outperformance of our credit gap improves respect to the two alternative gaps (HP and Linear-trend) the longer the lag such gap is introduced in the regression, thus the credit gap outperforming capacity increase with the number of lags which signals its good properties as an early warning tool which provides a longer period of time to implement policies to correct excess credit growth.

Finally, this individual comparison exercise has also suggests that the optimal lag of our indicator in a regression is a two-year one.

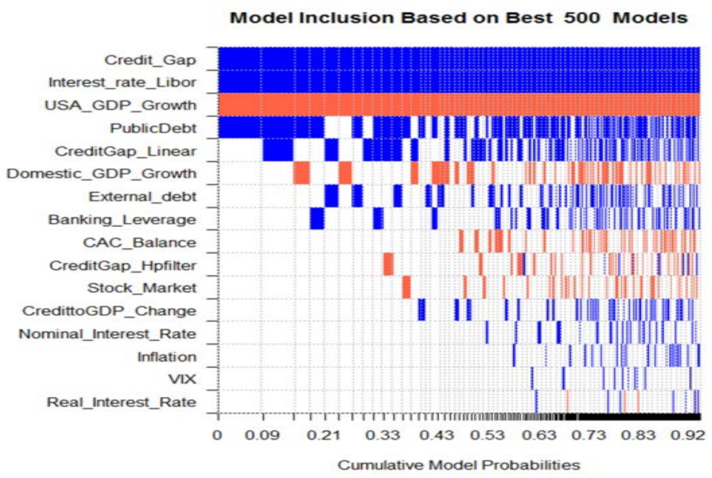
## 6.2 Multivariate tests: Bayesian Model Averaging (BMA) Results

The results of the BMA analysis can be checked In Chart 3 and Chart 4. In these, 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 Chart 3 we include our credit gap together with the rest of the alternative credit excess indicators in order to validate the individual our previous results in a multivariate framework. The results in Chart 3 clearly indicate that the indispensable variables in the analysis are our credit gap, the Libor interest rate and the GDP growth of USA (they display an estimated PIP equal to one). The analysis also shows that our credit gap is ranked as more informative than all other indicators and particularly to all the other indicators of excessive credit (linear-trend credit gap, HP credit gap and credit-to-GDP change), confirming its superiority in a multivariate framework.

Figure 3

**Posterior Inclusion Probability in a model of Banking Crisis Including all Alternative Gaps (explanatory variables introduced with a 2 year lag, blue=positive sign, red=negative sign)**

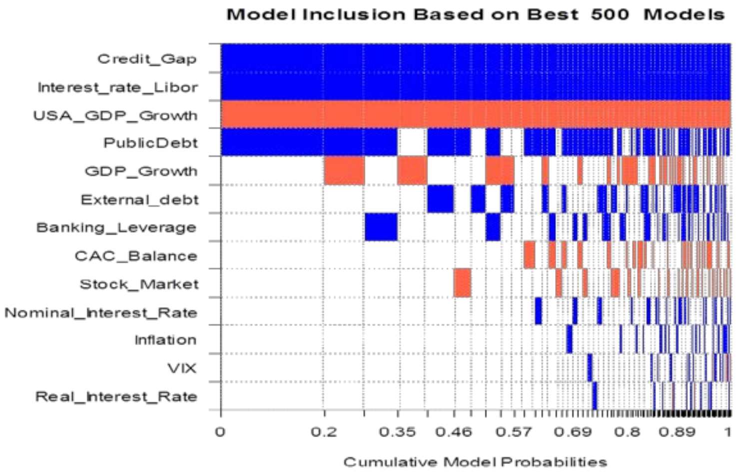



Source: BBVA Research

The results in Chart 3 clearly indicate that the indispensable variables in the analysis are our credit gap, the Libor interest rate and the GDP growth of USA (they display an estimated PIP equal to one). The analysis also shows that our credit gap is ranked as more informative than all other indicators and particularly to all the other indicators of excessive credit (linear-trend credit gap, HP credit gap and credit-to-GDP change), confirming its outperformance in a multivariate framework.

In Chart 4 we include only our estimated gap and the rest of the available EWIs. The different colors stand for the different sign of the explanatory variables (blue=positive sign, red=negative sign).

Figure 4
**Posterior Inclusion Probability in a model of Banking Crisis including Structural Gap only (explanatory variables introduced with a 2 year lag, blue=positive sign, red=negative sign)**




Source: BBVA Research

The results in Chart 4 confirm the superiority of our credit gap over the rest of indicators considered in the BMA analysis. It also confirms that the three variables regarded as indispensable are our credit gap, the Libor interest rate and the GDP growth of USA.

Other variables appear as important such as the public debt-to-GDP ratio, the GDP growth rate, the external debt-to-GDP ratio, the credit-to-deposits ratio (banking leverage) and to a lesser extent the CAC balance-to-GDP ratio. Although good coincident indicators , their properties as leading indicators is much lower than the indicators mentioned above.

This leaves us with our credit gap, the Libor interest rate and the GDP growth of USA, the credit-to-deposit ratio (banking leverage) and the CAC balance as the first variables ranked according to their PIP. Consequently, these are the variables that we are going to include in our final model. The BMA analysis also confirms several interesting features. First, our results highlight the importance of the global factors as robust indicators of banking crisis (Libor interest rate and US Global GDP growth). Global Market Risk indicators (VIX) can be suitable coincident indicators or triggers of banking crisis but could be less useful as leading indicators which has to be introduced in the regressions with some lags.

## 6.3 Choosing a Final Model as an Early Warning System (EWS)

Once we have checked the robustness and outperformance of our estimated Credit Gap, we also want to find out which is the best possible model that we can obtain using our credit gap and the other EWIs identified through the BMA analysis as the credit-gap, the Libor interest rate, the GDP growth rate of the USA, the credit-to-deposits ratio and the current account balance.

