UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI "M.FANNO"

CORSO DI LAUREA MAGISTRALE IN

BUSINESS ADMINISTRATION

TESI DI LAUREA

"EARLY WARNING SYSTEMS FOR SYSTEMIC BANKING CRISES: AN EMPIRICAL ANALYSIS"

RELATORE:

CH.MO PROF. GIOVANNI CAGGIANO

LAUREANDO: NICOLA COSTA

MATRICOLA N. 1080588

ANNO ACCADEMICO 2015 – 2016

Il candidato dichiara che il presente lavoro è originale e non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere.
Il candidato dichiara altresì che tutti i materiali utilizzati durante la preparazione dell'elaborato sono stati indicati nel testo e nella sezione "References" e che le eventuali citazioni testuali sono individuabili attraverso l'esplicito richiamo alla pubblicazione originale.

Abstract

This paper aims to contribute to that ever-expanding branch of the economic literature focused on the study of systemic banking crises and the development of statistical tools able to detect the signals of their occurrence in advance (Early Warning Systems). In particular, the three-step analysis, which constitutes the empirical core of the work, wants to provide robust evidences of the detrimental influence exerted by the *post-crisis bias* on the predictive efficacy of the binomial logistic model and whether this phenomenon magnifies along with the duration of the crisis episodes under scrutiny. Lastly, the average crisis duration, on a country basis, is linearly regressed on several economic, political and institutional metrics to further investigate which features usually distinguish those countries more prone to long-lasting defaults. Results show that models which do not account for any specific post-crisis solution systematically underperform the ones adopting it, across different sample compositions and crisis definitions. The *post-crisis bias* strikes harder as longer-lasting events are considered in the tests, suggesting that the bias-related distortion inflates for those countries which, at least historically, are more exposed to durable defaults. As emerged by the ultimate linear regressions, this peculiarity is associated to wealthier economies (in terms of GDP per capita), more distributed and open financial sectors.

Table of Contents

Chapter I:

Introduction

The effort exerted by the economic research community in studying the determinants of systemic banking crises and their employment to forecast these peculiar phenomena has a relative recent history. The first major event that triggered interest over this topic was undoubtedly the Great Depression of the late '20s, that burst in New York but spread by contagion through several countries all over the world. The huge real economic costs characterizing this episode, as well as others distinguished by a banking sector collapse, are difficult to be computed with precision but were enough deep to highlight the pressing need of understanding the nature of these phenomena and the signals that could detect their occurrence in advance. Since then, several papers have been produced in the attempt of identifying the causes and the best model able to describe and predict the oncoming of a systemic crisis, in other words an Early Warning System (EWS). At the light of the massive detrimental effect that a banking meltdown has on the level of a country's economic well-being, the contribution that a similar tool could give to the policy making process is unquestionable and priceless. However, despite the remarkable work produced so far, recently boosted by the latest waves of crises started in the early '80s, the existing literature has not been able to deliver a sufficiently extended set of standard indications regarding which framework and variables would consistently fit with most of the systemic episodes experienced so far by worldwide countries. This mismatch is due to either structural differences among the empirical exercises, regarding executive choices taken by the authors, or the distinctive and evolving nature of systemic banking defaults all over history. First of all, it still does not exist a universally agreed definition of "systemic banking crisis". This lack forced researchers to choose among few alternative solutions proposed by the creators of the main crisis dating databases which inevitably shaped the models and their results under a subjective perspective. Moreover, several different statistical models have been employed as EWS but, although just a handful of them have been used in the vast majority of the works, no one graduated as universally accepted standard due to their individual drawbacks. Other limitations, and so possible room for improvements, come from the selection of the indicators to be collected and included in the model to significantly predict the outburst of a crisis. During the last decades, tens of them have been evaluated, ranging from macroeconomic, through microeconomic and banking sector related ones, to variables reflecting the wealth of the real economy and the robustness of the institutional structure of a country. In spite of the effort, just few of them have demonstrated to be strongly relevant in most of the papers, while many are not even available for a large number of countries, often forcing researchers to test their models on a restricted number of observations. Scenario that makes it even harder to conduct a robust analysis on events such as system banking crises which are notoriously rare even over large timespans. Finally, discrepancies among the characteristics of the data sets, mainly due to the strongly different economic environment distinguishing some regional clusters of countries, or even country-specific ones, led to conflicting results and the increasingly solid feeling that a universally accepted EWS is still far from being achieved.

For what may concern the predictive models built so far, two frameworks have been leading the empirical scene on banking crises, especially during these last two fruitful decades. The first is the signal approach, a systematic statistical model, introduced by Kaminsky and Reinhart (1999) in their pioneering paper on the twin crises (i.e. the concomitant occurrence of banking and currency crises). The proposed method allowed for verifying the predictive capacity of some of the most discussed indicators of banking crises by individually testing them against a threshold value. Whether the variable had exceeded this safety level, then a warning signal would have been triggered. If the warning demonstrated to be able to predict the oncoming of a crisis within a horizon of 24 months, it would be considered a true signal. The approach accuracy have been progressively improved by setting the threshold value in order to minimize the so called noise-to-signal ratio that is the ratio between false signals and true ones. Nonetheless, since its first application, this counterintuitive approach presented some major drawbacks. It did not permit to measure the marginal predictive contribution of the single variable, as there was no room for distinction among different levels of abnormality in its behavior. Moreover, it did not count for the aggregated effect of the indicators, being limited at their stand-alone testing. Although the original authors (Kaminsky and Reinhart, 1999) attempted to partially cover these deficiencies by introducing composite indexes and multiple thresholds to account for different levels of deterioration in the variable value, the statistical solution implemented by Demirgüç-Kunt and Detragiache (1997) proved to overcome most of the signal approach weaknesses and outperform it as EWS. They considered the early warning indicators as explanatory variables integrated in a binomial logit regression model, where the dummy dependent variable took value ‘1’ in case of crisis period and ‘0’ otherwise. The behavior of the independent was shaped by a logistic function in order to restrict the possible outcomes within the unit interval. The parameters estimated along with the variables were interpreted as a measure of their contribution to the probability of experiencing a crisis, addressing the proper weight to each independent variable. This solution improved the model performances, both in-sample and out-of-sample, but at the same time introduced new dilemmas. For example, over the way in which the observed years following the burst of the crisis should be treated, as in those periods the values of the indicators are due to be deeply altered because of the stressing status of the banking system and the overall economy. Two possible treatments consist in either considering the periods in the aftermath of the financial collapse as simple tranquil periods, thus denying the existence of this issue, or eliminating them from the sample. Both solutions bear heavy costs, implying a reduced estimation efficiency and the loss of possible relevant information. With the purpose of avoiding the shortcomings due to be provoked by the so called *post-crisis bias* or *crisis duration bias*, a new econometric framework has entered the scene thanks to the effort of Bussiere and Fratzscher (2006) in their currency crises study. The model at issue is the multinomial logit approach that not only is an effective tool in avoiding the *post-crisis bias* but, in its first tests on financial crises, seems to outperform its predecessor in terms of predictive power (lower Type I and Type II errors). However, its application in the banking crises field is almost newly born and, despite its recent positive performances, its employment is still really limited. Nonetheless, even the author of this paper strongly believes that this path is the right one towards the development of a successful EWS for banking sector crises.

Accordingly with what exposed above, with this work the author wants to provide further and stronger evidence on the existence of the *post-crisis bias* and the effectiveness of the logit approach, either binomial or multinomial, in handling it. Together with this task, I will investigate the link (if any) that binds the crisis duration with the quality of the EWS performance, in order to possibly draft valuable policy advises on how to fruitfully pool sample of countries on the basis of their inclination toward long-lasting crisis experiences. To my knowledge, a similar challenge has never been taken before. These goals are deemed to be achievable through various empirical steps. First of all, it is needed to justify the suspects raised over this post-crisis specific drawback. This can be reached through the implementation of a horse-race comparison between the performances of binomial models that either adopt a well-known post-crisis treatment (or implementing a multinomial model) or do not account at all for the *crisis duration bias*. During this phase there will be also plenty of room for collecting and interpreting the information returned by the models on the correlations that bind candidate determinants and the probability of an oncoming crisis. Nonetheless, any detected relationship should be treated with caution as it may not necessarily reflect a direct causal link. Thereafter, by benchmarking the forecasting accuracy of binomial models on groups of countries discriminated on a crisis duration basis, I will target the duration-related goal. Conducting this latest exercise will possibly allow to shed a brighter light over the impact that the presence of long lasting crises has on the predictive ability of the model (in terms of AUC area) and the overall capacity of the proposed solution to tackle any emerging deficiency. In a second moment, the author will attempt to investigate which features of a country economic, political and institutional landscape are more correlated to the occurrence of long lasting crisis events. For this scope, a variable containing the average crisis duration over the explored time period and on a country basis would be regressed on a disparate and sizable list of variables in a linear and cross-sectional analysis. This experimental section should underline, if any, the characteristics of those economies that, on average, are more susceptible to durable banking crises and therefore, in implementing a specific EWS, should warmly consider to adopt a *post-crisis bias* appropriate solution. The outcomes of these tests are expected to provide valuable indications either on which statistical model would better fit the specific issues carried by the banking crisis prediction task and over how to set-up a solid early warning framework across several country-wide specifications.

Empirical results from the first logistic session suggest that lagged increments in inflation, a higher vulnerability to sudden capital outflows, a soaring credit from financial to private sector and a greater level of illiquidity in the banking system are all signals connected with an increase in the likelihood of experiencing a systemic crisis. Negative fluctuations in the nominal official exchange rate and the net open position ratio, as proxy for banking system FX exposure, are expected to exert the same effect on the stability of a country's financial sector, although these findings have not proved uniformly strong (see robustness tests). Beside considerations related to the individual determinants of a crisis, the binomial logit approach has been demonstrated valid in fulfilling the EWS duties, exhibiting satisfying post-estimation metrics and generating AUC areas always greater than 0.70. The multinomial approach, as well, performing slightly better than its binomial version, has confirmed its authority in the role of EWS and provided valuable information over the forces responsible to hold a country within the crisis state. In this matter, an enduring economic recession, growing terms of trade, an increasing susceptibility to capital outflows and the degree of illiquidity of the banking sector proved all worthy signals of a prolonged permanence under stressed conditions. Despite the predictive differences between these frameworks, either binomial or multinomial, results firmly confirmed the quality deterioration caused by the *post-crisis bias*. During the second empirical phase, the outcomes from several binomial logit tests have been plotted highlighting the positive relationship existing between an average crisis duration variable and the disruptive effect brought about by the *crisis duration bias*. The same duration variable, once employed as regressand in a linear regression analysis, has shown to be solidly and positively related to a country's income per capita and its level of institutional quality. To test the robustness of these findings, the whole empirical batch has been repeated employing a different crisis dating procedure (Laeven and Valencia, 2012). This double-checking process is useful not only to stress results under a different perspective but also to diversify the risk incurred by relying on a single crisis definition, undoubtedly affected by certain degree of subjectivity. Thanks to this additional analysis, main evidences on the *post-crisis bias* existence and the duration-quality relationship have been further corroborated and strengthened.

The paper proceeds as follows. Chapter 2 will review some of the most important findings achieved so far in matter of determinants of banking crises and their most supported theoretical interpretations, in the attempt of providing the reader with an adequately accurate portrait of the dynamics that tend to breed banking sector fragility. A brief recap on which features are deemed to characterize countries' tendency to prolonged crisis episodes is included too. In the closing part of this chapter I will also examine in detail the statistical models that gathered the strongest consensus in the role of EWS for systemic banking crises and among them the logistic ones that will play a leading role in my analysis. The specifics of the data sample, the dependent and the independent variables that will take part in the logistic exercises are all topics treated in Chapter 3, while the execution and results of the empirical tests are illustrated and described in Chapter 4. Robustness tests and sensitivity analysis, and the concluding remarks are exposed respectively in Chapter 5 and Chapter 6.

Chapter II:

Literature Review

2.1 An historical overview on systemic banking crises and their determinants

The latest systemic shock experienced by financial markets and national banking sectors, culminated in the Great Recession which officially started in 2007, has been the ultimate proof of the deteriorating power that a fragile banking system has on the real economy of a country once a crisis burst. Although the magnitude of this crash has been considered lower just compared to the Great Depression of the late '20s, it represents only the tip of the iceberg in the history of systemic banking failures. In fact, to find the very first documented episode in this matter we must go far back to the 33 A.D. in the Roman Empire, when Tiberius Caesar had to face a widespread closure of banking houses generated by a mix of factors (Calomiris 1989). Some of them are undoubtedly a thing of the past, as the sinking of some ships carrying uninsured cargo or the slave revolt, while others, such as liquidity draining by government-sponsored projects and international contagion, can be found even in some recent defaults. Since then the banking sector went through a radical evolution, however then as today it is still closely intertwined with elements of the real economy, the institutional and political frameworks. Its central position within the economic landscape of a country is the natural consequence of the intermediary role played by banks in the credit business. This feature made the banking system vulnerable from shocks of various origins, amplifying the scope of the study for those researchers that wanted to shed a light on the causes of a systemic banking default.

As anticipated in the introductory chapter, the interest over this topic has reached a remarkable level just after the events following the Wall Street stock exchange crash on October 1929 an its multiple banking collapses. The economic turmoil generated was tremendous, eroding GDP, employment level, international trades and production in all sectors (although primary sector industries, as agricultural and mining, suffered the most). From 1929 to 1933, banks all over the United States experienced sequential runs to deposits that led to the erosion of the liquidity in the system and the consequent bankruptcy of more than 15000 banks (Caprio and Klingebiel, 1997). Eventually, the banking downturn demonstrated to be more painful than the stock collapse, as the fiscal cost to bailout the banking sector and the losses due to the economic slowdown were huge. Thereafter, one major debate that developed during the 20th century was focused on the impact of the macroeconomic environment on the likelihood of a systemic banking crisis. For several decades after the Great Depression one major belief was that diffused runs on deposits were the results of depositors' self-fulfilling expectations, as firmly sustained by Diamond and Dybvig (1983), and the asymmetric information between banks and depositors over the quality of assets owned by financial institutions. This explanation was supported by the classic view under which individual bank runs could become systemic just in the presence of three elements: opaque information over bank assets that could lead to runs over solvent firms, sequential servicing which allows depositors to withdraw their funds until the bank closes and the lack of a credible lender of last resort. Under these assumptions, widespread bank runs were detonated by information shocks that, eventually, could have threaten the stability of solvent but illiquid banks. If this was the case, systemic defaults were principally a matter of unwarranted "panic" and "contagion of fear" (Friedman and Schwartz, 1963; Kindleberger, 1978). Although this theory has proved able to partially describe the contagion at the base of the Great Depression, to most of the researchers in the field this explanation seems over-simplistic and scarcely applicable to the new wave of banking distress started in the early '80s. These years will also mark the beginning of a fertile period for the literature, as systemic crises proliferated until the latest unfamous events in 2007. Between the end of World War II and the early '70s, the banking sector experienced a prolonged and steady tranquil period with almost no systemic crisis recorded. This peaceful timespan was principally due to a bunch of factors: a relative stable macroeconomic environment, low inflation, a spread economic growth, lax monetary policy, the introduction of the Bretton Woods system and a severe regulatory framework on banks' balance sheets to prevent them from taking excessive risks. As these elements fell apart, starting from the dissolution of the peg system in 1973 to an increasing macroeconomic volatility (partially fostered by the oil shock) and a loosening of the capital and overall regulatory requirements to the banking sector, a new wave of systemic defaults spread in both emerging and advanced economies.

2.1.1 Macroeconomic forces

One of the first in empirically supporting the causal connection between some macroeconomic key indicators and banking sector fragility was Gorton (1988), whose findings have been later on strengthen by several major authors in the field.

His study on the US National Banking Era supported the business cycle view according to which diffused runs were the results of shocks in some economics fundamentals. Thus, rather than being detonated by panic, a systemic meltdown would be the consequence of an underlying economic recession. Under this "business cycle" theoretical framework, crises normally follow periods of economic boom characterized by constantly growing GDP and large availability of liquidity in the market. Along with the wealth of the economy, stock market prices and speculative investments on specific businesses soar, facilitated by diffuse over-optimism among investors. Apart from an expansionary monetary policy, this speculative bubble is financed through a credit boom that in its turn is allowed by lax lending standards, a permissive financial regulation and a weak supervision. Once the speculative bubble reaches its peak and bursts, falling asset values and increasing share of non-performing loans, due to borrowers' inability in paying back their obligations, would deeply deteriorate banks' balance sheets that would found themselves stuck between the illiquidity of their assets and the liquidity of their liabilities (especially for current accounts). Once again, depositors would play a crucial role, as by withdrawing their funds they would relentlessly drain bank resources, leading the institutions towards illiquidity and insolvency. Several of these aspects can be also found in the latest US real estate bubble and similarly in the early '90s in Japan. Although the analysis of Gorton (1988) was based on a country-specific sample concentrated on a restricted timespan, it had the merits of demonstrating that banking crises were not just a consequence of extraneous random variables but had their roots in the economic environment. As a consequence, they were, at least partially, predictable events. From this milestone on, several macroeconomic indicators have been considered in most of those studies dealing with the determinants of systemic banking crises. Few of them have demonstrated to be strongly significant in predicting distress periods. In the attempt of explaining the forces able to foster the oncoming of a crisis, Demirgüç-Kunt and Detragiache (1997) found to be relevant and positive correlated to the probability of experiencing a banking crash the following fluctuations: a slowdown or decrease in GDP growth lagged by one period, high inflation, high real interest rates, a sudden spike in capital outflow and an increasing M2-to-Reserves ratio (suggesting that bank exposure to currency crises plays a role in setting the stage for a banking default). This paper represents a masterpiece both because of the innovative econometric approach employed (multivariate binomial logit) and its significant findings on some main macroeconomic variables. The robustness of its results has been repeatedly confirmed by numerous authors that tested the indicators with different models and data sets. Among them, Hardy and Pazarbaşioĝlu (1999), implementing a multinomial logit approach, proved the predictive capacity of some signals, as a sharp decline in real exchange rates, a collapse in capital inflows and a decrease in the terms of trade. A lagged reduction in the terms of trade, the ratio between exports and imports, is registered even by Davis and Karim (2008) which confirmed the variable, along with the growth rate of GDP, to be a valuable early warning indicator for full-fledged crisis episodes. Favorable terms of trade movements are due to proxy lower exchange rate-based market risk and lower chances of currency crisis. Kaminsky and Reinhart (1999), in one of the most-cited literature pillars, suggest that good proxies for banking system vulnerability are a downward movement of the real exchange rate, a decreasing short-term capital inflows on GDP, a collapse of the stock market, as well as confirming the role played by GDP growth and interest rates. The influencing capacity of stock market movements on the stability of a national banking sector has been gradually validated by the works of several researchers (Caprio and Klingebiel, 1997; Borio and Lowe, 2002).

However, in spite most of the variables just listed are all backed by a robust theoretical explanation, only few of them, such as the annual growth rate of GDP, inflation and interest rates, gathered consensus in the vast majority of the papers. As an example, in contrast with what discovered by Kaminsky et al.(1999), Demirgüç-Kunt et al. (1997) proved the real exchange rate variable to be insignificant and vice versa Kaminsky et al.(1999) found irrelevant the contributions of the M2-to-Reserves ratio and the terms of trade. As anticipated, this dissonance among the results is mainly due because of core differences both in the data samples employed and in the very nature of the banking crisis episodes all over the world. First Latin American turbulences in the early '80s were mainly prompted by external factors and exchange rate policies while the saving and loans debacle in US (S&L) was much related to banking sector deficiencies, a weak regulation and supervision, financial liberalization and generous deposit insurance schemes (Demirgüç-Kunt and Detragiache, 2005). As a consequence of this variety, researchers have not circumscribed their sphere of action to the only macroeconomic fundamentals. In fact, despite a healthy economic environment is proved of being able to lower the vulnerability of a banking sector to a systemic crash, it could, to some degree, have the reverse effect by eroding incentives for prudent banking (Caprio and Klingebiel, 1997). Macroeconomic developments, in Gavin and Hausmann (1996) "chain" analogy, represent just one of the forces that exerts tension on the chain, the banking system. Thus, economy-wide factors would not tell us anything about which is the weakest link and which are the flaws of the chain. From this statement, it can be easily deduced that even credit industry specific features could play some role in determining the vulnerability of the financial system.