To do that, we compare 32 different models combinations of this final variables and their lags. The final model is 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 Table 11 we show the forecasting accuracy of the three best estimated models according to the wide of the window at which we evaluate the accuracy of the signal: A window of 0 corresponds to the case in which we only consider a signal to be correct or false if a crisis occurs or not in the exact same year a signal is predicted for. For instance, if the model predicts that a crisis will occur in the year 2010, we only consider such signal to be correct if a crisis actually occurs in 2010, independently of what happens in the year 2009 or 2011. A window of 1 corresponds to the case in which we allow the signal to be correct/false taking into account the outcomes in a window of 1 year after/before the actual year a signal is predicted for. Obviously, the wider the window, the better the prediction performances because we allow the signals a larger opportunity to be correct.

Table 11
**Prediction performance of the three best estimated models, according to the wide of the window of years the accuracy of a signal is evaluated at**

| | Model 1 | | | | Model 2 | | | | Model 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | W - 0 | W - 1 | W - 2 | W - 3 | W - 0 | W - 1 | W - 2 | W - 3 | W - 0 | W - 1 | W - 2 | W - 3 |
| **Total in-sample 1990-2011** | | | | | | | | | | | | |
| ps-R2 | 0.20 | | | | 0.23 | | | | 0.18 | | | |
| AUROC | 0.78 | 0.85 | 0.89 | 0.91 | 0.81 | 0.87 | 0.90 | 0.90 | 0.77 | 0.83 | 0.87 | 0.89 |
| NSR* | 0.40 | 0.27 | 0.21 | 0.18 | 0.35 | 0.23 | 0.18 | 0.15 | 0.49 | 0.37 | 0.25 | 0.23 |
| SNR* | 3.26 | 4.34 | 5.33 | 7.07 | 3.16 | 4.05 | 4.54 | 4.76 | 3.34 | 4.52 | 4.28 | 4.67 |
| Loss* | 0.53 | 0.40 | 0.33 | 0.28 | 0.50 | 0.38 | 0.33 | 0.31 | 0.58 | 0.46 | 0.39 | 0.36 |
| True Positive % | 68% | 78% | 82% | 84% | 73% | 82% | 86% | 88% | 60% | 69% | 79% | 81% |
| False Positive % | 21% | 18% | 15% | 12% | 23% | 20% | 19% | 18% | 18% | 15% | 18% | 17% |
| **In-sample 1990-2007** | | | | | | | | | | | | |
| ps-R2 | 0.11 | | | | 0.14 | | | | 0.11 | | | |
| AUROC | 0.72 | 0.82 | 0.88 | 0.91 | 0.76 | 0.85 | 0.89 | 0.90 | 0.71 | 0.80 | 0.86 | 0.87 |
| NSR* | 0.47 | 0.29 | 0.21 | 0.15 | 0.41 | 0.25 | 0.11 | 0.18 | 0.55 | 0.43 | 0.34 | 0.31 |
| SNR* | 2.19 | 2.96 | 4.33 | 4.94 | 2.36 | 3.38 | 3.29 | 4.40 | 2.14 | 2.95 | 3.75 | 4.05 |
| Loss* | 0.63 | 0.48 | 0.36 | 0.30 | 0.59 | 0.43 | 0.36 | 0.34 | 0.68 | 0.56 | 0.47 | 0.44 |
| True Positive % | 68% | 79% | 83% | 88% | 71% | 81% | 92% | 86% | 60% | 67% | 73% | 75% |
| False Positive % | 31% | 26% | 19% | 18% | 30% | 24% | 28% | 19% | 28% | 23% | 19% | 18% |
| **Out-sample 2008-2011, All crises** | | | | | | | | | | | | |
| AUROC | 0.85 | 0.89 | 0.90 | 0.90 | 0.85 | 0.88 | 0.89 | 0.89 | 0.83 | 0.90 | 0.94 | 0.96 |
| NSR* | 0.31 | 0.22 | 0.14 | 0.12 | 0.25 | 0.20 | 0.17 | 0.17 | 0.39 | 0.25 | 0.17 | 0.12 |
| SNR* | 3.94 | 4.31 | 3.72 | 3.77 | 3.45 | 3.60 | 3.71 | 3.89 | 5.48 | 6.77 | 7.38 | 7.79 |
| Loss* | 0.44 | 0.37 | 0.35 | 0.33 | 0.42 | 0.39 | 0.36 | 0.35 | 0.46 | 0.34 | 0.27 | 0.22 |
| True Positive % | 75% | 82% | 89% | 91% | 81% | 85% | 87% | 87% | 66% | 78% | 85% | 89% |
| False Positive % | 19% | 19% | 24% | 24% | 23% | 23% | 23% | 22% | 12% | 11% | 11% | 11% |
| **Out-sample 2008-2011, New crises** | | | | | | | | | | | | |
| AUROC | 0.89 | 0.89 | 0.89 | 0.89 | 0.86 | 0.86 | 0.86 | 0.86 | 0.94 | 0.96 | 0.96 | 0.96 |
| NSR* | 0.13 | 0.13 | 0.13 | 0.13 | 0.28 | 0.28 | 0.28 | 0.28 | 0.17 | 0.11 | 0.11 | 0.11 |
| SNR* | 3.74 | 3.74 | 3.74 | 3.74 | 6.86 | 6.86 | 6.86 | 6.86 | 7.07 | 7.84 | 7.84 | 7.84 |
| Loss* | 0.34 | 0.34 | 0.34 | 0.34 | 0.36 | 0.36 | 0.36 | 0.36 | 0.27 | 0.21 | 0.21 | 0.21 |
| True Positive % | 0.90 | 0.90 | 0.90 | 0.90 | 0.75 | 0.75 | 0.75 | 0.75 | 0.85 | 0.90 | 0.90 | 0.90 |
| False Positive % | 24% | 24% | 24% | 24% | 11% | 11% | 11% | 11% | 12% | 11% | 11% | 11% |

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

After comparing the 32 models we find that the following model (Model 1) is the model that displays the best overall performance ("in" and "out" of sample) in terms of NSR, Loss value and AUROC. However, this model includes the Credit Gap with only a one year lag and the rest of the variables with a two year lag.It has the following specification:

$$Prob_{it} = f(CreditGap_{i,t-1}\,, Libor_{t-2}\,, US\ GDP\ growth\ rate_{t-2}, Liquidity_{i,t-2}, (\frac{CA}{GDP})_{i,t-2})$$

The second best overall model (Model 2) 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:

$$Prob_{it} = f(CreditGap_{i,t-2}\,, Libor_{t-2}\,, US\ GDP\ growth\ rate_{t-2}, Liquidity_{i,t-2}, (\frac{CA}{GDP})_{i,t-2})$$

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:

$$Prob_{it} = f(CreditGap_{i,t-1}\,, Libor_{t-1}\,)$$

In Table 12 we present the regression results of the three final model specifications where we can observe the estimated coefficients for each explanatory variable and their statistical significance in each specification.