2.1.2 Microeconomic factors

The relevance of the peculiar characteristics of the banking sector and its degree of connection with certain segments of the economy has been receiving attention from numerous authors during this last two decades. A one year lagged credit growth has been largely recognized as a strongly significant signal of oncoming banking issues. Demirgüç-Kunt and Detragiache (1997) emphasized credit growth role, especially if addressed to the private sector. The role of a credit boom in setting the stage for a systemic collapse has been recently confirmed by Navajas and Thegeya (2013) that tested the predictive power of some FSI (Financial Soundness Indicators). Among the most recent papers, the one developed by Boissay, Collard and Smets (2013) is noteworthy as it does not only confirms the widely approved positive correlation between an expanding credit availability and the likelihood of a systemic banking crisis but also, by entirely focusing on the credit boom phenomenon, further improves our knowledge on its relationship with a crisis. Larger credit growth rates seem to be associated to: higher crisis probability, smaller time-lag till the burst of the crisis and, in its aftermath, a deeper and longer recession. During my regression analysis on the average crisis duration there will be room to marginally verify this latest aspect. The credit boom related threat would eventually be amplified whether lending, as much as risk, was concentrated in a particular sector ("common risk factor"; Borio and Lowe, 2002). Lainà, Nyholm and Sarlin (2014) statistically proved it for the real estate market on a sample of European countries. As well as for advanced economies, lack of risk diversification is a major source of banking vulnerability even in emerging market countries where economy and so investments are normally focused on a limited bunch of businesses. Nonetheless, under these conditions, even the magnitude of the depression brought about by the crisis should exponentially inflate. Rojas-Suarez (1998), in their Latin America analysis, adopted an approach similar to the CAMEL framework used for the identification of individual distressed banks. They found to be good early warning signals: the loan-to-deposit interest rates spread, the deposit interest rate and interbank debt growth. Loan-to-deposit spread is just one of the proxies for financial liberalization that consolidates the theoretical belief that wants crises preceded by a deep deregulation of the financial sector and an increased competition among bankers. Findings in this regard are also those computed by Honohan (1997) that confirmed the validity of high loan-to-deposit spread and a high foreign borrowing-to-deposits ratio as good indicators for future banking instability. Even the currency mismatch between assets and liabilities that may emerge when banks borrow or lend abroad could have a potentially disruptive effect on the banking system profitability. As sustained by Demirgüç-Kunt and Detragiache (1997), if banks have a foreign exchange open position they would be much more susceptible to domestic currency fluctuations. By lending at home in foreign currency they could rebalance their exposure towards foreign exchange volatility but it would mean loading with the currency associated risks the domestic loans, thus increasing the share of non-performing assets in case of exchange rate shock. Caggiano et al. (2014) empirically demonstrated these allegations, finding relevant and negative the relationship between the occurrence of a crisis and the net open position of the banking system, measured as the ratio between net foreign assets and GDP. Still looking at banking balance sheet figures, they even verified the role played by a liquidity ratio as early warning indicator and its positive correlation with the likelihood of a full-fledged default. This outcome is largely supported by theory too. Proxied by a soaring credit-to-deposit ratio, a drop in liquidity of balance sheet items is expected to let a bank more vulnerable to sudden massive deposit withdrawals (a run on deposits). Another study that adopted single bank balance sheet figures to the systemic dimension was the one carried on by González-Hermosillo (1999). They found out a relevant deterioration in performing loan quotas and capital asset ratios right before the burst of a crisis, fostering the theory that wants a crisis preceded by a cyclical downturn. A decrease in capital requirements is also expected to increase the exposure of banks to depositor's run. To validate this intuition, Čihák and Schaeck (2007) tested and found significant a lagged decrease in the mandatory capital to risk-weighted assets, discovery that later on would be strengthen by Lainà, Nyholm and Sarlin (2014). In their pioneeristic study, Čihák and Schaeck used an aggregated version of bank ratios (FSI), which are usually taken into consideration during individual bank stress tests, as indicators of the overall system vulnerability toward shocks. A declining return on equity ratio had the strongest predictive power among the tested indexes (result confirmed by Navajas and Thegeya, 2013).

Other banking sector elements that lie outside of pure technical ratios regard banking ownership, concentration and the presence of foreign banks within a national financial system. Caprio and Martinez-Peria (2000) gave empirical support to the first one of them, finding that state-owned banking sectors are more prone to systemic collapses. The belief upon a possible negative effect of the presence of foreign banks has been matter for debates too. These were suspected to raise crisis probability through contagion, by withdrawing resources from the host country's branches to face problematic conditions at home, and their presumed short-term commitment in the local economic development. Worries that were empirically denied by Demirgüç-Kunt, Levine and Min (1998) that instead suggested a negative correlation between foreign banks presence and the risk of crisis. Ultimately, Beck, Demirgüç-Kunt, and Levine (2003) showed that banking crises are less likely where the sector is more concentrated and the regulation is favorable to competition. The role of banking structure is one of the most strongly debated issue that still does not find a widely-agreed conclusion. Large banks are supposed to bring some advantages to the system: more diversified risks, enhanced profits allowing for a safety buffer in case of adverse shock and an easier supervision process (Allen and Gale, 2000). On the other hand, the so called "concentration-fragility" view wants larger banks to be deeply affected by moral hazard as most of the times their capital is implicitly guaranteed by the government, in a "too big to fail" fashion. Furthermore, a concentrated banking sector could deteriorate information transparency and, by gathering market power on the hand of a very restricted number of institutions, it may induce borrowing firms to take higher risks to compensate for less competitive lending interest rates. This latest hypothesis has been also empirically tested and validated by Boyd and De Nicolò (2003). This heated debate will inspire my decision to include and test in my duration analysis quite a few proxies for banking concentration. In any case, the topic keeps on being controversial, leaving room for future investigations.

The role of competition, as well as concentration, has been a fertile field for controversy. Theoretically speaking, being a proxy for financial liberalization, a soaring competition would reflect a deregulation process being in place. For Kaminsky and Reinhart (1999), the same role would be fulfilled by an increasing real interest rate and money multiplier, reflecting a reduction in reserve requirements. Within the same investigation framework, it was found that almost 70% of the crises studied since 1970 was preceded by some kind of financial liberalization within the previous 5 years. A loose regulatory framework on this matter could consist in measures such as the removal of interest rates ceilings on deposits, easy access to the credit market for non-bank financial firms, lower restrictions on capital requirements and riskier activities that allow for a higher flexibility in the resource allocation process. Demirgüç-Kunt and Detragiache (1997) and Arteta and Eichengreen (2000) found that internal financial liberalization, as proxied by the removal of interest-rate controls, increases the risk of a banking crisis, respectively in advanced and in emerging economies. These regulatory provisions would eventually foster the level of rivalry within the banking sector, inducing banks to squeeze their loan-to-deposit margin and gamble on risky investments to preserve their profitability. Such a shaped regulatory framework joint with a faulty supervision would inevitably make bank managers more prone to deficient or even fraud practices. Once the distress reaches unsustainable levels and the risk of insolvency becomes concrete, the management could still decide to keep on bidding on junk assets and pursuing highly risky initiatives, magnifying bank exposures. This behavior could be the consequence of an extreme attempt to heal the bank balance sheet or to exploit the institution residual assets to realize a personal profit (looting). The whole risk taking process would be exacerbated by a widespread belief on the existence of an implicit governmental guarantee over bank capital. Caprio and Klingebiel (1997), in their comprehensive study on the causes and costs of systemic banking crashes, supported this view underlining poor management, weak supervision and fraud increases as major threats to the banking sector solidity, especially when associated to an ongoing liberalization. As this financial turmoil is due to follow a cyclical pattern, policy makers and bankers are expected to be able to refine their capacity to deal with these threats. However, market players, more often than not, during periods of strong economic growth risk to be affected by over optimism and disaster myopia. These states of mind are so frequent even because of the great time-space separating two crisis episodes, as if "…each generation would need to make is own mistakes."(Kindleberger, 1978).

2.1.3 Institutional quality and regional studies

So far, I have confined my theoretical recap exercise to macroeconomic and microeconomic dynamics. Nevertheless, the existing literature is not limited to these two categories. More recent papers have suggested a set of variables able to proxy, to a certain extent, the quality of the institutional framework in a country. In this direction, the effect of the presence of a deposit insurance, either explicit or implicit, has been widely discussed but, nowadays, results are still conflicting. A safety net on accounts should dissuade depositors to run on a solvent bank just as a consequence of an information shock, thus limiting systemic collapses by neutralizing self-fulfilling crises (Diamond and Dybvig, 1983). On the other hand, the presence of such a guarantee could trigger moral hazard and induce banks' managers to give up with prudent practices for more profitable and risky opportunities (Kane, 1989). Indeed, the US Savings & Loan crisis of the '80s has one of its roots on the moral hazard created by generous deposit insurances (Demirgüç-Kunt and Detragiache, 2002). Moreover, as depositors and overall creditors feel their resources as guaranteed their surveillance effort on bank executives' initiatives is due to get weaker, leaving greater room for mismanagement and market discipline deterioration. Both theoretical views are plausible and the empirical responses seem to sustain them equally. Demirgüç-Kunt and Detragiache (1997) showed a positive correlation between the presence of an explicit deposit insurance scheme and the likelihood of a systemic banking crisis. Thus, their findings would suggest that the negative

moral hazard effect on a banking sector outweighs the benefits deriving from a reduced risk of depositors' runs. Similar conclusions are drafted by Barth, Caprio, and Levine (2004). On the other side, the reverse view is encouraged by the results shown in Eichengreen and Arteta (2000) and Lambregts and Ottens (2006), whose studies were both carried on a set of developing countries. In these papers, the presence of an explicit deposit insurance is found respectively: insignificant and negatively correlated to the probability of a crisis. Apparently, in these circumstances, the role of deposit insurance turns out to have a beneficial effect on the banking system of developing countries, where bank liabilities are more short-term and panic-driven runs represent a huge threat. The most exhaustive document in terms of deposit insurance is certainly the one developed by Demirgüç-Kunt and Detragiache (2002), as they test for different features and coverage degrees of deposit insurances on a world-wide dataset. The results of their multivariate logit model suggest that the presence of a deposit insurance is detrimental for the stability of a banking system as much as extended is the coverage on the deposits. Its negative influence on bank stability becomes gradually stronger once there subsist other elements as: a lax regulation on bank interest rates, a weak institutional environment or the governmental nature of the insurance. Beyond the mere existence of a safety net on accounts, Demirgüç-Kunt and Detragiache (1997) in their first work considered even GDP per capita and the law and order index as further proxies for the quality of institution. They proved to have a significant predictive power with a negative relationship with the probability of a crisis.
With the same purpose, the level of transparency has been tested in the paper of Mehrez and Kaufmann (1999) whose findings confirm the intuitive existence of a linkage between a scarce degree of transparency and the likelihood of financial meltdown. Another variable considered proxy for the quality of institutions is the contract enforcement index, included by Arteta and Eichengreen (2000) in their emerging markets study. The evidence provided is weak and suggest no correlation between the variable and the onset of a banking crisis.

Some studies distinguish from others because built on a regionally or economically restricted data set of countries. This effort comes as consequence of the expected substantial differences among the causes of systemic banking crises between industrialized and developing countries or eventually different geographical-related cluster. Theoretically speaking, this mismatch comes as a result of profound discrepancies between developing and mature financial sectors. On average, emerging market economies are shallower, thus with a limited capacity to absorb economic shocks. The few available financial instruments do not allow to properly hedge risks that most of the times are concentrated in a very condensed number of businesses, as real economy still has to properly develop too. Regulation and supervision, as weak as the whole economic environment, are often ineffective in contrasting deficient management practices. In young financial sectors, few banks usually account for a large share of total assets while their liabilities have shorter maturity, therefore magnifying their exposure to sudden liquidity shocks (Rojas-Suarez and Weisbrod, 1996).Moreover, these banking systems tend to heavily rely on capital inflows and FDI (Foreign Direct Investments), leaving them vulnerable to unexpected high capital mobility. This was exactly the case of Thailand that experienced one of the unfamous Asian crises concentrated on the second half of the '90s. This episode appears as a perfect example of some of the elements just described. Banks used to borrow capitals from their foreign branches, subsequently investing them on the local real estate market and, thus, fostering the housing bubble through imported funds. High interest rates attracted huge amounts of resources from abroad, further fueling the economic boost of the country. Low regulatory standards and an aggressive lending (common aspect with advanced economies) eroded banks' assets quality and multiply their exposure towards the booming real estate market. When the US dollar appreciated against some Thai trading partners' currencies, the Thai exports went down as well, because of the dollar peg nature of the domestic currency. As speculators began to attack the fixed exchange rate and the capital inflows sudden turned in outflow, the country entered into a deep recession that climbed up until the burst of the housing bubble. Banking sector share of non-performing loans reached 46% of total asset. Its bailout cost was estimated in around $60 billion or 42% of GDP (Ergungor and Thomson, 2005).

Even the connection between the banking sector and the public finances seems stronger for developing countries. Arteta and Eichengreen (2000) showed that a fiscal deficit increase is associated with a higher financial instability in emerging economies, while the same variable resulted insignificant on the extended dataset employed by Demirgüç-Kunt and Detragiache (1997). A similar result on emerging market countries is obtained by Davis, Karim and Liadze (2010) that found the relevance of the fiscal surplus-to-GDP ratio to be secondary in strength just to the widely accepted GDP growth.

The real case previously exposed helps me introducing another topic concerning systemic banking crises over which researchers have given their responses. The aim of this effort was to clarify whether a particular exchange rate regime could increase the vulnerability of a financial system to external shocks. Even on this issue theory is divided. A floating exchange rate is expected to absorb, to some degree, external shocks on capital flows and terms of trade (as in the Thai episode). This view is supported by the works of Gavin and Hausmann (1996) and Mendis (1998). The insulation effect is not the only benefit attributed to a flexible regime.

Its presence, in fact, could detain banks' management from borrowing in foreign currency, as their exposure would skyrocket in case of domestic currency devaluation. On the other hand, a fixed rate, being affected by a lower volatility, especially for developing countries, would reduce the deterioration risk of bank balance sheets particularly exposed towards currency mismatch between assets and liabilities. Some weak empirical evidence supporting this view is brought by the paper of Arteta and Eichengreen (2000). A peg system could also help to discipline policy makers (Eichengreen and Rose, 1998) and to promote prudent banking practices as a consequence of the absence of a lender of last resort (Demirgüç-Kunt and Detragiache, 2005). Overall, the faint outcomes in both directions seem to provide further support to the findings of Kaminsky and Reinhart (1999) that proved a causality relationships going from banking crises to currency ones, not vice versa. Thus, empirically speaking, there is not any robust proof for either exchange rate regime to shelter or induce to systemic banking crises. It is rather demonstrated that the pressure triggered by a banking crisis on a country's currency could force policy makers to abandon a peg exchange rate, being the only way to properly carry out their role as lender of last resort.

This is exactly what happened in Mexico, during the so called "tequila" currency crisis. In November 1994, a sudden and steep capital outflow increase started eroding the Mexican foreign exchange reserves as a result of the central bank decision to defend the peg exchange rate with the dollar. Despite Mexican governmental finances appeared to be sound, the capital flight seemed to be triggered by political chaos and the choice by US Fed Reserve to raise its interest rates. Once the speculation on the dollarized peg became unbearable for the Mexican foreign exchange reserves, the central bank decided to leave the peso freely floating. By the end of December 1994, the currency devaluated by 35%. The peso collapse led to a strong increase in inflation and interest rates. Banking system came out affected both by its foreign exchange exposures and the hiking share of non-performing loans, as more and more borrowers could not afford the high interest rates. Ultimately, the government had to bailout the financial sector for up to 20% of GDP. Kaminsky and Reinhart (1999), in their widely-cited work, verified the huge economic cost associated to the so called twin crises, i.e. when banking and currency crashes occur simultaneously. What still does not find solid empirical support is the role of international shocks, as in this case the Fed interest rate movement, with respect to the likelihood of systemic banking crises. A restrictive monetary policy and an economic recession in industrialized economies are expected to negatively influence the economic environment in developing countries, especially dollarized ones. This was the case of Mexico in 1994, as well as other Latin American countries during the '80s. Eichengreen and Rose (1998) showed a significant positive correlation between monetary policy tightening or a growth slowdown in industrialized countries and financial sector fragility in developing ones. However, a later study by Arteta and Eichengreen (2002) on an extensive crises dataset, including the second half of the '90s (and therefore the Asian crises), found the same correlations to be weaker and less statistically significant. This result inevitably challenges the role of external factors as determinants in a comprehensive EWS for systemic banking crises, while strengthen the belief that banking crises evolve overtime. This is probably the main reason because both industrialized and developing countries could graduate from sovereign defaults but not from banking crises (Reinhart and Rogoff, 2008). The quest over the topic is still widely open, as the variety of its future research scenarios.

2.1.4 Findings and beliefs over crisis duration

Until now, I paid much attention on the causal forces that could lie behind the inception of a banking meltdown. However, unsurprisingly, some of these elements do not merely contribute to the burst of a crisis but they also play a role in determining the length of the downfall. As a consequence, most of the variables listed above that will not find place in my logistic analysis will still have a role in my forthcoming crisis length inquiry. The peculiar features of my duration analysis let my work cross the thin but still concrete boundaries that divide the EWS quest and the empirical branch aimed at shedding light on the severity of a crisis and the forces that drive it. With this purpose, Wilms et al. (2014) in their recent paper collected the most significant drivers of the real impact of a systemic banking crisis and tested them against nine different crisis severity measures, three of which regarding the event duration. They found that pre-crisis GDP per capita is strongly positively correlated with all their duration-based dependent variables which, in turns, suggest that defaults experienced by richer countries tend to last longer. Though my duration variable will be based on a raw mean computation instead of a GDP trend analysis, as the case of Wilms et al. (2014), I expect to find some matching results as well as brand-new empirical evidences. To a smaller extent, even financial openness and currency crisis related indicators demonstrated a positive and significant relationship with the severity of a crisis. However, the sign of the first correlation is controversial and matter of debate. Financial integration may offer risk-sharing opportunities and reduce the risk carried by a sudden stop in capital flows (Abiad et al., 2009). On the other hand, a loose regulatory framework on cross-border transaction may open the way for international financial shocks to spread into a country banking system. Similarly controversial is the role played by financial depth. More developed financial systems may suffer exponential losses during stressed period but at the same time deeper ones may provide a wider range of instruments to hedge from risk or optimally allocate it to stronger players, allowing a faster recover (Yanagitsubo, 2004). A steep growth of the financial sector size could highlight the creation of a credit bubble and an ongoing process of financial liberalization, along with lending standards rapidly deteriorating to permit the industry to keep on growing. Boissay, Collard and Smets (2013) found that greater financial booms are associated with longer and more painful recessions. Even exchange rate related downward fluctuations are expected to prolong a crisis period, especially when these movements set the stage for a currency crisis. The concomitant occurrence of banking and balance of payment crisis (twin crisis) would exacerbate the real costs and the persistence of the default (Berg, 1999; Kaminsky and Reinhart, 1999), especially when a consistent share of the country economic players is heavily exposed towards foreign currency liabilities.

An exhaustive literature review in matter of the mechanisms hiding behind the severity of banking crisis falls outside the scope of this empirical work. However, each of the forces described above will be represented, to some extent, by proxy indicators during my duration analysis.

2.2 The statistical models

The previous paragraph should have provided the reader with a summary up-to-date understanding on the forces behind a systemic banking crisis and their usefulness in building an effective early warning system. The aim of this section of the paper consists in gaining a better knowledge on the features of the main forecasting models employed so far. This further step would ultimately lead us to the end of the historical review and the introduction of the material purpose of this work. Similarly to what exposed on the determinants of a crisis, not all the early warning models that I will briefly describe further on would play a role in the execution of my empirical exercise but their description still remains valuable for the sake of a satisfactory knowledge over the empirical topic.

2.2.1 The signal extraction approach

The signal framework has been firstly developed and introduced by Kaminsky and Reinhart (1999), in their influential and widely-cited study on the twin crises. As briefly exposed in the introductory chapter, this method consists in testing individual variables behaviour against tranquil period values all over the 24 months preceding the burst of a crisis. If the variable exceeds a certain threshold, it would trigger a crisis signal that becomes true just if, within the following 24 months, a real crisis would really be registered. Otherwise, the variable would generate a false signal. The threshold is set to minimize the noise-to-signal ratio that looked like the ratio of false alerts to good warnings over a horizon of 24 months prior to the crises. The predictive power of the individual signals is then compared to each other performances on the basis of three yardsticks: the noise-to-signal ratio, the percentages of Type I and Type II errors, the conditional probability of crisis conditioned to a signal of the variable. A perfect indicator would correctly predict all crises without issuing misleading warnings. Unfortunately, in selecting the threshold for a variable, the analyst must compute a trade-off between Type I and Type II errors. A smaller threshold will inevitably lead to less Type I errors, as the number of missed crises decreases, and more Type II errors, as the number of false alarms increases. Although this choice should be made weighting the preferences of the policy maker, between the cost of missing to predict a crisis and the one incurred by implementing unnecessary measures because of a false alarm, Kaminsky and Reinhart opted for a statistical criterion like the noise-to-signal ratios (choice later on contested by Gaytán and Johnson, 2002). This methodology implicitly assumes that the cost of missing to identify a crisis (Type I) is higher that the loss generated by taking unnecessary precautionary initiatives because of an elevate number of false alarms (Type II). Even if the authors did not effectively own this kind of information. Under this assumption the three variables that graduated as the most significant ones have been the real exchange rate, the equity prices and the M2 money multiplier. The signal approach was subsequently adopted by various authors for their research exercises. Among them Rojas and Suarez (1998) that tested the predictive performance of aggregated banking ratios (CAMEL) on the likelihood of systemic banking crises. This early warning system has its key strengths in its transparency and relative simple understanding, which make of it a policy maker "friendly" approach, and the opportunity it offers to highlight early warning features of individual indicators. Benefits that seem to be far outweighed by drawbacks when the aim of the model is not anymore to measure single signals but to build a comprehensive early earning model based on a set of explanatory variables. First of all, the predictive accuracy of the model is deeply questioned. Either in the same original work (Kaminsky and Reinhart, 1999) or in other papers adopting the univariate signal approach, the share of correctly predicted crises within a horizon of 24 months after the issue of the signal is always lower than 30%. Performances that remain poor even when the test is conducted on out-of-sample crisis episodes. To improve the accuracy of the univariate version of the model several authors employed composite indicators. Borio and Lowe (2002) constructed aggregated indexes that would trigger a crisis signal just if all the variables included in the composition have crossed their thresholds simultaneously. Differently, Davis and Karim (2008), in their comparative framework on banking crises EWS, created composite indicators that gathered together variables on the basis of their individual NTSR (noise-to-signal ratio) and could issue a signal without necessarily that all the individual indexes contained in the composite have surpassed their cut-off. To some extent, this solution proved able to increase the model accuracy, by reducing both Type I and Type II errors, and to account for the aggregated effect of the variables included in the composite index. Other critics were directed to the very nature of the thresholds that separated a tranquil period behavior from a crisis one. Firstly, a single cut-off did not allow to register any difference in the marginal contribution of a variable to the probability of an oncoming crisis. In practice, the predictive role of an indicator should be relevant although its value has not fully reached and overcome a specific threshold, but for example just barely touched it. Secondly, the selection of common cut-off values for a cross-section dataset, considering them uniformly valid for all the countries included in it, appears to be an over-simplistic and inefficient solution. As suggested by Davis and Karim (2008), the optimal threshold may differ along with the peculiar structural characteristics of the institutional, political and economic environments of a specific country. As a consequence, the choice of country-specific thresholds would be preferred and could consistently improve the indicators performance. Finally, Gaytán and Johnson (2002), in their EWS review, stressed the inability of the signal approach to consider regional differences and to carry out study on the severity of a crisis. Most of the criticisms raised on the model, especially those regarding its lack of predictive efficacy, are not draft on absolute terms but have been highlighted by the comparison with other early warning frameworks. Indeed, few of them demonstrated their superior effectiveness in the role of early warning systems for banking crises and keep on gathering large consensus among the researchers committed in this field.

2.2.2 Binomial Logit Approach

Among them there surely is the multivariate logit approach, the one that probably experienced the most widespread diffusion since its first application on systemic banking crisis prediction. This model, as well as its probit version, is essentially a binary dependent variable model and its employment in predicting a banking crisis is based on the assumption that such events can be proxied by a response variable having just two possible outcomes (0, 1). The likelihood of a crisis is determined by a function of a vector of explanatory variables (x_k), which represent the right-hand side of the regression while their resulting coefficients (β_j) reflect the marginal contribution of each variable to the output probability. Logit and probit stand for two different cumulative distribution functions (G), standard logistic and standard normal, that ensure that the estimated probabilities are strictly between zero and one, for all possible values of the parameters and their associated variables.

$$P(y = 1|x) = G(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k)$$

$$0 < G < 1$$

The goal of the model is to understand which is the effect of the variables x_k on the response probability $P(y = 1|x)$, which in our case it is the likelihood of experiencing a crisis. In other words, the aim of the regression is to obtain the parameters β_j that can tell us some features of the relationship between x_k and the dependent variable. Information as the direction of the relationship and its statistical significance con be easily disclosed. However, because of the non-linear nature of the underlying relationship, the variables coefficients do not exactly match with the marginal magnitude of the regressors, as the case for a simple linear regression, and so cannot be strictly interpreted as the marginal increase or decrease in probability associated with a specific variable. To extract this information it would be necessary to run twice the regression, with and without the variable in question, while keeping all the other variables fixed and take the difference between the results. In other words, all the variables x_k must be taken into consideration. In any case, the effect exerted by the variable under scrutiny on the crisis probability would depend on the country's initial crisis probability. Whether the likelihood of a crisis would already be extremely high (or low), a change of a variable would not deeply affect the outcome of the regression, vice versa, starting from a probability of 0.5 would mean attributing a much more considerable weight to a movement of the variable.

A limited dependent variable model does not rely on an ordinary least squares estimator to obtain the parameters of the regression but on the maximum likelihood method (MLE) that better fit with a non-linear relationship. The MLE returns those coefficients that maximize the log-likelihood function:

$$\mathcal{L}(\beta) = \sum_i^n \ell(\beta) = \sum_i^n \{y_i \log[G(x_k\beta)] + (1 - y_i) \log[1 - G(x_k\beta)]\}$$

, being β the vector of coefficients and G, in case of logit approach, the logistic function:

$$G(x_k\beta) = \exp(x_k\beta) / [1 + \exp(x_k\beta)]$$

Although even the probit approach has been adopted in the attempt of building an early warning system for systemic banking crises, as demonstrated by the work realized by Eichengreen and Arteta (2000), the multivariate logit model has definitely achieved a greater success and implementation throughout the whole empirical literature on banking crises. As already anticipated, Demirgüç-Kunt and Detragiache (1997) have been the first in adopting a multivariate logit approach for studying banking sector crises. In their empirical session, they faced some obstacles in adapting the econometric model to the specific scope. As explained further on, the ways to tackle these complications still represent major opportunities for future improvements. First of all, they had to decide how to treat the years of data following the onset of a crisis, as in this period the relationship between the explanatory variables and the probability of a crisis would almost certainly be distorted by the crisis itself. They proposed two different solutions to tackle the problem, both capable to solve it even though at some costs. One of these consisted in eliminating all those observations following the year of banking crisis. This choice has the expensive drawback of wasting large amount of information (for example regarding multiple crises). Alternatively, they considered the possibility of discarding just those observations included between the burst and the end of the crisis. This solution had some advantages in terms of data gain and predictive power, as the dataset would have been considerably enlarged, but the choice of when the crisis was due to end was almost completely arbitrary, leaving room for the inclusion in the model of observations that could still be affected by the protracted influence of the crisis event. Since its first appearance, this phenomenon has been known by the author dealing with its treatment under the name of *post-crisis bias* (Bussiere and Fratzscher, 2006) or *crisis duration bias* (Caggiano et al., 2014).Another possible source of weakness (specifically speaking reverse causality) was the timing of some explanatory variables. Since their very first application, variables as the GDP growth and the credit growth were lagged by one or even two years (Demirgüç-Kunt and Detragiache, 1997) to both: better fulfill to their role as early warning indicators and avoid any kind of endogeneity bias. The preference towards lagged indicators further developed throughout the last decade until some of the most recent works, in which all the explanatory variables included in the regression were at least one year lagged (Davis, Karim and Liadze, 2010; Caggiano, Calice and Leonida, 2014).The endogeneity problem was not the last barrier that Demirgüç-Kunt and Detragiache (1997) had to overcome to set the model for the prediction of banking crisis. The presence of country fixed effects is often useful as it allows to account for the country-specific changes in the outcome probability. However, their employment requires that all those countries that have never experienced a crisis would be excluded by the dataset. Measure that would have implied a remarkable loss of information for the authors, whose cross-country database would have been more than halved. Under these circumstances, they preferred to preserve the entire dataset at the cost of removing the fixed effects. Other concerns were much less related to the specific features of the binomial logit while commonly shared by all the methods developed for the same purposes. Among them, the need to define a threshold probability presented again. With these regards, Demirgüç-Kunt and Detragiache (2000) adopted a methodology based on the loss function of a policy maker. Once some additional information, as the costs of taking preventive action and experiencing a crisis, are known this method is expected to be a more valid alternative to a selection based on a statistical index as the noise-to-signal ratio. Whilst a trade-off between missed crisis and false alarms is still necessary, a threshold selection centered on economic considerations would make the predictor much more efficient in the eyes of who is expected to finally exploit the model: the policy maker.

Once these issues have been set-aside, the binomial logit and probit models gave proof to outperform in terms of prediction accuracy most of the other models, whether tested in or out-of-sample. Since its first application as early warning system for banking crises, the efficacy of the multivariate logit approach has been repeatedly confirmed by those researchers that decided to rely on it to test their empirical hypothesis, as well as those authors that tried to compare different frameworks in a horse race fashion. Among them, it is worth mentioning the works of Davis and Karim (2008) and Alessi et al. (2015). Respectively, the first one benchmarked the performances of the signal and logit approaches for banking crises only, while the second one confronted nine different EWS together for a wider spectrum of financial crises (currency, sovereign debt, etc.). In both cases the models were ran on common cross-country datasets, crisis definitions and set of explanatory variables. Whereas the signal model demonstrated a better shape for a country-specific application (Davis and Karim, 2008), given the chance to set the threshold on a country by country tailored way, the multivariate regression models uniformly confirmed their superior predictive efficacy over all the other methods (in terms of correctly predicted crisis and non-crisis years). Nevertheless, both authors wisely promote prudence in relying entirely on a model rather than another. EWSs should fit as much as possible to the preferences of a specific policy maker and its reference country (Davis and Karim, 2008), while being supported by other models to further strengthen their forecasts robustness (Alessi et al. 2015). Despite the promising performances expressed so far by the binomial logit approach, its limitations still appear debilitating. In this regard, its multinomial version apparently provides room for further improvements.

2.2.3 Multinomial Logit Approach

Once again, as was the case for the binomial model, the multinomial innovation has been prompted by the emerging need of halting the remarkable accuracy leakages generated by the drawbacks affecting previous empirical frameworks. More specifically, some researchers tried to cope with the phenomenon that just recently has been reported under the names of *post-crisis bias* or *crisis duration bias* and whose compromising effects were known since the very first applications of the binomial multivariate models. The dilemma was centered on which modus operandi should be implemented to deal with the abnormal behavior of some explanatory variables in the years right after the crisis bust. As previously introduced, the way such data would be treated could lead to a biased model, whether the EWS was left free to erroneously consider them as tranquil period values, or to a relevant loss of information, in the case the researcher decided to tackle the problem by simply discarding the troubling periods (or, drastically, all the observations following a crisis episode). It was exactly the compelling necessity of handling differently the years right before the crisis that led Hardy and Pazarbașioĝlu (1999) to adopt, for the first time, the multinomial logit approach on the prediction of banking crises. Although, their exercise was not devoted at the resolution of the *post-crisis bias,* it still represents the very first attempt to introduce a third possible outcome in the regression, beyond the classical tranquil and crisis status supported by the binomial approach. The multinomial logit approach still belongs to the family of the limited dependent variable models but differently from its binary version it allows for more than two outputs. Thus, the additional outcome status for the response variable would introduce a higher flexibility, rather than being constrained to define a year either tranquil or crisis regime. Considering the post-crisis bias hypothesis, with three possible outcomes (S): one representing tranquil periods (S=0), another for the crisis times (S=1) and the additional one addressing post-crisis years (S=2), the theoretical outlook of the logit model would be as follow:

$$P(Y = S|X_k) = \frac{e^{\beta_S X_k}}{1 + e^{\beta_1 X_k} + e^{\beta_2 X_k}}$$

, where $X_k$is the vector of the regressors. If the tranquil period (Y=0) is adopted as control group than the vector of the coefficients β_1 will represents the marginal effect of the independent variables X_k on the probability of being in crisis period relative to the probability of being in a tranquil period. By the other hand, β_2 will represent the marginal contribution of the independent variables X_k on the probability of being in a post-crisis period relative to the probability of being in a tranquil period:

$$P(Y = 0) = \frac{1}{1 + e^{\beta_1 X_k} + e^{\beta_2 X_k}}$$

$$P(Y=1) = \frac{e^{\beta_1 X_k}}{1 + e^{\beta_1 X_k} + e^{\beta_2 X_k}}$$

$$P(Y=2) = \frac{e^{\beta_2 X_k}}{1 + e^{\beta_1 X_k} + e^{\beta_2 X_k}}$$

As a consequence, the relative probabilities are simplified as follow:

$$\frac{P(Y=1)}{P(Y=0)} = e^{\beta_1 X_k}$$

$$\frac{P(Y=2)}{P(Y=0)} = e^{\beta_2 X_k}$$

The maximum likelihood estimator keeps on being valid for the multinomial as well as the binomial logit model, given their common non-linear features.

This model has proved efficient in solving the *crisis duration bias* problem both in case of currency and banking crises. For what may concerns currency crises, Bussiere and Fratzscher (2006) developed a benchmark exercise between binomial and multinomial logit models, showing that the second one consistently outperforms the first. The advantages brought by the third possible outcome are visible through an overall increase in the predictive performance of the model, with a higher percentage of correctly predicted crises and a lower number of false alarms, as well as the full exploitation of the data on the post-crisis period. Results that have been lately confirmed in the banking crises field by the work of Caggiano et al. (2014) that tested the multinomial logit approach on a dataset of low-income Sub-Saharan countries. Similarly to the previously described work on currency crises, they compared the performances of the multinomial and the binomial models. Likewise, they obtained strong responses on the predictive superiority of a multinomial EWS, thus providing further proof of the existence of a *duration crisis bias* and the validity of this model to efficiently manage it. These improvements translated into a higher share of correctly predicted crisis (more than 3% of increment) and a lower percentage of false alarms (of around 2 %), relative to the binomial performance. Caggiano et al. (2014) adopt an additional method to measure the goodness-of-fit of the model. They draft the ROC curve (Relative Operating Characteristic; Figure 1) along with each regression executed. It plots the ratio of true crisis signals (True Positive, Sensitivity) against the ratio of false alarms (False Positive, 1-Specificity) for any possible threshold value. The result is a curve passing by the origin (0, 0) and the point (1, 1) whose extreme outcomes represent the worst and the best predictive performances. Respectively, a model with null predictive capacity (random guess) would produce a diagonal ROC curve starting on the origin and crossing the graph till the right corner point (1, 1). In this case, the

AUC (Area Under the Curve) would be equal to 0 and the model could tell as nothing more than a random guess on the forecasted event. Conversely, the perfect prediction method would lead to a ROC curve that starting in the origin would reach the right corner (1, 1) passing by the left corner (0, 1). Thus, the AUC would be equal to 1 and the model would be able to optimally forecast the crisis event. Therefore, as much as the AUC produced by the EWS is close to 1 as better would be its predictive accuracy.

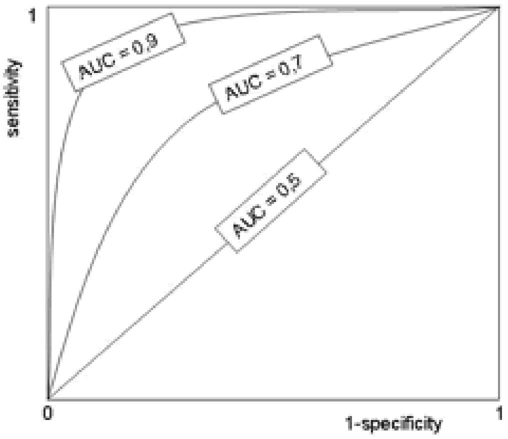


Figure 1: ROC curve and AUC illustrations.

Even under this valuation criteria, the multinomial model confirmed its improved predictive ability, with an AUC equals to 0.77 compared to 0.69 of the binary approach (Caggiano et al. 2014).

At the light of the progresses showed so far by this multiclass empirical framework, it is reasonable thinking at it as the EWS that could effectively move the empirical frontier on systemic banking crises towards a comprehensive model. However, its recent and scarce applications in the field do not provide enough evidences to already consider it completely reliable as a new standard. A secondary goal of this paper is to provide further support to the role of the multinomial logit model as best-in-class EWS for banking sector defaults.

Chapter III:

Setting the Empirical Stage

3.1 A definition for "systemic banking crisis"

Building a banking crisis variable requires a critical but necessary decision to be taken. The main complexity of this task derives from the very nature of the systemic event itself. It does not take place at a specific point in time but spreads out over a period whose boundaries are blurred. Identifying with precision its beginning and ending date is not possible as the economic phenomena (e.g. bank run) that could highlight the presence of an ongoing systemic meltdown could in turn have been the effects themselves of a preceding period of turbulence and local banking insolvency. Counterintuitively, the crisis could also be dated too early in the case its climax still has to be reached. Therefore, my analysis, as the ones conducted by the other authors in the field, should rely on a specific crisis definition that inevitably presents some degree of subjectivity. Rogoff and Reinhart, along with their comprehensive study on financial crises "This time is different"(2009), provide one of the most updated and recognized systemic banking crisis databases that covers the period between 1860 to 2014 (in its very last version). The definition proposed hereby splits banking crises in two different kind of events: "… (I) bank runs that lead to the closure, merge or takeover by the public sector of one or more financial institutions; (II) and if there are no runs, the closure, merging, takeover or large-scale government assistance of an important financial institution (or group of institutions), that marks the start of a string of similar outcomes for similar financial institutions." (Rogoff and Reinhart, 2009). The authors defined as "systemic" just those events described by the first definition (I) while the other crises (II) were considered as "…milder episodes of financial distress." (Rogoff and Reinhart, 2009). Upon these indications we have taken into account just the crisis dates that fall into the first crisis category (I), thus considering severe and systemic events only.

A strong limitation carried by this database regards the quite restricted spectrum of countries therein covered. With these premises, matching data availability from both left and right hand side variables of the regression would drastically shrink the count of effective countries included in my empirical process. To cope, at least marginally, to this sort of information leakage I relied on alternative dating sources that could have been integrated with the original one provided by Rogoff and Reinhart. First of them is the data set built by Caprio and Klingebiel (2003) which, according to Rogoff and Reinhart (2008), is "…authoritative, especially when it comes to classifying banking crises into systemic or more benign categories". In addition, to double-check my crisis integration process and make it as more robust and aligned as possible with the crisis definition described above, I have taken into account the papers of Boyd et al. (2009) and Chaudron and de Haan (2014), both comparing different techniques of banking crisis dating. The systemic crashes confirmed as such by all these three sources and compatible with the definition adopted by Rogoff and Reinhart would be reported by my crisis dependent variable. At the same time, those countries which demonstrated to have suffered just mild non-systemic defaults have been included in the base group of countries which have never experienced a systemic wide meltdown (e.g. Gabon, Guyana, Jordan). All those countries that lie outside of either one of these two eventualities have been excluded from the sample to preserve the integrity of the dependent variable. The same treatment has been reserved for those countries whose systemic banking crises, although strongly confirmed, had not a unanimously agreed duration (e.g. Republic of Congo, Niger, and Sierra Leone). This decision has been taken in order to avoid that misinterpreted information could bias the crisis duration analysis that will take place after the logistic regressions. This integration framework allowed me to preserve a satisfying size for my data set and enhance its heterogeneity by recovering information on several low income countries that otherwise would have been discarded.

Built upon these assumptions, the dependent variable accounts for 76 crisis episodes and a total of 324 years of crisis (out of 2380 total annual observations) within the full sample of countries. By restricting the data set to its smaller version (1998-2013) the amount of crisis episodes falls dramatically to 15 periods and 92 overall crisis years (within a total of 704 annual observations).