Table 12

**Regression Results of the three best final models. Period 1991-2011**

| **Banking Crisis (=1 if in year t there is a banking crisis, =0 otherwise)** | **Countries=68 Obs=1255 Pseudo-R2 = 0.20 Prob > chi2 = 0.0000** | **Countries=68 Obs=1191 Pseudo-R2 = 0.23 Prob > chi2 = 0.0000** | **Countries=68 Obs=1258 Pseudo-R2 = 0.18 Prob > chi2 = 0.0000** |
|---|---|---|---|
| | **Model 1** | **Model 2** | **Model 3** |
| Credit Gap (t-1) | 0.087*** | | 0.097*** |
| | 10.430 | | 12.060 |
| Credit Gap (t-2) | | 0.102*** | |
| | | 10.43 | |
| Libor (t-1) | | | 0.098** |
| | | | 2.300 |
| Libor (t-2) | 0.255*** | 0.377*** | |
| | 4.540 | 5.03 | |
| GDP Growth US (t-2) | -0.199*** | -0.26*** | |
| | -3.600 | -3.970 | |
| Current Account (t-2) | -0.027* | -0.029 | |
| | -1.700 | -1.630 | |
| Credit-to-Deposits (t-2) | 0.005** | 0.007** | |
| | 2.260 | 2.590 | |
| Constant | -3.22*** | -3.80*** | -2.80*** |
| | -10.910 | -10.820 | -11.050 |

***,**,* indicate significance at 1%, 5% and 10% respectively. Z-statistics are shown below the coefficients.
Source: BBVA Research

# 7 Potential uses: risk analysis tools

In this section we show the alternative uses derived from the final selected model. With these tools we can easily analyze the current systemic risk of more than 83 countries from both emerging and developed economies. These tools allow us to observe and analyze the estimated probabilities and the performance of the model delivering early warning signals.

The final model's chosen specification can also be expressed in the following way:

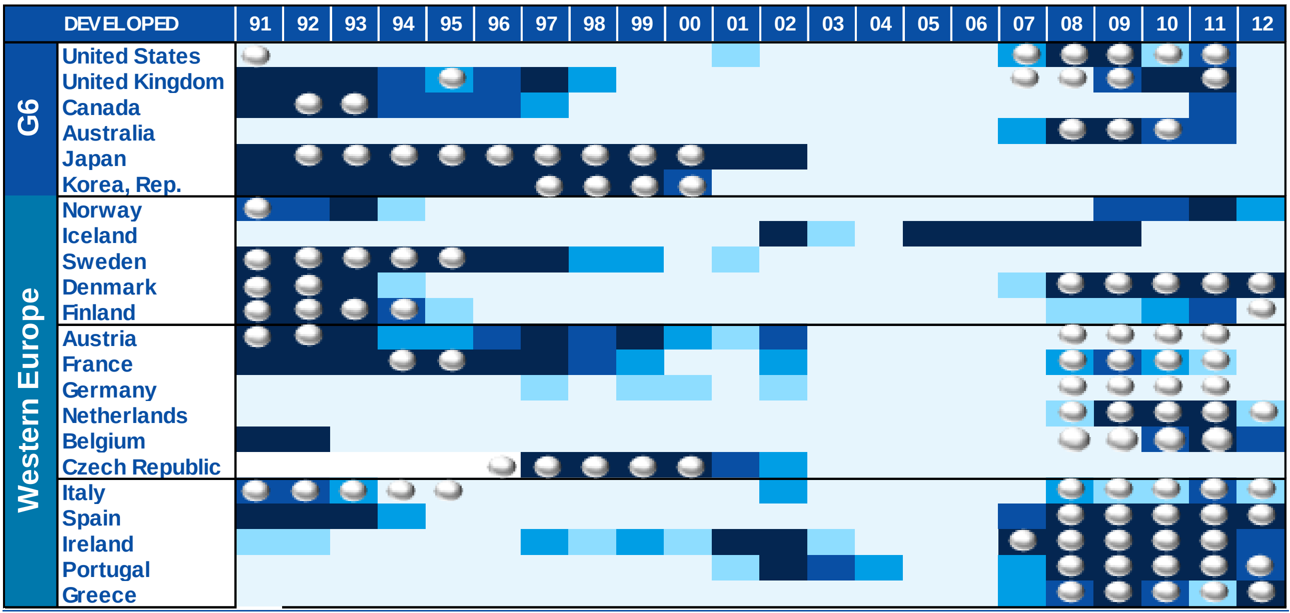
$$Prob.Crisis_{i,t+2} = f(CreditGap_{i,t}\ ,Libor_{i,t}\ ,US\ GDP\ growth\ rate_{i,t}, Liquidity_{i,t}, (\frac{CA}{GDP})_{i,t})$$

## Estimated Probability vs. Actual Crises

In Chart 5 and Chart 6 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 5 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 crises (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.