3.2 Explanatory variables

The selection of the independent variables has been inspired and dictated both by the theoretical framework behind the determinants of banking crises and data availability. In particular, my decision has been inspired by some of the most prominent papers in the field, almost mirroring the variable selection made by Caggiano et al. (2014) which, in turn, has been driven by the choices of Demirgüç-Kunt and Detragiache in their well-cited study back in 1997. As exposed previously during the review of the literature, these variables have found widespread empirical support from most of the previous banking-related studies and could be grouped in two main categories: macroeconomic and real sector concerned ones, financial and monetary the others. The following indicators belong to the first group:

- *GDP growth*: it proxies the economic wealth of a country and, consequently, the social fabric within which the banking system is embedded. A stagnation or decline of GDP should anticipate difficulties of the borrowers to respect their liabilities towards banks and therefore a deterioration of banks' balance sheets, as the amount of deteriorated assets soars. At the same time, a prolonged and sustained GDP growth could highlight an economic bubble being in place during the periods right before the burst of a systemic default;
- *Inflation*: a low rate of inflation indicates a relatively stable macroeconomics environment which, in turn, contributes to a greater solidity of the banking sector. Furthermore, a spike in inflation is meant to foster the formation of an economic bubble as borrowing gains popularity to the detriment of saving;
- *Nominal Exchange Rate Depreciation*: excessive foreign exchange risk exposure by the banking sector may leave it vulnerable to sudden movements of the exchange rate. In other terms, it could reveal vulnerability of the banking system to currency crisis;
- *Change in Terms of Trade*: a terms of trade drop is due to be one of the main factors behind bank insolvency (Caprio and Klingebiel, 1997). Even if this finding has not be strongly confirmed in all the empirical results achieved so far (see Kaminsky and Reinhart, 1999), the significance of a sudden deterioration in an economy's terms of trade in anticipating a banking collapse has been widely proved. Evidence that becomes more robust in developing countries and small open economies;

The second group of variables includes all those related to the monetary specifics of a country or the characteristics of the financial sector, with a focus on deposit banking institutions:

- *M2 to Total Reserves*: this ratio is a measure to the resistance of an economy, as well as its banking system, to sudden capital outflows (balance of payment crisis) and the country's ability to defend its currency. A higher value should correspond to a greater probability of financial system crisis;
- *Credit-to-GDP growth*: with credit intended as domestic credit provided by the financial sector, this ratio should proxy credit growth and financial liberalization (Pill and Pradhan, 1995). It should help detecting the existence of a credit bubble and/or decreasing lending standards and, as a consequence, should have a positive correlation with the probability of entering in a crisis;
- *Banking Capital-to-Assets (CAR):* built through the aggregation at the country level of individual banks' balance sheets data, this ratio should indicate the leverage of the

banking sector. A greater CAR would lower crisis likelihood by strengthening banks' ability to face credit deterioration and unexpected losses;

- *Liquidity ratio:* a higher banks' credits to banks' deposits ratio, thus high illiquidity, would mean lower resistance of the system to sudden massive deposit withdraws and, as a consequence, a greater probability of a disruptive run followed by a crisis;
- *FX Net Open Position:* it is included to highlight any currency mismatch between the value of banks' assets and liabilities which, in turn, should represent a threat to the banking system in case of domestic currency depreciation;

The variables just described are listed in Table 1a, along with their main statistics, and will be included in the first logit tests. Others, instead, will be introduced and involved at a later stage, during the average duration regression analysis. The choice over which indicators to include in the first benchmark session has been certainly shaped by the wish of reproducing some of the most robust achievements in terms of crisis determinants and possibly contributing to their deeper understanding. Moreover, by keeping the model in a relatively simple form, its comparison to other EWS would be facilitated while not necessarily sacrificing a significant amount of predictive power. However, as a matter of fact, data availability constraints have definitely set the final boundaries of the selection process. With this limited but heterogeneous handful of variables I could avoid to compromise the length of the period under examination and with it the whole list of collected countries. The only exception has been made for the Capital-to-Assets ratio (CAR) whose variable (*Leverage)* application has been restricted to the period 1998-2013. This solution has been adopted in order to both minimize the information costs incurred by including the variable in the regressions and stress the model under completely different conditions in terms of time-span and covered countries. All the explanatory variables employed in the logistic regression have been winsorized at the 1 and 99 percentiles to shelter the results from influential outliers.

As the main purpose of the paper is to build and test a EWS, all the variables will be lagged by one period, with the only exception of *GDP growth* which is expected to exhibit relevant early behaviors from two years prior the onset of the crisis. This settings is even aimed at avoiding the occurrence of reverse causality effects as the crisis itself could affect contemporaneous variables values.

Most of the data collected were obtained from the World Development Indicator database, made by the World Bank, and the International Financial Statistics from the IMF. When data were missing, central bank statistics and other country-specific sources were used as supplements. Data sources in details, along with their associated variables and a brief description, can be found in Appendix 4.

Table 1a: Summary statistics on the full data sample (Rogoff and Reinhart).

Variable	Mean	Std. Dev.	Min.	Max.	N
GDP growth	3.55	3.91	-10.06	14.72	2380
Inflation	14.85	34.77	-8.64	265.2	2380
Depreciation	13.49	42.29	-21.41	307.55	2380
M2-to-Reserves	8.95	14.54	0.58	89.2	2380
Credit-to-GDP	2.46	19.44	-65.81	100.02	2380
Liquidity	101.51	49.95	25.08	323.11	2380
Net Open Position	8.91	23.50	-65.66	95.29	2380
Terms of Trade	0.81	12.96	-38.26	56.26	2380
Leverage[a]	8.72	2.99	3.3	18.7	704

[a]Statistics computed on the reduced version of the data sample (1998-2013)

3.3 The data sample

The initial wider version of the panel data set, considering the limitations carried by both dependent and independent variables as previously exposed, accounts for 70 countries over a 34 years-time span, from 1980 to 2013. In its narrower form, dictated by the availability of the *Leverage* ratio, the data sample includes 44 countries whose yearly observations cover the period 1998-2013. Both versions consist in a world-wide mix of countries that have been strictly selected on the basis of their available data. As a consequence, among them, there are all sort of industrialized and developing economies accounting for profoundly different features. While the full data set is formed as a heterogeneous mix of developing and industrialize economies, its reduced version, being built on the constraint imposed by the *Leverage* availability, is constituted mainly by wealth countries. This feature is expected to exert some weight over the empirical estimations. With the crisis dates provided by Rogoff and Reinhart, out of the total count of countries (in the full data sample) 16 do not report any systemic crisis event, either because they have never experienced a banking default during the examined period or because these events have been considered non-systemic under the authors' definition of banking crisis. This bunch of nations will serve as a base group. As the count of countries involved in the tests shrinks to allow the assimilation of the *Leverage* variable, even its base group has been scaled down to a total of 7 clusters. For what may concern the composition of the sample and its degree of heterogeneity, in its final and extensive version, it includes a wide spectrum of economies ranging from deeply undeveloped to industrialized and wealthy ones. According to the national classification groups based on income per capita provided by the World Bank, the full data set accounts for: 8 Low Income Countries (LIC), 22 Low-Middle Income Countries (LMIC), 20 Upper-Middle Income

Countries (UMIC) and 20 High Income Countries (HIC). LIC presence completely disappears once the sample is downsized in its smaller form. An in-depth specification of the income classification criteria and a complete list of the countries included in the exercise can be found in the data Appendix 1. A world map illustrating the pool of countries included in the full data sample grouped by income per capita can be found in Appendix 3.
This collecting effort was strictly due as a higher number of observations was desirable, especially considering the infrequent nature of the systemic banking distress episodes compared to other financial crises (e.g. balance of payment crises).

3.4 The exercise step by step

Before entering deeply into the details of the empirical tests and their outcomes, I would briefly provide the reader with a summary glance over the purposes behind their execution. In order to understand whether a *post-crisis bias* effectively materializes when adopting a binomial model, I will treat the crisis years following the first in two ways: by considering them just as a tranquil period observations or by excluding them from the regression. This latest approach, already implemented by Demirgüç-Kunt and Detragiache (1997), should provide a shelter, even though marginal, to the distortion effect produced by the burst of the crisis on the variables values. Additionally to this task, during the following robustness checks I will implement a multinomial logit model that, along with its predictive performance, should be able to further clear the fog gathered around this phenomenon and eventually strengthen the role of this model for future early warning practices in the field. The statistics, shown in Table 1b, can be interpreted as a very first evidence of an underlying *post-crisis bias*. In fact, the averages of the variables differ across the three dependent variable status suggesting that mixing all the crisis periods without any discrimination could set the stage for a data mistreatment and the consequent misleading results. Once the first binomial benchmark session would be over, an additional exercise that would possibly provide useful information in this direction consists in taking into account of the average crisis duration on a country basis and produce multiple comparable results by partitioning the sample tested at different duration thresholds. Plotted in line charts, the AUC areas coming out from these regressions would allow us to graphically observe whether discarding post-crisis years does effectively increase the model predictivity and if these resulting benefits change along with the inclusion in the tested sample of increasingly lasting crisis episodes.

Whether in this step a correlation between crisis duration and models performances does effectively exhibit, in the last part of the empirical phase I will try to dig further this relationship by looking for aspects in the economical, institutional and political frameworks of a country that could exert any influence on the likelihood of being subject to longer lasting (and damaging) banking defaults. In its entirety, the empirical path just exposed has the ambitious aim of both contributing to the improvement of the empirical understanding of a specific issue such as the *crisis duration bias* and to provide scientific-robust advice in an EWS set-up process for those policy makers willing to efficiently employ it.

Table 1b: Independent variables' averages grouped by crisis states.

	GDP growth	Inflation	Depreciation	M2-to-Reserves	Credit-to-GDP	Liquidity	Net Open Position	Terms of Trade	Leverage[a]
Tranquil times (0)	3.83	12.57	10.31	8.08	2.68	100.56	10.11	0.31	8.84
First year of crisis (1)	1.71	32.16	29.85	15.17	4.72	124.13	0.58	2.76	6.95
Post-crisis periods (2)	1.75	28.41	34.86	14.21	0.00	112.04	1.53	4.28	8.09
Total	3.55	14.85	13.49	8.95	2.46	102.51	8.91	0.81	8.72

[a]Statistics computed on the reduced version of the data sample (1998-2013)

Chapter IV:

Results

4.1 A multivariate binomial benchmark

In this section I will report the outcomes of the binomial logit model, including in the regressions most of those primary variables (see Table 1a) that allowed me to exploit the data sample in its extensive form. The impact of *Leverage* is separately tested as the data set needs to be downsized to the period covering the last sixteen years of observations, between 1998 and 2013, for nearly two-thirds of the original number of countries.

As anticipated, the first set of logistic regressions is made with a binary dependent variable (*SBCRR*) which has been built upon the information provided by Rogoff and Reinhart (2009) in their comprehensive database on international financial crises. *SBCRR* takes the value of '1' in case of systemic crisis and '0' during tranquil periods. Since the very first estimation, all the explanatory variables employed will be lagged by one period, with the only exception of *GDP growth* whose indications are expected to be valuable two years in advance. Table 2 and Table 3 illustrate the results for the model applied to the full data sample which accounts for a total of 70 countries all over the time-span going from 1980 till 2013. The two tables differ from each other just for the treatment reserved to the crisis periods after the first. Respectively, in Table 2 post-crisis years have been set at '0', in other words normal times, while in Table 3 these observations have been dropped from the sample to limit the expected bias. From a maximum of eight independent variables, I have gradually removed the ones that showed the weakest correlation with the dependent variable, reaching a minimum base of four indicators. Due to its widely-diffused presence within the literature landscape, *GDP growth* has been kept as base variable whatever performance it has exhibited during the tests. All the explanatory variables have been winsorized at the percentiles 1 and 99 to limit the influence of outliers on the results.

Table 2: Binomial logit results on the full data sample (1980-2013). Crisis years next to the first are set as non-crisis periods. GDP growth is the annual growth rate of the country's GDP at market prices based on constant local currency(2005 U.S. Dollar). Inflation is measured by the annual growth rate of the GDP implicit deflator. Depreciation is the annual change in the official nominal exchange rate of the local currency unit with respect to U.S. dollar. M2-to-Reserves is computed as the ratio between money and quasi money aggregate (M2) on total reserves. Credit-to-GDP stands for the ratio between domestic credit provided by financial sector and GDP. Liquidity stands as the private credit provifed by deposit money banks as a share of demand, time and saving deposits in deposit money banks. Net open position is measured as net foreign assets in current local currency unit on GDP. Terms of Trade is intended as the annual growth rate of the ratio between exports and imports (in current LCU).

	(1)	(2)	(3)	(4)	(5)
GDP growth (-2)	0.037	0.027	0.026	0.032	0.033
	(1.07)	(0.83)	(0.77)	(0.97)	(0.98)
Inflation (-1)	0.015***	0.017***	0.017***	0.017***	0.017***
	(4.25)	(4.51)	(4.20)	(4.17)	(4.08)
Depreciation (-1)	-0.0059*	-0.0078**	-0.0085**	-0.0086**	-0.0086**
	(-1.92)	(-2.42)	(-2.37)	(-2.47)	(-2.42)
M2-to-Reserves (-1)	0.023***	0.019***	0.020***	0.019***	0.019***
	(8.82)	(6.60)	(6.61)	(6.15)	(6.11)
Liquidity (-1)		0.0059***	0.0059***	0.0059***	0.0059***
		(3.53)	(3.51)	(3.55)	(3.51)
Credit-to-GDP (-1)			0.0055	0.0054	0.0055
			(1.43)	(1.39)	(1.40)
Net Open Position (-1)				-0.0067	-0.0066
				(-1.43)	(-1.43)
Terms of Trade (-1)					0.0028
					(0.35)
Observations	2240	2240	2240	2240	2240
Pseudo R^2	0.043	0.056	0.057	0.059	0.059
AIC	599.6	593.9	595.0	595.6	597.5
N of countries	70	70	70	70	70
Degrees of freedom	4	5	6	7	8
Wald chi-squared	101.7	100.8	101.6	102.4	101.1
Likelihood-ratio	-294.8	-291.0	-290.5	-289.8	-289.8

t statistics in parentheses

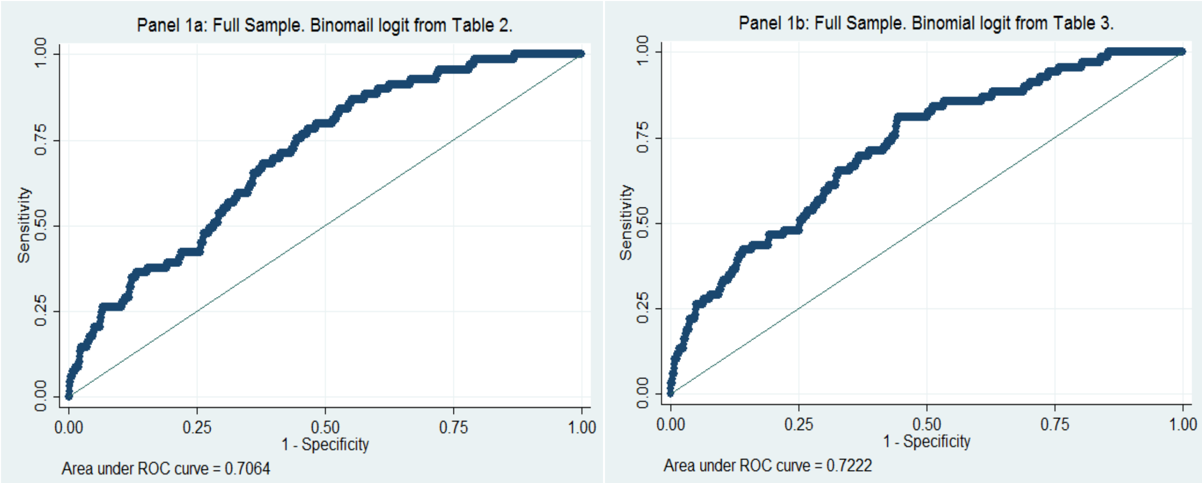
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Binomial logit results on the full data sample (1980-2013). Crisis years next to the first are excluded from the sample.

	(1)	(2)	(3)	(4)	(5)
GDP growth (-2)	0.031	0.018	0.014	0.021	0.020
	(0.83)	(0.50)	(0.39)	(0.56)	(0.54)
Inflation (-1)	0.022***	0.024***	0.027***	0.026***	0.026***
	(3.71)	(3.90)	(3.94)	(3.81)	(4.05)
Depreciation (-1)	-0.011**	-0.014**	-0.015***	-0.015***	-0.016***
	(-2.10)	(-2.54)	(-2.64)	(-2.64)	(-2.87)
M2-to-Reserves (-1)	0.028***	0.025***	0.025***	0.024***	0.023***
	(8.14)	(6.53)	(6.56)	(5.60)	(5.43)
Liquidity (-1)		0.0067***	0.0068***	0.0068***	0.0069***
		(3.60)	(3.74)	(3.71)	(3.71)
Credit-to-GDP (-1)			0.0076**	0.0076**	0.0083**
			(2.09)	(2.02)	(2.16)
Net Open Position (-1)				-0.0056	-0.0055
				(-1.06)	(-1.04)
Terms of Trade (-1)					0.0098
					(1.16)
Observations	1997	1997	1997	1997	1997
Pseudo R^2	0.060	0.076	0.078	0.080	0.081
AIC	573.9	566.5	567.1	568.1	569.2
N of countries	70	70	70	70	70
Degrees of freedom	4	5	6	7	8
Wald chi-squared	106.0	109.6	141.2	172.8	193.4
Likelihood-ratio	-282.0	-277.2	-276.5	-276.1	-275.6

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$



Most of the results coming from these first set of regressions confirm the findings of the majority of the papers draft in this field. Quite surprisingly, the role of *GDP growth* as a determinant for a systemic banking crisis did not prove to be strong with all the specifications tested, with the null hypothesis that has never been rejected at a satisfying significance level (lower than 10%). On the other hand, the positive sign of its coefficient is in line with the economic cycle view that wants a financial debacle preceded by a wealth bubble, inflated by an excessive credit availability and a viral over-optimism among economic players.
As the crisis inception date gets closer (within the previous 12 months), the relationship between the variable and the likelihood of entering in a crisis state turns negative, suggesting that the bubble has finally burst and the economy has started to spiral down into the recession. In light of this performance, the two-year lag, twice with respect to the other variables, revealed excessive, as a movement of the GDP gains statistical consistence as much as we get closer to the crisis outburst date. In fact, once tested contemporaneously on the same set of countries, the variable turns highly significant. This thesis finds support in the conclusions reached by Demirgüç-Kunt and Detragiache (1997) and Kaminsky and Reinhart (1996) whose analysis have been conducted on a similar world-wide heterogeneous country data set. Nonetheless, the robust result of the contemporaneous variable could be affected by reverse causality, thus reflecting the influence of the crisis on the economy rather than vice versa. A decline of *GDP growth* has demonstrated to manifest with a significant lag prior to a crisis for specific cluster of countries built on regional or income criteria (Caggiano et al., 2014). *Inflation* performed significantly better in all the specifications with a p-values often lower than 1% and the expected positive correlation with the probability of crisis. These evidences confirm the negative impact in the stability of a banking system provided by weak macroeconomic conditions during the period right before the oncoming of a crisis. The positive coefficient of the *Liquidity* variable shows that the more the banking sector is

exposed to credit risk the more it would be vulnerable to credit deterioration and therefore to a systemic meltdown. The ratio between banking sector credits on its deposits resulted deeply statistically significant in all the regressions, confirming its relevance as early warning indicator. As suggested by Caggiano et al. (2014), this relationship is especially threatening for emerging economies where, on average, banks are heavily exposed towards a restricted number of industries. As well, developing countries should be even much more sensitive to foreign exchange fluctuations as they are more than often characterized by dollarization and the financial sector balance between credits and liabilities is exposed to FX risk. However, employing the FX *Net Open Position* I could not confirm its relevance as powerful determinant for a banking crisis. As expected, a declining *Net Open Position* increases the likelihood of experiencing a crisis but the results displayed in the table are not able to strongly corroborate the statistical significance of this relationship. For what may concern *M2-to-Reserves*, it was expected to show a positive correlation with the outcome probability and so it is. This ratio, which works as a proxy for capital outflow vulnerability, proved statistically relevant in all the specifications, strengthening the theoretical allegations towards a sudden fall (rise) in capital inflows (outflows) harmful effect on a country's banking sector. *Depreciation* and *Credit-to-GDP* growth variables, as well as the previous ones, demonstrated to be valid early warning indicators. Nonetheless, their relationships with the output variable can find an explanation in the main theoretical frameworks. In fact, a drop in the nominal exchange rate is expected to destabilize the balance sheets of those institutions characterized by heavy foreign exchange exposures. Vice versa, the positive correlation showed by *Credit-to-GDP* growth confirm the role of a credit boom in setting the stage for a systemic banking crisis, at least as far as one year prior to the dated onset of the crisis. With these regards, the differences in the treatment reserved to the post-crisis years markedly affected the variable significance level, supporting the idea that the deceiving influence exerted by the mistreatment of these distressed observations can extend beyond the sole model predictive quality. Nevertheless, caution is required, as *Credit-to-GDP* variable is unique in this feature. Ultimately, a change in the *Terms of Trade* did not incisively emerge as early warning indicator performing poorly in both full-sample exercises. Its flimsy result can find a marginal explanation in the weak presence of low income developing countries within the sample.