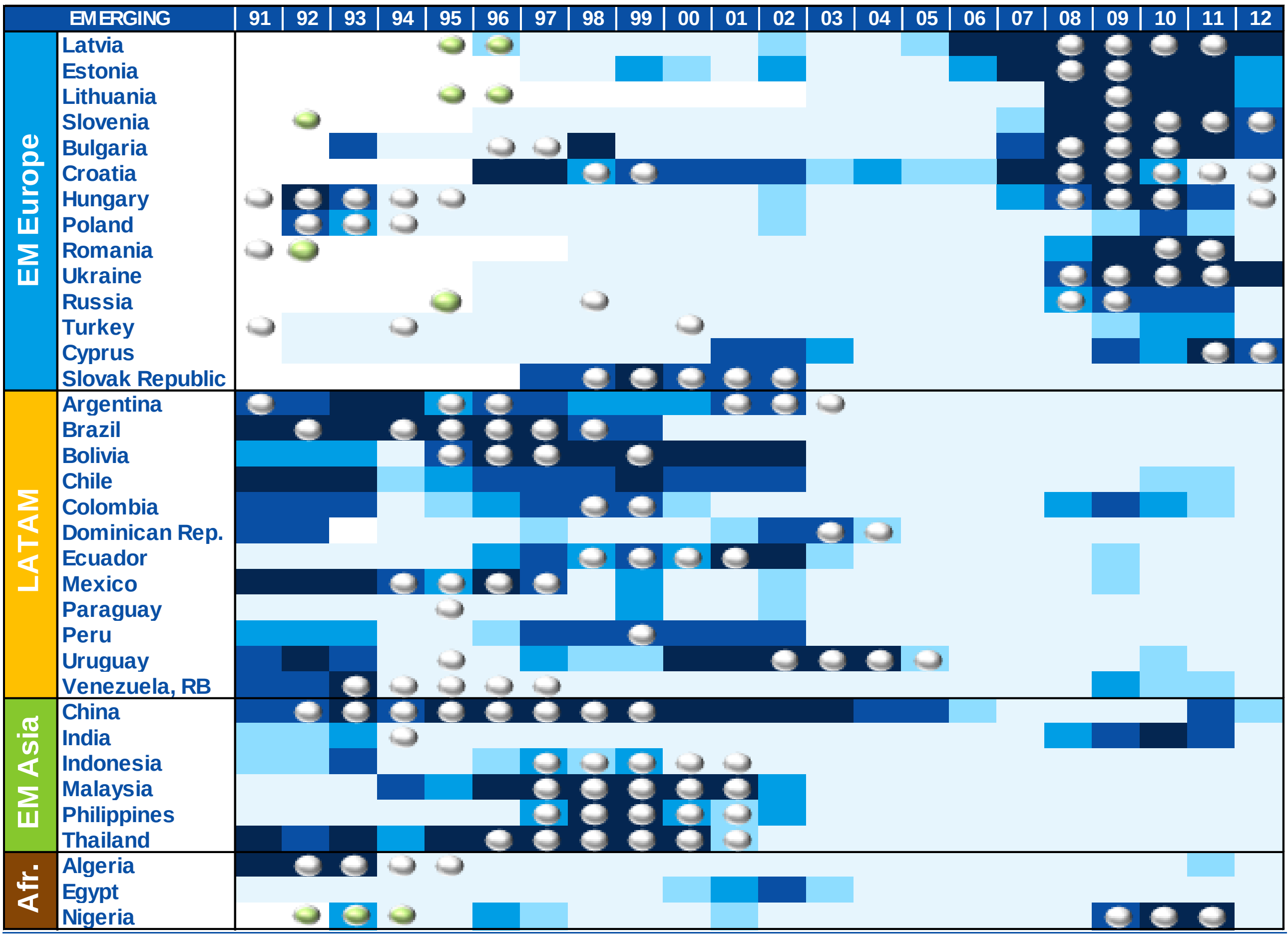
Figure 5
**Probability of Banking Crisis (2-years ahead) vs. Actual Crises. Developed Economies**




Source: BBVA Research

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, as it can be observed in Chart 6. In the case of the Asian crisis of 1997 the model correctly anticipates the financial crises 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.

Figure 6
**Probability of a crisis (two years ahead) vs actual crises. Emerging Economies**




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
Source: BBVA Research

The model warns also about the existence of imbalances 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 leveraging and de-leveraging processes.

## Dynamic Thresholds

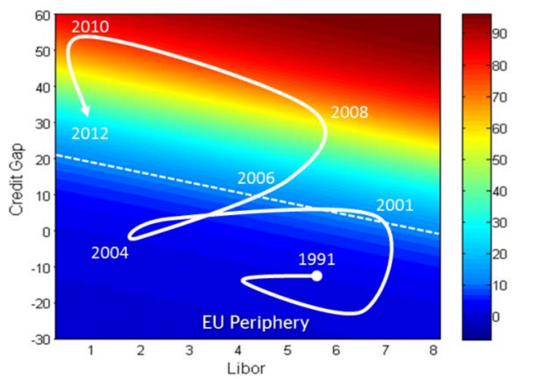
The estimated model allows us to compute and analyze early warning thresholds of our structural credit gap, i.e. the levels of the estimated probability that would trigger a warning signal. We can also analyze how these thresholds change depending on the values of the different explanatory variables.

In Chart 7 to Chart 10 we observe the regions of probability that are jointly determined by the Credit Gap (vertical axis) together with each one of the other explanatory variables (horizontal axis). In every case, the region of probabilities depends on the value of the two chosen variables, keeping the other variables constant at their median value.

Inside each region we can observe a risk threshold (a dash line) that represents the optimal estimated probability of a crisis that should trigger a warning signal. This optimal threshold probability is estimated by maximizing the difference between the percentage of correct signals and the percentage of false signals, which is equivalent to minimizing the sum of the type I and type II errors.