Table 4: Binomial logit results on the reduced data sample (1998-2013). Crisis years next to the first are set as non-crisis periods. Leverage is intended as the ratio between banks' capital and banks' asset on a systemic basis.

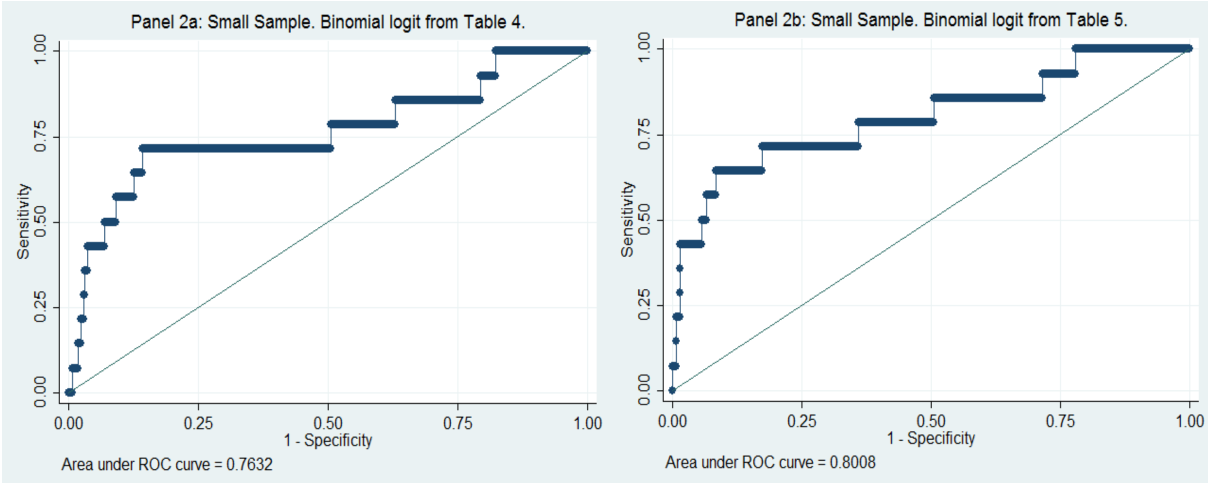
	(1)	(2)	(3)	(4)	(5)
GDP growth (-2)	0.022	0.066	0.066	0.076	0.065
	(0.29)	(0.71)	(0.70)	(0.77)	(0.70)
Inflation (-1)	0.031***	0.030***	0.030***	0.024**	0.027***
	(4.39)	(4.43)	(4.51)	(2.16)	(3.46)
Depreciation (-1)	-0.0032	-0.0028	-0.0028	-0.0052	-0.0067
	(-0.37)	(-0.32)	(-0.32)	(-0.47)	(-0.45)
M2-to-Reserves (-1)	0.034***	0.033***	0.032***	0.033***	0.039***
	(4.34)	(4.00)	(3.28)	(3.24)	(2.78)
Liquidity (-1)	0.0091***	0.011***	0.011***	0.011***	0.012***
	(2.90)	(4.33)	(4.53)	(4.29)	(3.76)
Credit-to-GDP (-1)		-0.00045***	-0.00045***	-0.00044***	-0.00041***
		(-3.76)	(-3.74)	(-3.68)	(-3.28)
Net Open Position (-1)			0.00061	0.0011	0.0011
			(0.06)	(0.10)	(0.09)
Terms of Trade (-1)				0.025	0.023
				(0.95)	(0.97)
Leverage (-1)					0.10
					(0.78)
Observations	660	660	660	660	660
Pseudo R^2	0.101	0.117	0.117	0.124	0.130
AIC	133.8	133.8	135.8	136.8	137.9
N of countries	44	44	44	44	44
Degrees of freedom	5	6	7	8	9
Wald chi-squared	31.9	40.7	40.7	68.2	73.3
Likelihood-ratio	-60.9	-59.9	-59.9	-59.4	-59.0

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Binomial logit results on the reduced data sample (1998-2013). Crisis years after the first are excluded from the sample.

	(1)	(2)	(3)	(4)	(5)
GDP growth (-2)	-0.037	0.0095	0.0020	0.034	0.021
	(-0.38)	(0.08)	(0.02)	(0.25)	(0.16)
Inflation (-1)	0.034***	0.033***	0.034***	0.025***	0.027***
	(4.16)	(4.20)	(4.28)	(4.69)	(5.02)
Depreciation (-1)	0.0033	0.0035	0.0033	-0.0043	-0.0040
	(0.35)	(0.36)	(0.35)	(-0.34)	(-0.30)
M2-to-Reserves (-1)	0.045***	0.043***	0.042***	0.044***	0.050***
	(5.29)	(4.93)	(4.02)	(4.00)	(3.24)
Liquidity (-1)	0.012***	0.014***	0.013***	0.015***	0.016***
	(3.81)	(5.58)	(5.59)	(4.94)	(3.96)
Credit-to-GDP (-1)		-0.00046***	-0.00046***	-0.00046***	-0.00043***
		(-3.61)	(-3.64)	(-3.58)	(-3.22)
Net Open Position (-1)			0.0040	0.0035	0.0032
			(0.37)	(0.30)	(0.27)
Terms of Trade (-1)				0.047	0.044*
				(1.58)	(1.67)
Leverage (-1)					0.094
					(0.69)
Observations	591	591	591	591	591
Pseudo R^2	0.147	0.161	0.162	0.182	0.187
AIC	125.0	125.1	127.0	126.4	127.7
N of countries	44	44	44	44	44
Degrees of freedom	5	6	7	8	9
Wald chi-squared	45.1	53.9	55.7	68.4	70.7
Likelihood-ratio	-56.5	-55.6	-55.5	-54.2	-53.9

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$



The second set of estimations carries two major differences with respect to the one previously commented. Firstly, the variable *Leverage* enters the scene in a regression, along with the main eight variables. Its correlation is not statistically significant and the positive sign of its coefficient is not anticipated by theory: a larger share of capital with respect to assets is due to act as a shelter and thus reduce the likelihood of a crisis. However, the model shows how this phenomenon does actually manifest getting closer to the time in which the crisis is expected to start. In fact, once employed the variable value in the current period T, the relationship with the probability of a crisis turns out to be effectively negative. I will not dig deeper into this peculiarity as it falls outside the scope of the exercise. Secondly, the data sample under scrutiny is restricted both in terms of covered time-span and of countries. In particular, low income countries completely disappear from the sample unbalancing its composition towards wealthy economies. As a result of this metamorphosis, the latest episodes, concerning the notorious Great Recession, acquire considerable weight. Therefore, not only the amount of observations drastically shrinks but even the results drawn upon them are expected to differ from the ones obtained using the entire data set. Looking at the outcomes of the regressions displayed in Table 4 and Table 5, the variable *Depreciation* is no longer statistically significant and its relationship with the output probability vary across the six different specifications. Looking closely at the residual data, once the sample has been squeezed out, it becomes pretty evident the fact that by leaving out many low income countries and overall large time intervals subjected to exchange rate shocks, the sample shape has profoundly changed. Notably, even the notorious episodes known as the Mexican "Tequila" crisis and the Asian crises, largely characterized by a drastic drop in the domestic currency valuation, are excluded in this occasion. Therefore, it comes weaker the effect of a currency collapse on the probability of observing a systemic banking crisis, as it should weight the most for those countries whose economies strongly rely on foreign currencies. This is especially true for all those economies, included in the full sample, that are dollarized and are consequently much more sensitive to exchange rate fluctuation. As well as for *Depreciation*, this reason is due to be also at the base of the robust insignificance reported by the *Net Open Position*. Essentially because low income countries' financial systems have normally worst net foreign assets balances and therefore are more exposed to a sudden negative shock of the domestic currency. A part from these two explanatory variables, the others confirmed their relevance (or irrelevance) as early warning indicators, highlighting some possible similarities in the ingredients necessary to set the stage for a systemic-wide banking crisis across different countries and time periods. A little criticism can be raised on the sign of the coefficient of *Credit-to-GDP*. Unexpectedly, in these last regressions it turned negative, contradicting the role of a credit boom in setting the stage for a banking crisis. What it is deemed to be one of the most supported theoretical belief, however, can vanish if in period T-1 the available domestic credit has already started to drop. If this is the case, this result appear much more plausible as once the credit bubble burst the lack of liquidity in the system is expected to make the economic recession, as well as the banking instability, even worse. Rogoff and Reinhart systemic definition is not exempt from this possible drawback as the event deemed to signal the beginning of the crisis episode, such as a run or a large-scale government assistance, could be the result of a downward spiral started well before. As partially verified with the full data set, the regression resulting coefficients do not change along with the deletion of the post-crisis observations.

In order to test the goodness of the models, I mainly relied on the in-sample predictive accuracy and the magnitude of the area under the ROC curve (AUC) generated by the model. Since post-estimation measures, as the percentage of correctly predicted crises (Sensitivity), Type I and Type II errors (respectively missed crises and false alarms), strictly depend on the subjective selection of the threshold probability needed to trigger a crisis signal, I will give priority to the model that has been able to produce the greater AUC. The ROC curve is more informative than the post-estimation classification table since it summarizes the predictive power for all possible thresholds. Furthermore, I took into account three statistical tests: the pseudo R-squared, the Wald chi-squared statistic and the Akaike Information Criterion (AIC). The first two of them, the pseudo R-squared, as suggested by McFadden, and the Wald statistical test, convey measures of the model quality by comparing the performance of its unrestricted version to its intercept-only one. Higher values for these metrics would correspond to a better goodness-of-fit of the model to the data. The AIC criterion provides a relative measure of the quality of a model, computing together the log likelihood of the estimation and the number of estimated parameters. Being the sign of the likelihood term reversed, smaller AIC values mean greater quality in the peculiar model specification (relative to the others being tested). Back to my results, it seems fairly striking that the models tested on the downsized data sets performs better that the ones which exploited all the available data. These conclusions are backed by an increased pseudo R-squared and Wald chi-squared. In all probability, by deeply reducing the data set heterogeneity and the number and variety of systemic events included within the tested period (15 episodes against 76 for the extended sample), the resulting cluster of countries is more efficient than naively pooling together all the available information. As sustained by van der Berg et al. (2008), this feature is particularly distinctive of financial crises. Because of their mutating tendency across countries and time-spans, these phenomena should be studied in optimal clusters that account for their specific characteristics. For what may concern the in-sample predictive ability check, to optimally perform this duty it would be necessary additional information on the costs that a government should bear in case of failing to signal a crisis and the cost of implementing useless preventive actions once a crisis is detected by the model. In other words, it would be needed the utility function of the policy maker. As a consequence, modelling a EWS on a country-specific basis represents the ideal solution to maximize its effectiveness, as confirmed by several authors. However this grade of specificity falls outside the real purposes of this work. Being unable to shape this threshold on the basis of a policy maker's preferences, I decided to set the cutoff probability equal to the ratio between the available crisis episodes within the data sample and the total amount of observations (once the variables have been lagged). Consequently, these percentages range from a maximum of 3.4% to a minimum of 2.1%, respectively for the full and reduced data sets. Based on these premises, by dropping the years following the burst of the crisis we obtained better results than simply considering these observations as tranquil periods. The improvements highlighted in Table 6 undoubtedly suggest the existence of the so called *post-crisis bias*. Taking into consideration just the models accounting for this issue, the widest specification from Table 3 correctly detected about a half of the crises, while the rate of false alarms (Type II error) amounts to a percentage of about 74.6%. Overall, the model correctly predicts a satisfactory 73.8% of the observations. By adding *Leverage* and restricting the data sample, as done in Table 5, the overall predictive ability of the model remarkably increases to a percentage barely lower than 80%. Both, the rate of crisis correctly predicted (Sensitivity) and the rate of non-crisis periods detected (Specificity) increase to 71.4% and 80.1%, respectively. In both cases, full and small sample, the models implementing the bias solution outperformed the remaining models in all the post-estimation metrics exposed above .To definitively identify whether a clear performance boost effectively emerges and which test actually outperformed the others, I looked to the ROC and the Area Under the Curve (AUC). Independently of the data set employed, the models adopting a dependent variable adjusted for the *crisis duration bias* (Panel 1b, Panel 2b) proved better than by erroneously considering post-crisis years as normal periods (Panel 1a, Panel 2a). On the basis of this measure, the quality of the regressions run on the reduced data sample fairly overcome the one executed on the full set, with a maximum AUC of 0.80 (reached by dropping the post-crisis observations; Panel 2b).

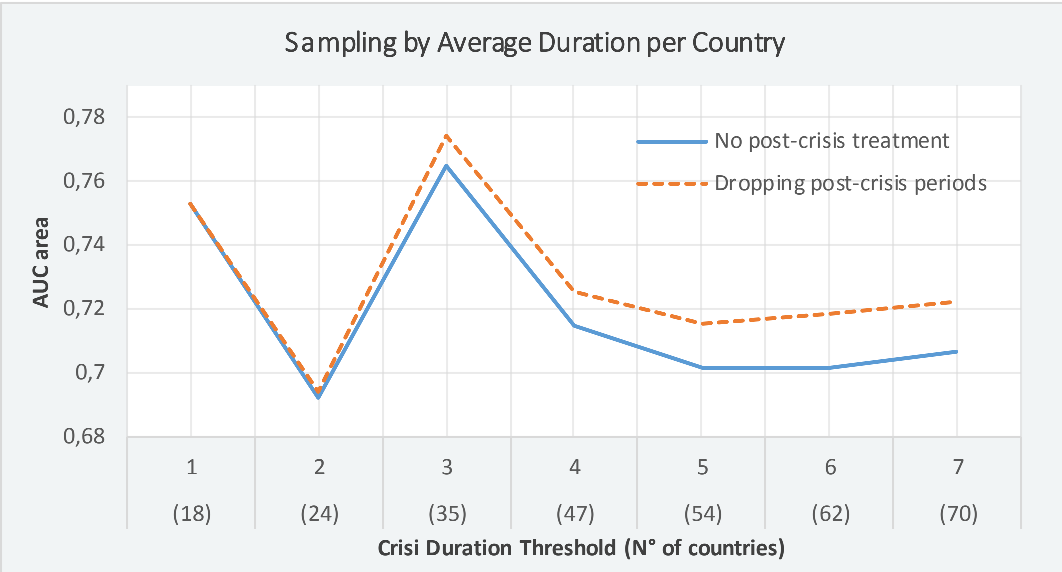
Table 6. Post-estimation statistics for all four binomial logit exercises.

Rogoff and Reinhart (2009)		**Binomial**	
	Metrics	*Treatment*	*No Treatment*
Full Data Sample (1980-2013)	Sensitivity (%)	50.7	49.3
	Specificity (%)	74.6	72.3
	Overall accuracy (%)	73.8	71.6
	AUC	*0,7222*	*0,7064*
Small Data Sample (1998-2013)	Sensitivity (%)	71.4	71.4
	Specificity (%)	80.1	74.6
	Overall accuracy (%)	79.9	74.6
	AUC	*0,8008*	*0.7632*

4.2 Graphing the Duration Effect

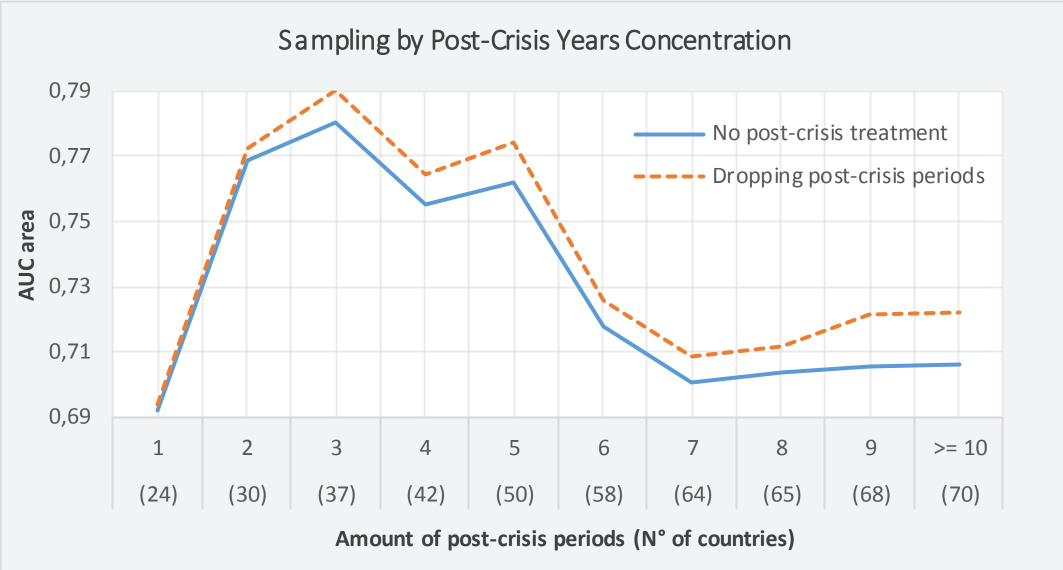
The following exercise shares the same spirit of the previous logistic analysis as it aims at adding further weight to the *post-crisis bias* hypothesis and providing, if any, a graphical representation of the relationship that binds together the duration of a crisis and the quality of the model in predicting these events. To the best of my knowledge, an exercise set up as follow has never been implemented before. More practically, I have tried to strengthen the convictions over the existence of a crisis duration related bias through plotting the AUC areas, as supreme measure for the model accuracy (being independent of any policy maker utility function), of multiple binomial logit regressions whose data sample has been discriminated on the basis of a crisis duration variable. This latest indicator has been computed as the average length of time of all the crises experienced by a country during the time-span covered by the full data set (1980-2013). The resulting graphs will show whether the predictive quality of the model moves along with an increasing presence of long-lasting crisis episodes in the sample. In fact, by gradually including countries whose average crisis duration is increasingly higher, there should be greater room for the *post-crisis bias* to manifest and influence the predictive goodness of the specification. As done for the previous benchmark exercise, the same regressions will be run both dropping the post-crisis observations and treating them as tranquil times in a horse race fashion. Whether a post-crisis effect does effectively persist and magnify along with the raising duration of the crises under scrutiny, the exclusion of the post-crisis observations should benefit to the predictive ability of the evaluated models. Because of the comparative meaning of this empirical session, the regressions therein are carried on considering only the more extensive form of the data sample and the eight initial explanatory variables illustrated in Table 2 (winsorized at the 1st and 99th percentiles). The reference crisis definition is the one proposed by Rogoff and Reinhart (2009), already employed in the previous logistic tests. The information reported for each logistic regressions has been limited to the AUC areas produced by their ROC curves and the number of countries involved in the regressions.

Graph 1. Effect on AUC areas of longer-lasting crisis episodes.



Graph 1 illustrates the results for a total of seven regressions. At each marginal increment in the *Crisis Duration Threshold*, all the countries that have experienced defaults whose average duration is equal or below this cutoff are included in the tested data sample until its maximum size is reached (70). A couple of exceptions has been made for the Central African Republic and Israel whose systemic crises did effectively start during the late 70s, protracted within my period of interest and ultimately vanished during the first observed decade. In order to avoid to exclude these observations or, even worse, simply truncate the crisis and align its beginning with the one of the covered period (1980-2013), I have considered these two episodes in their integrity when computing the respective average crisis duration value. The count of countries per regression is reported under the X axis of the graph in brackets. What stands out pretty evident from the graph is the irregular trend of the model accuracy while the list of countries keeps on inflating towards its maximum dimension. At a first glance, this result could firstly seem quite abnormal as more countries should mean more observations and therefore more useful information. However, this efficiency argumentation becomes quite pointless when studying financial crises because the factors expected to lie behind these events can profoundly change from country to country. Van den Berg, Candelon and Urbain (2008) found that not only a richer list of countries could not improve the model efficacy but it could even deteriorate it. This is exactly what we can observe from Graph 1. Their findings suggest that optimal clusters of countries can be even smaller than the regional dimension that has been purpose of studies for several researchers. Therefore, the AUC deterioration reported by both models can be the detrimental effect of the additional countries which effectively spoil better shaped clusters. Another plausible explanation consists in the existence of a negative relationship between the average crisis duration and the model predictive accuracy. Differently from the "optimal cluster" interpretation, this hypothesis could also explain why the spread between the two plotted lines gradually widens together with the average crisis duration of the countries included and, simultaneously, the average duration of the whole data sample climbs to its peak. In fact, at least partially, the trigger behind this quality deterioration seems to coincide with the *post-crisis bias* and this phenomenon magnifies along with the increasing length of the crises included in the sample. The spread between the two lines in Graph 1 confirms the effectiveness of the treatment adopted in tackling this problem. Nonetheless, there could exist frameworks (as the multinomial logit model) that could effectively allow a greater power recovery rather than just roughly discarding the suspected observations. Whatever solution is adopted, the take-away conclusion that emerges from the widening gap between the dotted and the flat lines suggests that the greater is the average crisis duration of the entire sample the more a proper treatment for the post-crisis years becomes beneficial for the sake of the predictive quality of the EWS. The performance differential between the two post-crisis treatments stabilizes once the data sample has almost regained its complete form. To a slightly limited degree, an ever increasing *post-crisis bias* is also what stands out from the exercise displayed in Graph 2 that plots the outcome of a test similar to the previous one. The main executive difference consists in the criteria that lays behind the partitioning of the data sample. While previously the list of countries grew as the average duration requirement was getting looser, this time it is the number of post-crisis periods (years of crisis after the first) per country that determines the data sample composition. Here, too, countries which have frequently gone through long-lasting systemic distress periods weight more than other countries, heavily affecting the improvement brought in by the duration bias solution. Once again the *post-crisis bias* hypothesis finds further support. Indeed, the model implementing the bias solution always outperforms the simple binomial framework and their performance gap hits its climax once the most crisis affected countries are taken into consideration.

Graph 2. Effect on AUC areas of post-crisis periods.



4.3 The hidden forces behind crisis duration

The results just obtained and plotted in the previous graphs suggest a positive correlation between the length of the crisis episodes and the predictive quality of the models tested upon them. Under this assumption, countries which are more prone to prolonged financial recessions should truly consider to adopt a *post-crisis bias* solution when implementing a logistic early warning system concerning systemic banking crises. Nonetheless, in practical terms, this hint comes pretty difficult to handle and exploit from a governmental point of view. In the attempt of providing policy makers with more quantifiable and representative indexes of the exposure of their country towards future enduring banking crises, I have collected a relevant number of variables that could eventually show some relationship among the political, economic and institutional framework of a country, and its tendency of experiencing above-average lasting defaults. Moreover, by regressing the duration variable on these metrics I could eventually uncover some of hidden forces that lie behind the duration of a banking crisis and, therefore, able to explain, to some extent, its severity. So far, these kind of investigations have been peculiar of that research area involved in estimating the real costs of a financial meltdown. In fact, crisis duration based indexes have been employed as proxies of the impact of the crisis on a country's economic environment. Wilms, Swank and de Haan (2014) did it in matter of systemic banking crises and found several indexes, ranging from institutional to more financial ones, significantly related to the duration of a crisis. High GDP per capita, in particular, emerged as symptomatic of prolonged recessions. Income per capita is considered to be a proxy for the quality of institutions and, in turn, is likely to be positively correlated with the effectiveness of the supervision effort over the relative banking industry (Demirgüç-Kunt and Detragiache, 1997). Under these assumptions and being it readily available for my entire set of countries I used it as base explanatory variable for the oncoming linear regressions. My dependent variable, previously employed as *Crisis Duration Threshold* in executing the AUC-Duration graphical exercise (Graph 1), is the average crisis length on a country basis and therefore constant for each of them. To accommodate this feature and the purpose of this test, the data set ceases to have a panel data shape while turning into a simple cross-sectional data sample with a single observation per country. As anticipated, the explanatory variables have very diverse natures among each other, ranging from indicators of specific features of the financial sector, such as its concentration, liberalization and depth, to institutional quality indexes. Their main selection criteria is focused on the existing literature in matter of real impact of banking crises, as efficiently summarized by Wilms, Swank and de Haan (2014), complemented with a pinch of the author's own curiosity. A part from those variables that were static by nature, as for the *Latitude*, the others had to be averaged over the entire sample (1980-2013), where available, or across a restricted time-span. Table 7 lists the whole set of variables involved in this analysis along with their coverage period, a brief description and the rationale behind their involvement.

Table 7. Independent variables for the duration-related OLS linear regressions.

Variable	Rationale	Description	Period
GDPmean	Wealth, Institutional Quality	GDP per capita in US dollar $ (current value) *	1980-2013
Latitude	Geographic position	Absolute value of the latitude of a country. Ranges from 0 to 1 (south to north)	-
OECD	Membership	Binary variable.	2013
Conc_Beck	Banking Sector Concentration	Fraction of assets held by the three largest banks *	1988-1997
Conc_Caprio	Banking Sector Concentration	Share of deposits held by the five largest banks *	1980-1997
Conc_3	Banking Sector Concentration	Fraction of assets held by the three largest banks *	1997-2013
Conc_5	Banking Sector Concentration	Fraction of assets held by the five largest banks *	1997-2013
RCC	Credit Boom	Total credit booms years by country based on the variation of real credit per capita from its long term trend	1980-2011
Qcluster	Institutional Quality	Weighted means of economic, political and legal institutional quality indexes (1-5)	1990-2010

Gov_Quality	Institutional Quality	Composite of: voice and accountability, political stability, government effectiveness, regulatory quality, rule of law, and corruption. From 0 to 1, from worse to better governance.	1998
frac_ent	Financial Liberalization	Number of entry applications denied as a fraction of the total number of applications received from domestic and foreign entities.	2001
restrict	Financial Liberalization	Sum of four measures that indicate the degree of restriction of bank activities in the securities, insurance and real estate markets and ownership and control of nonfinancial firms. From 1 (unrestricted) to 4 (prohibited)	2001
rr	Financial Liberalization	Ratio of reserves required to be held by banks	2001
bfree	Financial Liberalization	Indicator of relative openness of banking and financial system *	1995-1997
KAOPEN	Financial Openness	Also known as the Chinn-Ito index is an indicator of a country's degree of capital account openness. *	1980-2013
BCreditGDP	Financial Depth	Ratio of domestic credit to private sector provided by banks on GDP *	1980-2013
HHI	Market Concentration	The Herfindahl-Hirschmann Index (Product HHI), is a measure of the degree of product concentration. *	1995-2013
CurrencyCount	Currency Crisis	Total years of currency crisis experienced by the country.	1980-2010
SovereignCount	Sovereign Debt Crisis	Total years of sovereign debt crisis experienced by the country.	1980-2010
ConflictCount	Armed Conflict	Total years of armed conflict experienced by the country.	1980-2013
Foreign	Foreign Ownership	Percentage of banking system assets in banks that are 50 % or more foreign owned.	2001
State	State Ownership	Percentage of banking system assets in banks that are 50 % or more state owned.	2001
Muslim	Cultural discriminant	Variables that capture the percentage of population that is Muslim.*	1988-1997
Catholic	Cultural discriminant	Variables that capture the percentage of population that is Catholic.*	1988-1997

*Average over the period.

Unsurprisingly, most of these independent variables show strong multicollinearity with each other, especially among those sharing the same rationale. Therefore, bunching them together, although not affecting the predictive power of the models, would result in radically altered coefficients, as information redundancy skyrockets. By keeping just the average GDP per capita (*GDPmean*) as base variable and combining it with a different predictor at a time, I partially stem the phenomenon. Nonetheless, it still manifests pretty evident with some specific variable (Table 8). Another possible source of estimation bias could come from those countries which do not carry any information over their systemic banking default episodes (just mild episodes). This cluster of countries that was considered as part of the base group in the previous logistic regressions is mostly composed by low and middle income economies. It is a fact that some of these, particularly due to their tiny size, have never experienced a truly disruptive systemic episode or at least no one that could eventually fall under the crisis definition adopted by Rogoff and Reinhart (2009). However, including them in the linear regressions as free-crisis countries could distort the resulting coefficients. Therefore, the regressions have been run just considering those countries which had at least one documented systemic crisis, leaving out the entire base group previously employed during the logistic tests. Despite the information cost of this decision is burdensome, this solution was required to preserve the robustness of the analysis and the veracity of the relationships delivered by the regressions.

Table 8: Correlation and collinearity (VIF) indexes between GDPmean and the other variables.

IV	Correlation	VIF
GDPmean	1.00	1.00
Latitude	0.71	1.99
OECD	0.84	3.35
Conc_Beck	-0.29	1.09
Conc_Caprio	-0.06	1.00
Conc_3	-0.14	1.02
Conc_5	-0.31	1.11
RCCount	-0.47	1.29
Gov_Quality	0.89	4.72
frac_ent	-0.24	1.06
restrict	-0.38	1.16
rr	-0.25	1.07
bfree	0.47	1.28
ConflictCount	-0.15	1.02
CurrencyCount	-0.20	1.04
SovereignCount	-0.34	1.13
Qcluster	0.83	3.18
KaoMean	0.69	1.90
BCreditGDP	0.75	2.25
HHI	-0.39	1.18
Foreign	-0.21	1.05
State	-0.20	1.04
Muslim	-0.31	1.11
Catholic	-0.03	1.00

The number of countries involved in each regression (given each variable availability) and the resulting coefficients along with their t-statistics are displayed in Table 9.

Table 9: Linear regression on average crisis duration.

Duration	GDPmean	Latitude	OECD	Conc_Beck	Conc_Caprio	Conc_3	Conc_5	RCCount
Coefficient	-	-2.13	0.83	-1.69	-0.55	1.50	2.08	0.073
	-	(-1.06)	(0.72)	(-1.32)	(-0.28)	(1.02)	(1.12)	(0.45)
GDPmean	0.000050*	0.000076**	0.000022	0.000046*	0.000038	0.000059**	0.000057**	-0.00029
	(1.99)	(2.16)	(0.48)	(1.90)	(1.23)	(2.48)	(2.20)	(-1.26)
Obs.	55	55	55	44	28	54	51	31
F-Statistic	3.95*	2.54*	2.22	3.69**	0.82	3.32**	2.51*	0.94

Duration	Gov_Quality	frac_ent	restrict	rr	bfree	CurrencyC	SovereignC	ConflictC
Coefficient	0.67	0.91	0.21	0.063	-0.034	-0.015	0.068	0.035
	(0.81)	(0.84)	(1.45)	(1.50)	(-0.08)	(-0.29)	(1.66)	(1.16)
GDPmean	0.000013	0.000056**	0.000065**	0.00011**	0.000051*	0.000035	0.000061*	0.000054**
	(0.24)	(2.23)	(2.24)	(2.69)	(1.81)	(1.19)	(2.01)	(2.14)
Obs.	44	43	32	21	43	43	43	55
F-Statistic	2.53*	2.53*	2.73*	3.99**	2.01	0.97	2.37	2.66*

Duration	Qcluster	KaoMean	BCreditGDP	HHI	Foreign	State	Muslim	Catholic
Coefficient	0.56*	0.36	-0.00026	1.64	0.0072	0.0034	0.0045	-0.0091
	(1.83)	(1.08)	(-0.02)	(0.92)	(0.41)	(0.22)	(0.44)	(-1.23)
GDPmean	-0.0000064	0.000024	0.000050	0.000059**	0.000053**	0.000052**	0.000054**	0.000050**
	(-0.16)	(0.70)	(1.33)	(2.19)	(2.13)	(2.08)	(2.12)	(2.09)
Obs.	54	55	55	55	44	44	44	44
F-Statistic	4.60**	2.56*	1.94	2.40	2.26	2.19	2.28	3.00*

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Information related to the goodness-of-fit of the various models have been omitted in order to focus the reader's concentration on the correlations showed by the tested variables. The displayed outcomes do not seem to carry robust and valuable information as most of the variables resulted poorly correlated to the average duration dependent variable. Nevertheless, the positive relationship showed by *GDPmean* is statistically significant in most of the specifications and backed by the results obtained by Wilms et al. (2014), which found the correlation between the crisis duration and the income per capita to be the strongest among all their duration-related results. A part from the income-related variable, just the *Qcluster* variable, proxy for the quality of a country's institutional framework, proved faintly significant. These results suggest that countries with healthy institutions and above average wealth per capita are more subjected to long-lasting banking crises, once they have already plummeted in it. Interestingly, in spite of being backed by past empirical evidences, none of the four variables expected to measure the degree of financial liberalization, (*frac_ent*, *restrict*, *rr,* bfree), together with the count of currency crises (*CurrencyC*) have been able to show a significant relationship. As anticipated, behind the unsatisfactory report produced by the regressions there could be the negative influence of collinearity bias. With this regard, the author decided to report, just beneath the resulting coefficients from each specification, the F-Statistic of the individual analysis. This metric, outcome of a statistical test (F-test) in which the test statistic has an F-distribution, conveys the capacity of rejecting the joint hypothesis that the coefficients are both zero. Along with the correlation between the coefficients and the variance inflation factor (VIF) for each pair *GDPmean-Variable(IV)*, as displayed in Table 8, the F-statistic would become a valuable tool to highlight the presence of detrimental levels of collinearity between the tested variables. Understanding if a problem of this nature is actually affecting the results would not represent a priority whether the aim of the exercise was to learn more about the quality and the goodness-of-fit of each specification. Nonetheless, as the true goal of the linear regressions consists in highlighting any significant relationship between dependent and independent variables, the presence of undisputable signs of multicollinearity among these variables and *GDPmean* could threat the reliability of the exercise, by distorting the obtained coefficients and their significance levels. In this direction, anomalously high VIFs (which quantify the degree of variance distortion caused by collinearity to an estimated coefficient), greater than 1.9, accompanied by correlation values higher (or lower) than 0.5 (or -0.5) would trigger solid alarm signals. Additional concerns would arise whether an F-statistic (among those reported in Table 9), which strongly rejected the joint null hypothesis, was related to insignificant or scarcely significant coefficients, thus, unable to firmly discard the null hypothesis for the individual parameter. This investigation process led to a bunch of explanatory variables whose connection with GDP per capita was more or less intuitive. For example, it is deeply recognized that OECD members (*OECD*) are on average wealthy countries and that income per capita works even as measure for a country's institutional quality, mirroring the rationale behind the employment of *Gov_Quality* and the ordinal variable *Qcluster*. As it is well known that private credit to GDP ratio (*BCreditGDP*) strongly correlates with income level. A bit more blurred, instead, are the causes of the collinearity exhibited by *KaoMean*, the average of a financial openness index, and the banking concentration metric provided by Beck et al.(2000), *Conc_Beck*, although more concentrated banking sectors are generally associated with young growing financial systems. To definitively test whether a collinearity issue effectively exists I have regressed the duration dependent variable against the individual independent variables over which concerns of bias were strongest. Results from this analysis can be observed in Table 10.

Table 10: Variables characterized by high collinearity with income per capita.

Latitude	0.93						
	(0.63)						
OECD		1.30**					
		(2.07)					
Conc_Beck			-2.39*				
			(-1.88)				
Gov_Quality				0.84**			
				(2.26)			
Qcluster					0.52***		
					(3.06)		
KaoMean						0.52**	
						(2.16)	
BCreditGDP							0.012
							(1.44)
_cons	4.16***	4***	5.89***	4.17***	2.98***	4.31***	3.81***
	(9.96)	(12.17)	(6.38)	(14.52)	(6.03)	(15.40)	(8.06)
Observations	55	55	44	44	54	55	55

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As supposed, several of the variables expected to be affected by collinearity problems with *GDPmean* demonstrated a relevant and, in some cases, even strong direct correlation with the duration dependent variable. From this latest regressions, it emerges that the degree of a country's capital account openness and its banking system concentration can both play a role in determining the inclination of a specific economy towards longer-lasting banking defaults. More in detail, countries characterized by milder restrictions on cross-border financial transactions could be more exposed to drastic capital flights or contagion from international shocks in case of systemic default, while a fragmented (distributed) banking industry is due to be much harder to recover than just concentrating the recovery effort on few institutions. The same weakness is expected to be shared by those countries whose economic, political and institutional infrastructures are wealthier and robust. Nonetheless, it must be remembered that these considerations gain validity once a country is effectively plummeted into a crisis state. Therefore, if on average a wealthy and well-structured country demonstrated to be more exposed to lasting crisis events, the opposite relationship can become true once the variables in question are employed as early warning indicators in the attempt of forecasting those specific events. The results reached by Demirgüç-Kunt and Detragiache (1997) on GDP per capita and Law and Order index support this view. Differently, the positive sign exerted by *BCreditGDP*, the proxy measure for financial depth, is supported by previous findings, as a greater availability of instruments would allow to quickly relocate credit risk or eventually provide some shelter through sophisticated hedging strategies. Unfortunately, even when tested separately from income per capita, *BCreditGDP* coefficient could not prove statistically robust at a satisfying significance level.

Chapter V:

Robustness and Sensitivity Analysis

5.1 Determinants check-up

In the first part of the robustness analysis, I verify whether the results obtained on the determinants of a banking crisis are more or less solid. To do so, I keep the variables that showed a sound correlation with the oncoming of a crisis during the previous binomial exercises and I retest their performances under changed conditions, in terms of covered time-span, independent and dependent variable. A plausible doubt can be raised on the possible misleading influence that integrated crisis dates on the original Rogoff and Reinhart (2009) definition could have brought in the regressions. Once these data have been discarded from the model, the count of countries under scrutiny squeezes and the resulting outcome is illustrated in Table 11 (1). No noteworthy changes in the significance level of the variables coefficients or in their signs materialize, confirming that the integration process has been managed with adequate care. Thereafter, I have questioned the specific dating procedure and crisis definition employed so far, reproducing the logistic analysis on a dependent variable built on the information provided in the systemic banking crisis database by Laeven and Valencia (2012). The verdict of this test, as displayed in Table 11 (7), confirmed the beliefs gathered around the predictive quality of some indicators as the growth rate of the *Credit-to-GDP* ratio, the *M2-to-Reserves* and the *Liquidity* indexes while at the same time denying the centrality of the role played by the change in the exchange rate and the inflation rate. Along with the eclipsing of these variables, other indicators gained a stronger position in the analysis. *GDP growth* rate and *Net Open Position* are subjected to sharp movements toward satisfying significance levels. Both their negative signs have already been observed in the literature, telling of an increasing crisis likelihood in case of an economic downturn starting with a two-year lag and a one-year lagged negative net open position, highlighting a currency mismatch between banks' assets and liabilities that, in turn, magnifies any balance sheet deterioration caused by an eventual drop in the domestic currency valuation. Again, all these findings gather wide support within the dedicated literature, confirming the choice of the banking crisis definition as one the most troublesome decision that a researcher and/or a policy maker have to take in implementing a EWS. Moving from the left to the right-hand side of the regression equation, four additional explanatory variables have been tested, all being lagged by one period. Among these, the strongest signal is produced by the short-term lending real interest rate (Table 11/2) which confirms its relevance as early warning indicator and the detrimental effect of a lending interest rate shock on the solidity of banks' balance sheets. Higher rates are deemed to increase the chance of banking crisis one year in advance, both because they reflect an ongoing financial liberalization and an increasing risk of greater shares of non-performing loans. Notably, the interest rate variable proved to be a good proxy for financial deregulation as it contemporaneously drown out the effect carried by the growth rate of private credit, as both variables pursues the same rationale. Other weaker but still significant correlations are showed by the capital account openness index (*Kaopen*; Table 11/4) and the armed conflict dummy variable (*Conflicts*; Table 11/5). Respectively, these results suggest that countries with a tighter regulatory framework in matter of cross-border financial transactions and/or which are taking directly part into an armed conflict are more likely to experience a systemic banking crisis in the following year. In particular, a higher grade of financial openness is expected to make banks more resilient to surrounding shocks as they can lend abroad to partially hedge from domestic economy volatility. The export concentration index (HHI; Table 11/3) was deemed to capture to which extent an economy is focused on a restricted number of industries and, therefore, whether the banking sector degree of risk diversification could impact the probability of crisis. Although its sign is in line with the theoretical assumption that wants more concentrated economies highly exposed to full-fledged financial meltdowns, its coefficient did not proved statistically able to reject the null hypothesis at a satisfying significance level. Ultimately, the dummy deposit insurance variable did not show a significant correlation with the likelihood of experiencing a crisis (Table 11/6). Overall, the variables that during the main logistic analysis proved able to solidly perform as banking crisis predictors confirmed the robustness of their valuable contribution in the attempt of building an effective EWS.