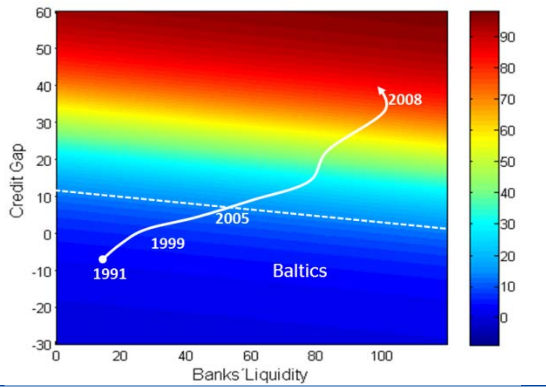
The pattern of colors oscillates between dark blue (meaning zero or low probability of crisis as can be observed in the right bar), other lighter blue colors around the risk threshold and the yellow and red areas (where the probability of crisis is clearly above the risk threshold).As observed in the graphics the optimal threshold is dynamic in the sense that changes with different values of the risk indicators[14]. The following examples show how the credit threshold can be affected by changes in the different variables:

Figure 7
**Banking Crisis Probability surface (Credit Gap and the Libor interest rate): EU Periphery**



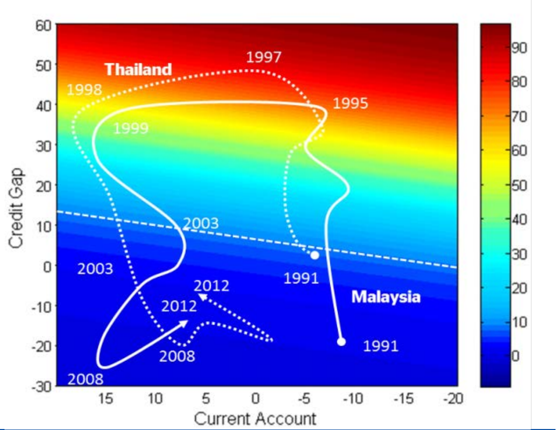

Source: BBVA Research

Figure 8
**Banking Crisis Probability surface (Credit Gap and Credit-to-Deposits ratio): The Baltics**



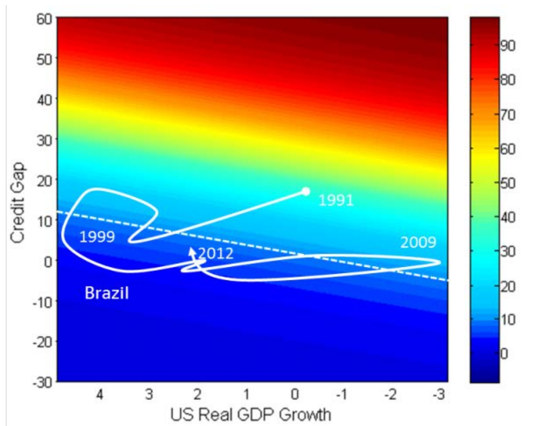

Source: BBVA Research

Figure 9
**Banking Crisis Probability surface (Credit Gap and Current Account Balance): South East Asia**




Source: BBVA Research

Figure 10
**Banking Crisis Probability surface (Credit Gap and US GDP growth): Brazil**




Source: BBVA Research

14: A similar approach is used by Lund-Jensen (2012).Monitoring systemic risk based on Dynamic Thresholds. IMF Working Paper 12/159.

Chart 7 shows the joint dynamic path of both the credit gap and the global interest rate (Libor) and the optimal threshold for the EU Periphery countries between 1991 and 2012. As it can be seen, when global interest rates decreased in 2003-04 the structural credit gap was nil (i.e. credit ratio 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. Thus vulnerability rapidly increased once the interest rates rose sharply Furthermore, 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 (smaller slope). The second graph (Chart 8) shows how Baltic countries moved rapidly right-upwards with both the structural credit gap and bank´s liquidity increasing quickly 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 joint evolution of the structural credit gap and the current account balance in the South East Asian countries (Malaysia and Thailand) is represented in Chart 9. 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 the credit excess experienced a sharp reversal which was followed by a long lasting de-leveraging process.

Finally, the Brazilian example (Chart 10) shows us how the US GDP growth can affect the early warning threshold and probability of crisis even when the credit gap remains constant.

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

# 8 Conclusions

In this paper we have developed and a new Early Warning System (EWS) of systemic banking crises for a large group of 83 developed and emerging economies. In the model the most important explanatory variable is a measure of excess domestic private credit. Rather than using simple Credit-to-GDP growth or excess Credit measures resulting from ad hoc filter or trends we propose a new leading indicator of systemic crises, based on a panel-data methodology developed in a previous study.

We have tested the properties, advantages and forecasting accuracy of our credit gap measure according to several in-sample and out-of-sample statistics, through both univariate and multivariate (BMA) analyses. There are several advantages according to the results of the analysis.

First, our measure allows us to analyze the underlying factors behind the structural ratios and gaps as our measure is the result of an estimated panel data model. Second, and contrary to the alternative measures, our credit gap measure seems robust to the so-called “end-of-sample” problem. Third, the credit gap displays a better forecasting performance, especially when comparing out-of-sample performance. Besides, it presents good timing properties (its outperformance increases with the number of lags) and stability (it maintains its characteristics in different samples). Fourth, our credit gap measure not only have good warning signals but also predicts well its duration and intensity, contrary to other EWI which lose their characteristics soon after the start of a crisis. Furthermore, and different from other EWIs, it might be further improved.

Finally we develop an Early Warning System for banking crises. For this, we have combined our new credit gap with other variables that are usually regarded as indicators of financial imbalances, such as the current account balance and the credit-to-deposits ratio. Importantly, our analysis show the importance of global financial variables usually considered as triggers of a systemic crisis, such as the Libor interest rate and the growth rate of GDP of US.

With the final model we provide different tools that could be used to analyze the current levels of systemic risk in a wide variety of countries. The model allows us to estimate dynamic risk thresholds of our credit gap in combination with other variables by calculating probabilities of crisis. Risk thresholds can change significantly with the level of global interest rates (as it was the case in the EU periphery) so countries should enhance macro-prudential policies in order to anticipate the future increases in interest rates in western countries. Similarly, countries with high Credit-to-Deposit ratios and/or Current account deficits should make extra efforts in monitoring the performance of the private sector credit.

# 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.

Babecký, J, Havránek,T, Matějů,J Rusnák,M, Šmídková K and Vašíček, B (2012). Banking, Debt and Currency Crises. Early Warning Indicators for Developed Economies. ECB WP 1485

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 2002

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, (1998), "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.

Demirgüç-Kunt, A., and E. Detragiache (2005), "Cross-Country Empirical Studies of Systemic Bank Distress: A Survey", National Institute Economic Review, 2005

Drehmann, M., Borio, C., Gambacorta, G.J., Trucharte, C. (2010), "Countercyclical capital buffers: exploring options", BIS Working Paper No 317

Drehman, M. and Juselius, M. (2013), "Evaluating early warning indicators of banking crises: Satisfying policy requirements", BIS Working Paper No 421, August 2013, Monetary and Economic Department.