Table 11: Binomial robustness tests on the full data sample (1980-2013). Crisis years next to the first are excluded from the sample. A part from Export Concentration, Kaopen and Conflicts, all variables have been winsorized at the percentiles 1 and 99. Real Interest Rate stands for the short-term lending nominal interst rate adjusted for inflation as measured by the GDP deflator. Kaopen, or Chinn-Ito index, reflects a country's degree of capital account openness. Conflicts is a dummy variable that indicates in which years a country has been involved in an armed conlfict. Explicit is a dummy variable that take value of 1 when is present an explicit deposit insurance scheme.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
M2-to-Reserves (-1)	0.021***	0.031***	0.032***	0.024***	0.024***	0.023***	0.018***
	(3.89)	(5.52)	(4.77)	(4.91)	(5.60)	(5.03)	(3.44)
Liquidity (-1)	0.0052***	0.012***	0.014***	0.0064***	0.0068***	0.0066***	0.0070***
	(2.91)	(6.54)	(7.76)	(3.41)	(3.68)	(3.86)	(4.85)
Credit-to-GDP (-1)	0.012***	0.010	0.012	0.0078**	0.0080**	0.0077**	0.011***
	(2.58)	(1.28)	(1.31)	(2.29)	(2.25)	(2.21)	(2.92)
Inflation (-1)	0.034***	0.042***	0.030***	0.028***	0.030***	0.029***	0.0058
	(5.29)	(2.96)	(4.32)	(4.62)	(4.79)	(4.85)	(1.11)
Depreciation (-1)	-0.022***	-0.000086	0.0019	-0.018***	-0.018***	-0.018***	-0.00026
	(-3.97)	(-0.01)	(0.21)	(-3.38)	(-3.34)	(-3.40)	(-0.06)
Real Interest Rate (-1)		0.041***					
		(2.78)					
Export Concentration (-1)			0.077				
			(0.04)				
Kaopen (-1)				-0.15*			
				(-1.81)			
Conflicts (-1)					0.53*		
					(1.81)		
Explicit						-0.061	
						(-0.26)	
Net Open Position (-1)							-0.012***
							(-2.65)
GDP growth (-2)							-0.051*
							(-1.88)
Terms of Trade (-1)							0.00074
							(0.08)
Observations	1475	1126	1152	2056	2064	2064	2302
Pseudo R^2	0.073	0.144	0.132	0.081	0.080	0.075	0.072
AIC	494.8	269.2	216.5	599.5	607.0	610.0	624.9
n of countries	51	63	70	70	70	70	78
Degrees of freedom	5	6	6	6	6	6	8
Wald chi-squared	131.9	92.8	117.3	153.4	120.9	141.8	257.7
Likelihood-ratio	-241.4	-127.6	-101.2	-292.7	-296.5	-298.0	-303.5

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.2 A different definition for banking crisis

A part from the outcomes in matter of crisis determinants, one major element that could have been key in misleading my results consists in the choice of the crisis dating source used to build my dependent variable. Therefore, to further consolidate my findings over the *post-crisis bias* and the role of the crisis duration in amplifying it, I have repeated the main exercises switching my initial dependent variable, built on the information provided by Rogoff and Reinhart (2010), with the one which relies on Laeven and Valencia (2012) dating procedure. Together with the work of Rogoff and Reinhart, the database built by Laeven and Valencia, whose latest version dates back to 2012, represents the state-of-the-art in terms of most updated financial crisis archives. Therein, a crisis episode is registered whether these two conditions coexist: significant signs of financial distress in the banking system (1) and significant banking policy intervention measures in response to significant losses in the banking system (2). Signs of financial distress are defined as e.g. bank runs, losses in the banking system and/or bank liquidations, while a policy intervention is considered "significant" whether it met some quantitative prescriptions, for example the bank restructuring gross cost exceeds the 3 percent of GDP, or qualitative ones, such as system-wide bank nationalizations, guarantees, deposit freezes and bank holidays. Built on these assumptions, the new data set allows the inclusion of countries that had to be previously discarded due to a lack of information under the crisis definition provided by Rogoff and Reinhart. Most of the recovered countries belong to the lower GDP per capita slots, thus contributing to an increasing diversification in the composition of the sample. The following economies, from almost all over the world, find a place in this analysis: Bangladesh, Belize, Chad, Republic of Congo, Dominica, Kuwait, Niger and Sierra Leone. With these premises and considering the extended version of the data sample (now including 78 countries), the resulting dependent variable counts in 80 systemic banking crises and 276 overall distressed annual observations. The adoption of this variable as regressand led to the results of the countercheck exercises on the logit models reported in Table 12. Here, I condensed the output information coming from the logistic regressions keeping what is effectively relevant in determining the power of the model as EWS and, if any, the existence of a *post-crisis bias*. The independent variables are the ones employed for the previous main logistic tests. As it can be clearly seen, the predictive quality improvement obtained by dropping the crisis years after the first is confirmed.

Table 12. Logit models predictivity power once Laeven and Valencia (2012) dating system is employed.

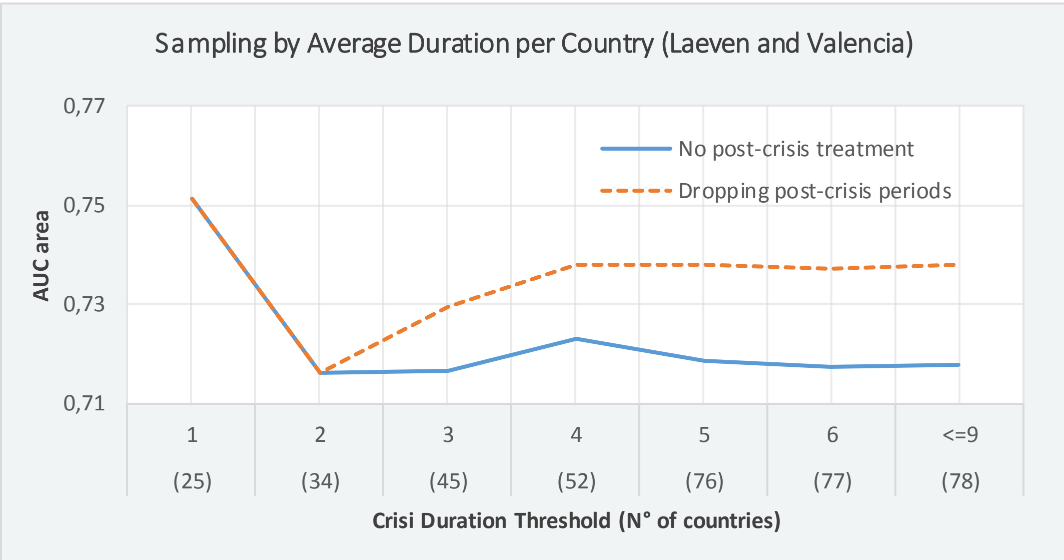
Laeven and Valencia(2012)		**Binomial**	
	*Metrics**	*Treatment*	*No Treatment*
Full Data Sample (1980-2013)	Sensitivity (%)	55.4	56.8
	Specificity (%)	72.2	70.5
	Overall accuracy (%)	71.6	70.1
	AUC	*0,7381*	*0.7176*
Small Data Sample (1998-2013)	Sensitivity (%)	75.0	75.0
	Specificity (%)	80.3	78.8
	Overall accuracy (%)	80.2	78.7
	AUC	*0.8548*	*0.8158*

*Specific cutoff probabilities computed as the ratio between the crisis periods and the total number of observation in each test (ranging from 1.7% to 3.2%).

Confirming the weight that the choice of the crisis dependent variable exerts on the performances of the individual warning indicators, *Inflation* and specially *Depreciation* lose their relevance while *Net Open Position* gains a stable and significant position in the model. Nonetheless, no coefficient turns its sign from positive to negative or vice versa.

For what may concern the second step of the duration analysis, a similar pattern in the AUC-Average Duration relationships can be appreciated in Graph 3 and Graph 4. The duration-based threshold variable had to be adapted to the new dependent variable while the methodology did not change. Still the spread between the two lines, representing respectively the two diverse post-crisis years treatments, keeps on widening as much as the countries included in the sample have a greater average crisis length. This result further confirms the existence of a relationship between crisis duration and *post-crisis bias* and therefore the quality of the EWS.

Graph 3. Crisis average duration analysis.



Graph 4. Incremental post-crisis bias.

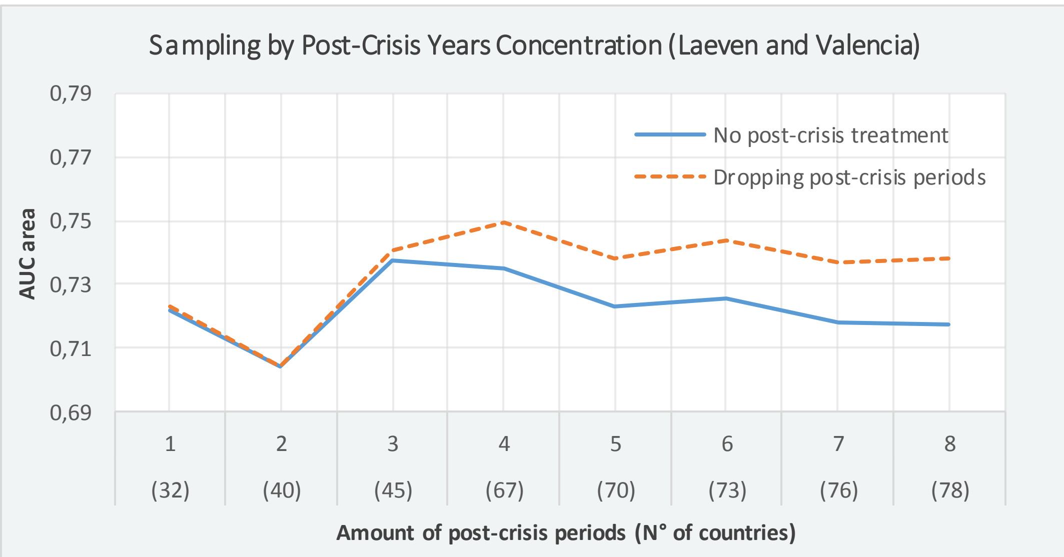


Table 13 and Table 14 report the results of the linear regressions on the average crisis duration. The coefficients and all their t-statistics therein confirm the relationship between a country's grade of wealth and the time length of the crises it is usually subjected to. Even the quality of institution confirms its relevance while the concentration of the banking system and the financial openness proxies lose their significance turning irrelevant. In their place, a proxy for financial liberalization stands out along with the export concentration index (*HHI*) and the count of currency crises (*CurrencyC*). These two latest findings, although empirically supported by some authors, completely lose their relevance once tested separately from income per capita while the negative correlation reported by *frac_ent* endures, conveying a thicker evidence. This outcome gives some hint on the possible relationship between financial liberalization and crisis duration, suggesting that the more the financial system is restricted into a tight regulatory jacket the shorter it would be the recession in case of full-blown banking crisis.

Table 13: Linear regression on average crisis duration (Laeven and Valencia).

Duration	GDPmean	Latitude	OECD	Conc_Beck	Conc_Caprio	Conc_3	Conc_5	RCCount
Coefficient	-	-0.75	0.18	1.21	0.95	0.29	1.34	-0.024
	-	(-0.55)	(0.22)	(1.29)	(0.85)	(0.27)	(1.00)	(-0.19)
GDPmean	0.000056***	0.000064***	0.000050	0.000063***	0.000067***	0.000055***	0.000060***	0.00010
	(3.08)	(2.66)	(1.60)	(3.38)	(3.55)	(2.95)	(2.97)	(1.35)
Obs.	62	62	62	47	29	60	55	35
F-Statistic	9.50***	4.84**	4.70**	5.75***	6.68***	4.35**	4.44**	0.93
Duration	**Gov_Quality**	**frac_ent**	**restrict**	**rr**	**bfree**	**CurrencyC**	**SovereignC**	**ConflictC**
Coefficient	-0.58	-1.48*	0.043	0.0063	-0.43	0.066*	0.041	-0.024
	(-1.13)	(-1.99)	(0.41)	(0.22)	(-1.59)	(1.70)	(1.32)	(-1.09)
GDPmean	0.000089**	0.000048***	0.000072***	0.000082***	0.000070***	0.000073***	0.000075***	0.000051***
	(2.57)	(2.76)	(3.63)	(3.08)	(3.49)	(3.38)	(3.23)	(2.74)
Obs.	48	46	33	22	47	43	43	62
F-Statistic	5.65***	7.39***	7.34***	5.04**	6.10***	5.91***	5.22***	5.35***
Duration	**Qcluster**	**KaoMean**	**BCreditGDP**	**HHI**	**Foreign**	**State**	**Muslim**	**Catholic**
Coefficient	-0.077	-0.15	-0.011	2.07*	-0.013	-0.015	-0.0015	0.0030
	(-0.32)	(-0.65)	(-1.24)	(1.93)	(-1.11)	(-1.56)	(-0.22)	(0.55)
GDPmean	0.000063*	0.000067**	0.000080***	0.000068***	0.000051***	0.000050***	0.000054***	0.000055***
	(1.93)	(2.65)	(2.99)	(3.61)	(2.85)	(2.80)	(2.89)	(3.11)
Obs.	61	62	62	62	47	47	48	48
F-Statistic	4.49**	4.91**	5.56***	6.83***	5.45***	6.17***	4.90**	5.06**

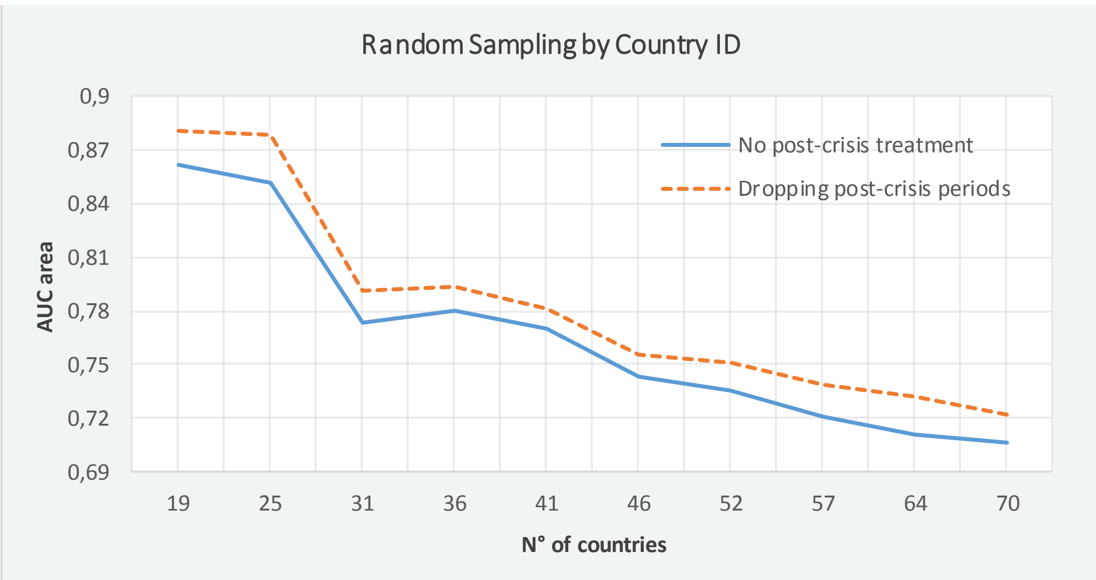
t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Variable characterized by high collinearity with income per capita. (Laeven and Valencia)

Latitude	1.67						
	(1.53)						
OECD		1.25**					
		(2.58)					
Conc_Beck			0.20				
			(0.21)				
Gov_Quality				0.56**			
				(2.04)			
Qcluster					0.31**		
					(2.24)		
KaoMean						0.28	
						(1.60)	
BCreditGDP							0.0087
							(1.38)
_cons	3.22***	3.25***	3.39***	3.48***	2.84***	3.52***	3.17***
	(11.10)	(14.12)	(4.63)	(16.36)	(7.47)	(16.86)	(9.45)
Observations	62	62	47	48	61	62	62

Besides changing the dependent variable, I have checked whether the same trend in the AUC-based lines could be achieved by randomly adding additional countries to the sample. To verify this hypothesis, instead of the average duration threshold as sampling variable, I used the id number built on the country's position in the alphabetical order. In order to recreate the same conditions as in the duration exercise, I have included all the zero-crisis countries (base group in the logit estimations) since the very first regression. Graph 5 shows that, running this exercise, the same ever widening spread between the lines cannot be observed anymore. This result strengthens the belief that simply adding crisis episodes to the sample is not a sufficient condition to improve the efficacy of the proposed *post-crisis bias* solution. Noteworthy, the model adopting the bias solution always outperforms the other.

Graph 5. A random sample test to confirm crisis duration influence on the model predictive power.



5.3 Multinomial logit model

In order to switch from a binomial to a multiclass framework, the discrete dependent variable had to adapt, from being a dummy to a categorical one. Once the change has taken place, it allows for three possible outcome states: tranquil (no crisis) periods $Y_{i,t} = 0$ (where i stands for the country and t for the period of observation), first year of crisis, $Y_{i,t} = 1$, and following years of crisis after the first, $Y_{i,t} = 2$. As anticipated, this distinction should permit me to avoid the distortion produced by those annual observations that, even though treated as tranquil times, are undoubtedly influenced by the economic distress triggered by the burst of the crisis. This phenomenon, so far called *post-crisis bias* or *crisis duration bias* and expected to stand out particularly accentuated for systemic banking crises, due to their long-lasting duration and effect on a country's economic environment, has already boldly emerged during the previous binomial exercises. The benefits generated by the multinomial model should not be limited to an improvement in the model predict power. In fact, by recycling the information carried by the post-crisis observation, the model should provide access to a deeper understanding of the variables partial effects on the crisis probability. To our knowledge, this represents the very first attempt in banking crisis literature of employing a multinomial logit model on a heterogeneous and extended country data set for EWS, with *post-crisis bias* related purposes. Hardy and Pazarbașioĝlu (1999), although similarly employing a mixed data set, exploited the incremented flexibility carried by the categorical dependent variable to distinguish between pre-crisis and crisis years and between severe and full-blown episodes. Other previous works, as the ones already mentioned, have focused on clusters exclusively composed by open market economies (Bussiere and Fratzscher, 2006) or region-oriented (Caggiano et al., 2014).

The multinomial tests fit the same specifics of the first binomial regressions. The crisis dating procedure employed is the one provided by Rogoff and Reinhart (2009), already examined in depth in the previous chapters, and the variables included as early earning indicators are the ones listed in Table 1a. Table 15 reports the results of the multinomial tests on the full data sample. The first panel displays the resulting coefficients from comparing the probability of entering in a crisis ($Y_{i,t} = 1$) against the one of experiencing tranquil times ($Y_{i,t} = 0$), our base outcome. The coefficients of the variables and their t-statistics, apart from the case of *Credit-to-GDP*, almost fully match with the outcome produced by the binomial model. No relevant differences are highlighted as variables significance levels, as well as the extent of their correlations, coincide with the results previously described. This outcome overlapping was fairly expected as both dependent and independent variables employed to test the probability of an oncoming crisis did not change along with the model definition. Furthermore, the similarity among the parameters offers a preliminary check for the existence of the *Independence from Irrelevant Alternatives* (IIA) assumption. Also known as *binary independence*, this assumption holds just in case the odds of experiencing the crisis event with respect of being in tranquil times does not depend on whether some other alternative is present or absent. This condition is essential for the multinomial model in order to be unbiased and valid. With this matter, the necessary evidences to verify whether this assumption statistically holds are provided by the Hausman-McFadden test (1984) and the SUE-based (Seemingly Unrelated Estimation) Hausman test. Based on the results obtained by running these analysis on my multinomial model (which can be observed in Appendix 3), I can conclude that the null hypothesis of independent alternatives cannot be rejected and, therefore, confirm that the IIA hypothesis holds. For what may concern the novelties brought in by the multinomial model, the results for the post-crisis period ($Y_{i,t} = 2$) are disclosed in the second panel of Table 15. Here, the variables contribution on the probability of remaining in a crisis state with respect of going back to a more tranquil period are reported. The peculiar behavior of some variables during post-crisis years (see averages in Table 1b) works again in favor of the *post-crisis bias* hypothesis, suggesting that these observations cannot be correctly compared to tranquil periods. Looking more specifically at the results, a prolonged economic recession is expected to further worsen the banking sector stability or at least to contrast its recovery. Indeed, *GDP growth* is strongly statistically significant in all the specifications of the model. *Inflation,* while turning its sign from positive to negative, loses most, if not all, its statistical consistency, signaling that once the crisis has already climbed or it is starting to manifest (being the variable one year lagged) the macroeconomic environment does not hold the main role it had as warning indicator in molding the future probability of the country to keep on experiencing the crisis or leave this state. As the case of the economic recession, the same deteriorating effect should be provoked by a growing ratio between banks' credits and banks' deposits which works as a proxy for the banking sector liquidity position. The displayed results confirm this belief with a *Liquidity* ratio significant and positively correlated with the likelihood of enduring in the crisis state. Even the *M2-to-Reserves* and the *Credit-to-GDP* ratios maintain the same signs as in the first panel but the p-value of the second one soars well above the 10% significance threshold, turning *Credit-to-*GDP statistically irrelevant. On the other hand, vulnerability from sudden capital outflow preserves its incisiveness on the probability of remaining under stress. *Depreciation* shift towards a positive correlation is supported by theory too. As the national current account benefits from the pre-crisis domestic currency depreciation, a turnaround in this matter would mean a drop in exports and a consequently worse balance of payment outfit. In this matter, the impact exerted by the change in rate of exchange vanishes once the change in the terms of trade is taken into consideration. *Terms of Trade* variable resulted highly statistically relevant in the fifth and more exhaustive specification (Table 15/5), apparently absorbing the effect of an exchange rate movement. In this context, an upward movement of the ratio between exports and imports appear to be detrimental to the state of the economy increasing the likelihood of a prolonged banking depression. The *Net Open Position* variable preserves its negative coefficient suggesting that even after the burst of the crisis, a higher favorable foreign currency mismatch between banking sector assets and liabilities can contribute to a faster recovery, as the threat of domestic currency devaluation comes weaker. However, its t-statistic is still too low to be able to discard the null hypothesis with a satisfying degree of certainty.