Drehman, M. and Tsatsaronis, K. (2014), "The credit-to-GDP gap and countercyclical capital buffers: questions and answers", BIS Quarterly Review, March 2014.

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.

Eichengreen, B. and Rose, AK. (1998), "Staying afloat when the wind shifts: External factors and emerging-market banking crises", NBER Working Paper Series 1998.

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.

Giannone, D., M. Lenza, and L. Reichlin (2010): "Money, Credit, Monetary Policy and the Business Cycle in the Euro Area," mimeo.

Hardy, D. and Pazarbasioglu, C. (1998), "Leading Indicators of Banking Crises: Was Asia Different", IMF Working Paper WP/98/91

Kaminsky G., Lizondo, S., and Reinhart, C. (1998), "Leading Indicators of Currency Crises", IMF Staff Papers Vol. 45, No. 1 March 1998.

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.

Saurina, J and Repullo, R (2011). •The Countercyclical Capital Buffer of Basel III: A critical Assessment

Zeugner, S. (2011), “Bayesian Model Averaging with BMS” Working Paper 2011.

# Appendix A1

Table A.1

**Final performance score based on all statistics across all 15-methodological variations. Regressions including variables with a 1 year lag**

| | Credit Gap | HP- Gap | Credit/GDP change | Credit/GDP change>5 | LT-Gap |
|---|---|---|---|---|---|
| TOTAL | 190 | 172 | 37 | 81 | 162 |
| In sample | 60 | 60 | 1 | 24 | 55 |
| In sample 2007 | 49 | 65 | 9 | 14 | 63 |
| Outsample (All) | 41 | 23 | 12 | 18 | 24 |
| Outsample (First) | 40 | 25 | 15 | 27 | 17 |

NSR=(%Type II)/(1-%Type I);. Loss*=(%Type II)+(%Type I). The NSR, Loss, True Positive and False Positive values are estimated at the optimal cut-off probability that minimizes the loss function.
Source: BBVA Research

Table A.2

**Final performance score based on all statistics across all 15-methodological variations. Regressions including variables with a 2 year lag**

| | Credit Gap | HP- Gap | Credit/GDP change | Credit/GDP change>5 | LT-Gap |
|---|---|---|---|---|---|
| TOTAL | 220 | 133 | 53 | 79 | 157 |
| In sample | 73 | 39 | 12 | 15 | 61 |
| In sample 2007 | 63 | 49 | 3 | 21 | 62 |
| Outsample (All) | 43 | 22 | 16 | 18 | 21 |
| Outsample (First) | 41 | 23 | 22 | 25 | 13 |

NSR=(%Type II)/(1-%Type I);. Loss*=(%Type II)+(%Type I). The NSR, Loss, True Positive and False Positive values are estimated at the optimal cut-off probability that minimizes the loss function.
Source: BBVA Research

Table A.3

**Final performance score based on all statistics across all 15-methodological variations. Regressions including variables with a 2 year lag**

| | Credit Gap | HP- Gap | Credit/GDP change | Credit/GDP change>5 | LT-Gap |
|---|---|---|---|---|---|
| TOTAL | 217 | 107 | 142 | 63 | 122 |
| In sample | 77 | 30 | 34 | 16 | 43 |
| In sample 2007 | 72 | 45 | 29 | 2 | 51 |
| Outsample (All) | 40 | 12 | 36 | 16 | 16 |
| Outsample (First) | 28 | 20 | 43 | 29 | 12 |

NSR=(%Type II)/(1-%Type I);. Loss*=(%Type II)+(%Type I). The NSR, Loss, True Positive and False Positive values are estimated at the optimal cut-off probability that minimizes the loss function.
Source: BBVA Research

Table A.4

**Final performance score based on AUROC across all 15-methodological variations. Regressions including variables with a 1 year lag**

| | Credit Gap | HP- Gap | Credit/GDP change | Credit/GDP change>5 | LT-Gap |
|---|---|---|---|---|---|
| TOTAL | 56 | 37 | 11 | 16 | 40 |
| In sample | 14 | 11 | 1 | 3 | 11 |
| In sample 2007 | 10 | 14 | 3 | 1 | 12 |
| Outsample (All) | 16 | 8 | 2 | 2 | 12 |
| Outsample (First) | 16 | 4 | 5 | 10 | 5 |

NSR=(%Type II)/(1-%Type I);. Loss*=(%Type II)+(%Type I). The NSR, Loss, True Positive and False Positive values are estimated at the optimal cut-off probability that minimizes the loss function.
Source: BBVA Research

Table A.5

**Final performance score based on AUROC across all 15-methodological variations. Regressions including variables with a 2 year lag**

| | Credit Gap | HP- Gap | Credit/GDP change | Credit/GDP change>5 | LT-Gap |
|---|---|---|---|---|---|
| TOTAL | 59 | 29 | 16 | 20 | 35 |
| In sample | 16 | 6 | 4 | 4 | 10 |
| In sample 2007 | 12 | 11 | 1 | 3 | 12 |
| Outsample (All) | 16 | 7 | 4 | 3 | 10 |
| Outsample (First) | 15 | 5 | 7 | 10 | 3 |

NSR=(%Type II)/(1-%Type I);. Loss*=(%Type II)+(%Type I). The NSR, Loss, True Positive and False Positive values are estimated at the optimal cut-off probability that minimizes the loss function.
Source: BBVA Research

Table A.6

**Final performance score based on AUROC across all 15-methodological variations. Regressions including variables with a 2 year lag**

| | Credit Gap | HP- Gap | Credit/GDP change | Credit/GDP change>5 | LT-Gap |
|---|---|---|---|---|---|
| TOTAL | 56 | 18 | 41 | 22 | 23 |
| In sample | 15 | 6 | 9 | 2 | 8 |
| In sample 2007 | 15 | 8 | 6 | 0 | 11 |
| Outsample (All) | 16 | 2 | 11 | 9 | 2 |
| Outsample (First) | 10 | 2 | 15 | 11 | 2 |

NSR=(%Type II)/(1-%Type I);. Loss*=(%Type II)+(%Type I). The NSR, Loss, True Positive and False Positive values are estimated at the optimal cut-off probability that minimizes the loss function.
Source: BBVA Research

# Working Paper

## 2015

15/02 **Álvaro Ortiz Vidal-Abarca and Alfonso Ugarte Ruiz:** Introducing a New Early Warning System Indicator (EWSI) of banking crises.

15/01 **Alfonso Ugarte Ruiz:** Understanding the dichotomy of financial development: credit deepening versus credit excess.

## 2014

14/32 **María Abascal, Tatiana Alonso, Santiago Fernández de Lis, Wojciech A. Golecki:** Una unión bancaria para Europa: haciendo de la necesidad virtud.

14/31 **Daniel Aromí, Marcos Dal Bianco:** Un análisis de los desequilibrios del tipo de cambio real argentino bajo cambios de régimen.

14/30 **Ángel de la Fuente and Rafael Doménech:** Educational Attainment in the OECD, 1960-2010. Updated series and a comparison with other sources.

14/29 **Gonzalo de Cadenas-Santiago, Alicia García-Herrero and Álvaro Ortiz Vidal-Abarca:** Monetary policy in the North and portfolio flows in the South.

14/28 **Alfonos Arellano, Noelia Cámara and David Tuesta:** The effect of self-confidence on financial literacy.

14/27 **Alfonos Arellano, Noelia Cámara y David Tuesta:** El efecto de la autoconfianza en el conocimiento financiero.

14/26 **Noelia Cámara and David Tuesta:** Measuring Financial Inclusion: A Multidimensional Index.

14/25 **Ángel de la Fuente:** La evolución de la financiación de las comunidades autónomas de régimen común, 2002-2012.

14/24 **Jesús Fernández-Villaverde, Pablo Guerrón-Quintana, Juan F. Rubio-Ramírez:** Estimating Dynamic Equilibrium Models with Stochastic Volatility.

14/23 **Jaime Zurita:** Análisis de la concentración y competencia en el sector bancario.

14/22 **Ángel de la Fuente:** La financiación de las comunidades autónomas de régimen común en 2012.

14/21 **Leonardo Villar, David Forero:** Escenarios de vulnerabilidad fiscal para la economía colombiana.

14/20 **David Tuesta:** La economía informal y las restricciones que impone sobre las cotizaciones al régimen de pensiones en América Latina.

14/19 **David Tuesta:** The informal economy and the constraints that it imposes on pension contributions in Latin America.

14-18 **María Abascal, Tatiana Alonso, Santiago Fernández de Lis, Wojciech A. Golecki:** A banking union for Europe: making virtue of necessity.

14-17 **Angel de la Fuente:** Las finanzas autonómicas en 2013 y entre 2003 y 2013.

14-16 **Alicia Garcia-Herrero, Sumedh Deorukhkar:** What explains India's surge in outward direct investment?

14-15 **Ximena Peña, Carmen Hoyo, David Tuesta:** Determinants of financial inclusion in Mexico based on the 2012 National Financial Inclusion Survey (ENIF).

14-14 **Ximena Peña, Carmen Hoyo, David Tuesta:** Determinantes de la inclusión financiera en México a partir de la ENIF 2012.

14-13 **Mónica Correa-López, Rafael Doménech:** Does anti-competitive service sector regulation harm exporters? Evidence from manufacturing firms in Spain.

14/12 **Jaime Zurita:** La reforma del sector bancario español hasta la recuperación de los flujos de crédito.

14/11 **Alicia García-Herrero, Enestor Dos Santos, Pablo Urbiola, Marcos Dal Bianco, Fernando Soto, Mauricio Hernandez, Arnulfo Rodríguez, Rosario Sánchez, Erikson Castro:** Competitiveness in the Latin American manufacturing sector: trends and determinants.

14/10 **Alicia García-Herrero, Enestor Dos Santos, Pablo Urbiola, Marcos Dal Bianco, Fernando Soto, Mauricio Hernandez, Arnulfo Rodríguez, Rosario Sánchez, Erikson Castro:** Competitividad del sector manufacturero en América Latina: un análisis de las tendencias y determinantes recientes.

14/09 **Noelia Cámara, Ximena Peña, David Tuesta:** Factors that Matter for Financial Inclusion: Evidence from Peru.

14/08 **Javier Alonso, Carmen Hoyo y David Tuesta:** A model for the pension system in Mexico: diagnosis and recommendations.

14/07 **Javier Alonso, Carmen Hoyo y David Tuesta:** Un modelo para el sistema de pensiones en México: diagnóstico y recomendaciones.

14/06 **Rodolfo Méndez-Marcano and José Pineda:** Fiscal Sustainability and Economic Growth in Bolivia.

14/05 **Rodolfo Méndez-Marcano:** Technology, Employment, and the Oil-Countries' Business Cycle.

14/04 **Santiago Fernández de Lis, María Claudia Llanes, Carlos López- Moctezuma, Juan Carlos Rojas and David Tuesta:** Financial inclusion and the role of mobile banking in Colombia: developments and potential.

14/03 **Rafael Doménech:** Pensiones, bienestar y crecimiento económico.

14/02 **Angel de la Fuente y José E. Boscá:** Gasto educativo por regiones y niveles en 2010.

14/01 **Santiago Fernández de Lis, María Claudia Llanes, Carlos López- Moctezuma, Juan Carlos Rojas y David Tuesta.** Inclusión financiera y el papel de la banca móvil en Colombia: desarrollos y potencialidades.

## 2013

13/38 **Jonas E. Arias, Juan F. Rubio-Ramrez and Daniel F. Waggoner:** Inference Based on SVARs Identied with Sign and Zero Restrictions: Theory and Applications

13/37 **Carmen Hoyo Martínez, Ximena Peña Hidalgo and David Tuesta:** Demand factors that influence financial inclusion in Mexico: analysis of the barriers based on the ENIF survey.

13/36 **Carmen Hoyo Martínez, Ximena Peña Hidalgo y David Tuesta.** Factores de demanda que influyen en la Inclusión Financiera en México: Análisis de las barreras a partir de la ENIF.

13/35 **Carmen Hoyo and David Tuesta.** Financing retirement with real estate assets: an analysis of Mexico

13/34 **Carmen Hoyo y David Tuesta.** Financiando la jubilación con activos inmobiliarios: un análisis de caso para México.

13/33 **Santiago Fernández de Lis y Ana Rubio:** Tendencias a medio plazo en la banca española.

13/32 **Ángel de la Fuente:** La evolución de la financiación de las comunidades autónomas de régimen común, 2002-2011.

13/31 **Noelia Cámara, Ximena Peña, David Tuesta:** Determinantes de la inclusión financiera en Perú.

13/30 **Ángel de la Fuente:** La financiación de las comunidades autónomas de régimen común en 2011.

13/29 **Sara G. Castellanos and Jesús G. Garza-García:** Competition and Efficiency in the Mexican Banking Sector.

13/28 **Jorge Sicilia, Santiago Fernández de Lis and Ana Rubio:** Banking Union: integrating components and complementary measures.

13/27 **Ángel de la Fuente and Rafael Doménech:** Cross-country data on the quantity of schooling: a selective survey and some quality measures.

13/26 **Jorge Sicilia, Santiago Fernández de Lis y Ana Rubio:** Unión Bancaria: elementos integrantes y medidas complementarias.

13/25 **Javier Alonso, Santiago Fernández de Lis, Carlos López-Moctezuma, Rosario Sánchez and David Tuesta:** The potential of mobile banking in Peru as a mechanism for financial inclusion.

13/24 **Javier Alonso, Santiago Fernández de Lis, Carlos López-Moctezuma, Rosario Sánchez y David Tuesta:** Potencial de la banca móvil en Perú como mecanismo de inclusión financiera.

13/23 **Javier Alonso, Tatiana Alonso, Santiago Fernández de Lis, Cristina Rohde y David Tuesta:** Tendencias regulatorias financieras globales y retos para las Pensiones y Seguros.

13/22 **María Abascal, Tatiana Alonso, Sergio Mayordomo:** Fragmentation in European Financial Markets: Measures, Determinants, and Policy Solutions.

13/21 **Javier Alonso, Tatiana Alonso, Santiago Fernández de Lis, Cristina Rohde y David Tuesta:** Global Financial Regulatory Trends and Challenges for Insurance & Pensions.

13/20 **Javier Alonso, Santiago Fernández de Lis, Carmen Hoyo, Carlos López-Moctezuma and David Tuesta:** Mobile banking in Mexico as a mechanism for financial inclusion: recent developments and a closer look into the potential market.

13/19 **Javier Alonso, Santiago Fernández de Lis, Carmen Hoyo, Carlos López-Moctezuma y David Tuesta:** La banca móvil en México como mecanismo de inclusión financiera: desarrollos recientes y aproximación al mercado potencial.

13/18 **Alicia Garcia-Herrero and Le Xia:** China's RMB Bilateral Swap Agreements: What explains the choice of countries?

13/17 **Santiago Fernández de Lis, Saifeddine Chaibi, Jose Félix Izquierdo, Félix Lores, Ana Rubio and Jaime Zurita:** Some international trends in the regulation of mortgage markets: Implications for Spain.

13/16 **Ángel de la Fuente:** Las finanzas autonómicas en boom y en crisis (2003-12).

13/15 **Javier Alonso y David Tuesta, Diego Torres, Begoña Villamide:** Projections of dynamic generational tables and longevity risk in Chile.

13/14 **Maximo Camacho, Marcos Dal Bianco, Jaime Martínez-Martín:** Short-Run Forecasting of Argentine GDP Growth.

13/13 **Alicia Garcia Herrero and Fielding Chen:** Euro-area banks' cross-border lending in the wake of the sovereign crisis.

13/12 **Javier Alonso y David Tuesta, Diego Torres, Begoña Villamide:** Proyecciones de tablas generacionales dinámicas y riesgo de longevidad en Chile.

13/11 **Javier Alonso, María Lamuedra y David Tuesta:** Potentiality of reverse mortgages to supplement pension: the case of Chile.

13/10 **Ángel de la Fuente:** La evolución de la financiación de las comunidades autónomas de régimen común, 2002-2010.

13/09 **Javier Alonso, María Lamuedra y David Tuesta:** Potencialidad del desarrollo de hipotecas inversas: el caso de Chile.

13/08 **Santiago Fernández de Lis, Adriana Haring, Gloria Sorensen, David Tuesta, Alfonso Ugarte:** Banking penetration in Uruguay.

13/07 **Hugo Perea, David Tuesta and Alfonso Ugarte:** Credit and Savings in Peru.

13/06 **K.C. Fung, Alicia Garcia-Herrero, Mario Nigrinis Ospina:** Latin American Commodity Export Concentration: Is There a China Effect?.

13/05 **Matt Ferchen, Alicia Garcia-Herrero and Mario Nigrinis:** Evaluating Latin America's Commodity Dependence on China.

13/04 **Santiago Fernández de Lis, Adriana Haring, Gloria Sorensen, David Tuesta, Alfonso Ugarte:** Lineamientos para impulsar el proceso de profundización bancaria en Uruguay.

13/03 **Ángel de la Fuente:** El sistema de financiación regional: la liquidación de 2010 y algunas reflexiones sobre la reciente reforma.

13/02 **Ángel de la Fuente:** A mixed splicing procedure for economic time series.

13/01 **Hugo Perea, David Tuesta y Alfonso Ugarte:** Lineamientos para impulsar el Crédito y el Ahorro. Perú.

Click here to access the list of Working Papers published between 2009 and 2012

Click here to access the backlist of Working Papers:

Spanish and English



Contact Details:
**BBVA Research**
Paseo Castellana, 81 – 7º floor
28046 Madrid (Spain)
Tel.: +34 91 374 60 00 y +34 91 537 70 00
Fax: +34 91 374 30 25
bbvaresearch@bbva.com
www.bbvaresearch.com