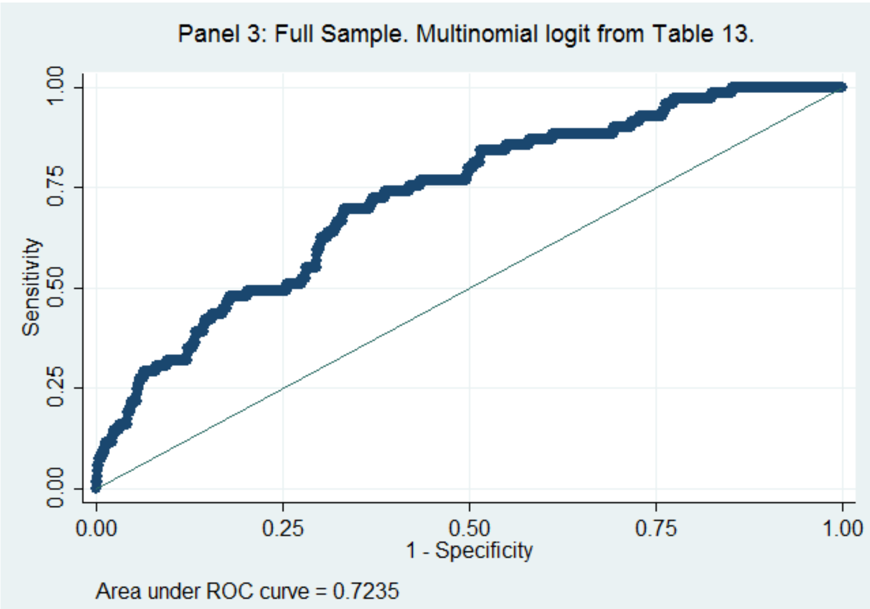
Table 15: Multinomial logit results on the full data sample (1980-2013).

	(1)	(2)	(3)	(4)	(5)
First year of crisis:					
GDP growth (-2)	0.0219	0.00934	0.00780	0.0148	0.0149
	(0.60)	(0.27)	(0.22)	(0.42)	(0.42)
Inflation (-1)	0.0146***	0.0163***	0.0172***	0.0165***	0.0166***
	(4.30)	(4.72)	(4.45)	(4.34)	(4.36)
Depreciation (-1)	-0.00371	-0.00530*	-0.00597*	-0.00625*	-0.00651*
	(-1.19)	(-1.65)	(-1.67)	(-1.81)	(-1.92)
M2-to-Reserves (-1)	0.0276***	0.0241***	0.0241***	0.0232***	0.0232***
	(7.30)	(5.68)	(5.68)	(5.20)	(5.18)
Liquidity (-1)		0.00710***	0.00705***	0.00698***	0.00698***
		(3.56)	(3.56)	(3.53)	(3.53)
Credit-to-GDP (-1)			0.00587	0.00576	0.00640*
			(1.59)	(1.54)	(1.74)
Net Open Position (-1)				-0.00711	-0.00704
				(-1.36)	(-1.34)
Terms of Trade (-1)					0.00715
					(0.92)
Following crisis years after the first:					
GDP growth (-2)	-0.0947***	-0.0989***	-0.0994***	-0.0943***	-0.0915***
	(-4.48)	(-4.73)	(-4.73)	(-4.45)	(-4.31)
Inflation (-1)	-0.00181	-0.00133	-0.00115	-0.00177	-0.000113
	(-0.28)	(-0.20)	(-0.17)	(-0.28)	(-0.02)
Depreciation (-1)	0.00832*	0.00797*	0.00789*	0.00757*	0.00573
	(1.93)	(1.83)	(1.79)	(1.76)	(1.31)
M2-to-Reserves (-1)	0.0231***	0.0203***	0.0203***	0.0191***	0.0195***
	(4.25)	(3.54)	(3.54)	(3.07)	(3.10)
Liquidity (-1)		0.00583***	0.00579***	0.00572***	0.00574***
		(3.04)	(3.01)	(2.90)	(2.86)
Credit-to-GDP (-1)			0.00162	0.00130	0.00242
			(0.37)	(0.29)	(0.53)
Net Open Position (-1)				-0.00643	-0.00624
				(-1.10)	(-1.07)
Terms of Trade (-1)					0.0199***
					(3.49)
Observations	2240	2240	2240	2240	2240
Pseudo R^2	0.070	0.082	0.083	0.085	0.092
AIC	2009.3	1986.2	1989.2	1988.4	1977.3
N of countries	70	70	70	70	70
Degrees of freedom	8	10	12	14	16
Wald chi-squared	131.1	121.0	162.8	182.8	330.6
Likelihood-ratio	-994.6	-981.1	-980.6	-978.2	-970.6

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Ultimately, in order to complete the benchmark framework between the two logistic models I have replicated the post-estimation analysis to compute the predictive quality expressed by the model. In doing so I could not rely on the same techniques adopted to elaborate the ROC curves and the related AUC areas for the binomial model. I have therefore implemented a tailored algorithm for multiclass ROC curves able to isolate and test separately each pair of outcomes among the three accounted by my crisis dependent variable. While the AUC was computed over the full range of possible cutoff probabilities, to determine the other post-estimation metrics, as Specificity and Sensitivity, I have adopted the same threshold previously used with the binomial model which accounted for the *post-crisis bias* (3.4%). Under these assumptions the multinomial model slightly outperforms its symmetric binomial version, as well as outperforming the binomial specification that does not account for any treatment in matter of *crisis duration bias*. Numerically speaking, the multinomial model had successfully predicted 44.9 % of the crisis episodes (Sensitivity), issuing false alarms (Type II error) just in the 17.1% of the cases (Specificity was 82.9%) and 81.9% overall observations correctly classified. The generated AUC, plotted in Panel 3 along with the ROC, resulted barely greater than that of the "treated" binomial (0.7222). Difference, however, that can be noticed just at the third decimal place of the value.



Chapter VI:

Conclusions

The latest full-blown "subprime" crisis, whose inception officially dates back to the end of 2007, reminded the international community, especially industrialized and highly financially-developed countries, how severely such kind of shocks can strike and how painful the recovery path to leave them behind could be. Though mutating their characteristic aspects through time, these events demonstrated to preserve some common features that can hopefully lead towards the implementation of tools able to predict and eventually prevent them from happening. Early Warning System models endorse these goals and, as the forecasting quality and precision offered by these instruments climb up, fostered by the interest gathering over the topic and the effort of the scientific research, policy makers are more than ever considering to effectively employ these frameworks during their financial decision process. In light of the pressing need to further refine these instruments, this paper firstly analyses the performances of some widely-recognized variables as early warning indicators for banking crises and the capacity of the binomial and multinomial logistic models to fulfill the EWS duties. From the regressions run on an extensive and heterogeneous panel data sample, including a wide range of crisis episodes and countries, it emerges a scenario that has been largely sustained by the theory in the field. An increasingly unstable macroeconomic environment, detected by a rising inflation, fosters the chance of a banking system to fall into a recession. The role of a lagged credit boom in setting the stage for a credit crunch and the consecutive deterioration of banks' balance sheets have been confirmed by most of the tests, especially those exclusively accounting for the Great Recession periods (limiting the sample from 1998 to 2013) which have been notoriously characterized by the burst of huge credit and real estate bubbles. An expansive credit market is due to be backed by a financial deregulation being in place, with falling lending standards. With this regard, increasing real interest rate one year prior to the burst of the crisis found robust empirical support and with it the belief that wants a systemic default preceded by an ongoing financial liberalization. Along with the downsizing of the sample, as it is the case for the application of the real lending interest rate (1994-2013) and the leverage ratio (1998-2013), the significance of a change in the exchange rate is drained away to firmly reappear with a negative coefficient when the regressions are run again on the extended panel (1980-2013). This behavioral swing proves either the deep differences in the causal force behind each specific crisis episode and, most often than not, the weakening effect of a drop in the rate of a country domestic currency exchange rate on the stability of the respective banking system. This is particularly due to the more or less wide exposure of the banks towards foreign exchange risks which are fed, for example, by borrowing abroad and simultaneously lending in domestic currency. The tested models unanimously agree on the positive relationship between the lagged values of the M2 to Reserves and the Liquidity ratio with the likelihood of experiencing a crisis. Respectively, the first ratio, by measuring the capacity of an economy to react to sudden capital outflows or a strong devaluation of the currency, indicates that, as its value increases, the exposure of the banking system and the overall economy to unexpected liquidity shocks sky-rockets. Either because depleted foreign exchange reserves limit government and central bank freedom of action in case of a rampant currency devaluation or because expanding monetary policy could foment inflation as well as currency devaluation too. Liquidity ratio, instead, is much related to the state of health of the banking system consolidated balance sheet and the capacity of these financial institutions to deal with tough, perhaps unexpected, credit deterioration. Unsurprisingly, the vulnerability of the banking system follows the trend of this ratio. When illiquidity expands, the system is more exposed to uncovered losses and vice versa. For what may concern the impact of the economic boom, the double lag imposed on the GDP growth variable seems fairly excessive, at least when adopting the crisis definition proposed by Rogoff and Reinhart (2009). In fact, its coefficient struggle to show significant early warning signs until the variable is tested contemporaneous to the crisis period (less than one year prior to the burst). Under these conditions, the variable confirms the strong negative impact that an economic recession could have on the banking system, as loan deterioration sky-rockets, but its merit as early warning indicator gets weaker. Kaminsky and Reinhart (1996), who have examined monthly data around the crisis, found that the GDP growth starts declining almost 8 months prior to the onset of the crisis, giving credit to these allegations. Nonetheless, the robust tests built on an alternative crisis dating procedure (Laeven and Valencia, 2012) and the outcome produced by the regional exercise of Caggiano et al. (2014) suggest that the timing with which an economic recession shows up before a banking crisis varies across different sample specifications and crisis definitions. An in depth examination of this specific interpretation is undoubtedly required, especially considering the rich literature sustaining the leading role of this variable.

A part from building a deeper understanding of which the determinants of a banking crisis could be, the paper focused on how to improve the performance of the logit model as Early Warning System for these events. In particular, it had the scope of shedding a brighter light on the so called *post-crisis* or *crisis duration bias* whose detrimental effect on the model predictivity, until now, has just been marginally investigated. With this regard, the results from the binomial comparison framework strongly confirmed the importance of the adoption of a *post-crisis bias* solution to boost the forecasting efficacy of the model. The presence of the bias and the validity of the proposed solution have been confirmed across several framework specifications (both binomial and multinomial) which differed in terms of country sample, dependent and explanatory variables. During the second empirical phase, I have elaborated an exercise that could return some hints on the relationship between the *post-crisis bias* and the duration of crisis event. To the best of my knowledge, this represented the first attempt to track down the features of the relationship between a logit EWS predictive quality and the duration of the crisis episodes. Keeping the AUC of each tested model as supreme measure for its forecasting power and by partitioning the full data sample in clusters on the basis of an average crisis duration variable, I have found that: the longer the crises under scrutiny last, the more the *post-crisis bias* negatively influences the model quality and, therefore, the more advisable the adoption of a dedicated solution should be.

In order to make this suggestion more viable to policy makers inclined to exploit it, I have regressed the average duration per country on a long series of institutional, economic and political indicators, crossing the boundary that divides two main research areas that float around systemic banking crises: one dedicated to the study of the determinants of a crisis and their implementations as early warning indicators, the other one involved in unveiling which forces can influence the severity of a default. Severity that in few cases is proxied by crisis duration variables, similarly, to some extent, to the exercise reported in this paper. The results describe a strong positive relationship between income per capita (GDP per capita) and the average crisis duration. Once a crisis occurs, economies characterized by above-average wealth per capita, healthy institutions, distributed financial sectors and elevate financial openness (low restrictions on cross-border financial transactions) are more prone to persist longer periods in crisis state. In view of these considerations, countries that fit with the characteristics just described, more than other, should strongly consider to adopt a *post-crisis bias* solution when implementing a logit EWS for systemic banking crisis. This decision would be crucial to tailor the model to their specific economy and sharpen the most its predictive accuracy.

However, these latest conclusions should be treated with caution and should be matter of further analysis to meet desirable reliability standards. The linear regression exercise, in particular, is new in its execution. Never before a crisis-duration variable has been regressed against such a number of static metrics which, until now, have proved their validity just during panel data based exercises. Most of these indicators proved validly correlated with crisis duration when tested around the crisis date along with their annual or monthly timing, while largely losing their relevance once averaged over the period.

To a smaller extent, other aspects of my work inevitably present some limitations. The subjective choice of the dependent variable is certainly one. I have tried to cope with this drawback by verifying my findings on two different banking crisis definitions (Rogoff and Reinhart, 2009; Laeven and Valencia, 2012). Despite the main results targeted by the paper are proven satisfactorily robust, the choice of banking crisis definition still detains a potentially deep impact on the output of the tests and, therefore, on the indications provided by the EWS. Future research shall firmly tackle this issue in order to clear the path towards a widely agreed and effective early warning framework. For what may regard the relationship between the quality of the EWS and the duration of the crises, similar tests on a larger sample, either in terms of number of countries and time-span, are required to further strengthen the statements achieved in this paper, allocating them the authority that the mild experience of the author in the field does not permit him to do.

Nonetheless, the author believes that his work has the merit to highlight a new path towards a higher understanding of these systemic phenomena, drawing the first empirical conclusions on the duration-EWS relationship and, thus, laying the foundations for its future thorough comprehension. The quest for the implementation of an exhaustive early warning framework is still widely open and the path towards its accomplishment is full of complexities that wait to be properly handled. In any case, the ultimate prize will definitely worth the effort.

Appendix

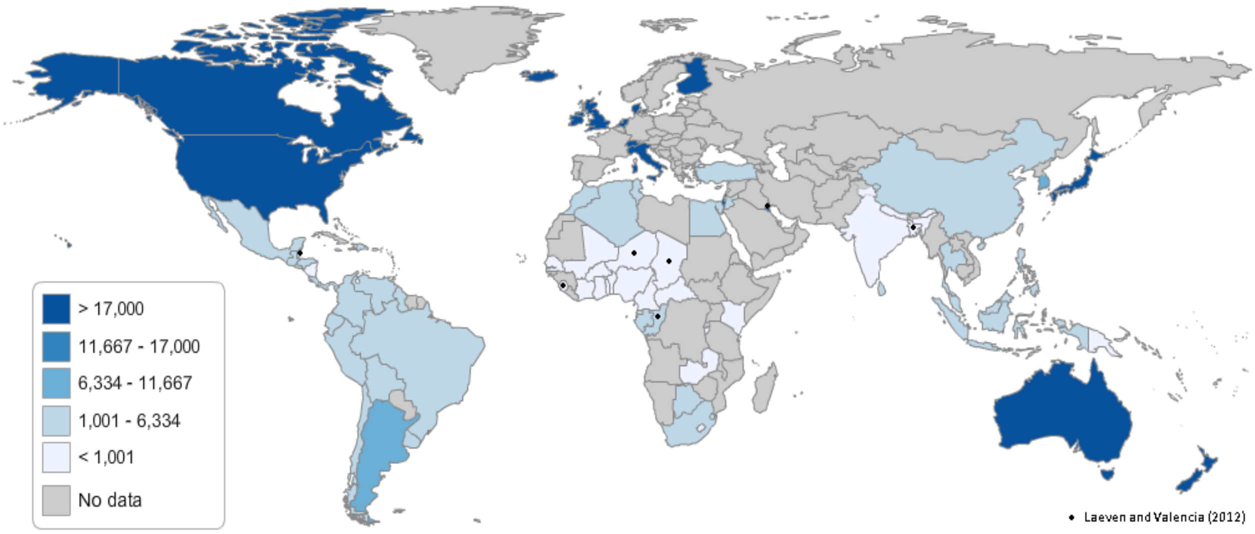
Appendix 1. Sample composition and crisis periods included in each data set. Countries are grouped on the basis of their GNI per capita as proposed by the World Bank (2014). Income Groups: Low Income (lightest blue; $1,045 or less), Low Middle Income (from $1,046 to $4,124), Upper Middle Income (from $4,126 to $12,175), High Income (darkest blue; $12,175 or more).

Country	**Rogoff & Reinhart (2009)**	**Laeven & Valencia (2012)**	**RS**
Benin	1988-1990	1988-1992	
Burkina Faso	1988-1994	1990-1994	
Burundi	1994-1995	1994-1998	
Central African Rep.	(1976)-1982, 1988-1999	1995-1996	
Chad*		1992-1996	
Gambia			
Mali	1987-1989	1987-1991	
Nepal	1988	1988	
Niger*		1983-1985	
Sierra Leone*		1990-1994	X
Togo	1993-1995	1993-1994	
Bangladesh*		1987	X
Bolivia	1986-1987, 1994-1996, 1999	1986, 1994	X
Cameroon	1987-1993, 1995-1998	1987-1991, 1995-1997	
Congo (Rep.)*		1992-1994	
Côte d'Ivoire	1988-1991	1988-1992	
Egypt	1981-1983, 1990-1995	1980	X
El Salvador	1989-1990	1989-1990	X
Ghana	1982-1989, 1997	1982-1983	X
Guatemala			
Guyana		1993	
Honduras			X
India		1993	X
Indonesia	1997-2002	1997-2001	X
Kenya	1985-1988, 1992-1995	1985, 1992-1994	X
Lesotho			
Morocco	1983-1984	1980-1984	X
Nicaragua	1987-1996	1990-1993, 2000-2001	
Nigeria	1992-1995, 1997, 2009-2013	1991-1995, 2009-2012	X
Papua New Guinea			
Philippines	1981-1987, 1997-2001	1983-1986, 1997-2001	X
Senegal	1988-1991	1988-1991	X
Sri Lanka	1989-1993	1989-1991	
Swaziland	1995	1995-1999	
Zambia	1995-1998	1995-1998	
Algeria	1990-1992	1990-1994	
Belize*			
Botswana			

Brazil	1985-1986, 1990-1991, 1994-1997	1990-1998	X
China	1997-1999	1998	X
Colombia	1982-1987, 1998-2000	1982, 1998-2000	X
Costa Rica	1987-1991, 1994-1996	1987-1991, 1994-1995	X
Dominica*			
Dominican Republic	2003-2004	2003-2004	X
Ecuador	1981, 1996-2002	1982-1986, 1998-2002	X
Gabon			
Jamaica	1995-2000	1996-1998	
Jordan		1989-1991	X
Malaysia	1985-1988, 1997-2001	1997-1999	X
Mauritius			
Mexico	1981-1982, 1993-1997	1981-1985, 1994-1996	X
Panama	1988-1989	1988-1989	X
Peru	1983-1990, 1999	1983	X
South Africa	1989		X
Thailand	1983-1987, 1996-2000	1983, 1997-2000	X
Tunisia	1991-1995	1991	
Turkey	1982-1985, 2000-2001	1982-1984, 2000-2001	X
Argentina	1980-1982, 1989-1990, 1995-1996, 2001-2003	1980-1982, 1989-1991, 1995, 2001-2003	X
Australia			X
Canada			X
Chile	1981-1984	1981-1985	X
Denmark	2008-2013	2008-2012	X
Finland	1991-1994	1991-1995	X
Iceland	2007-2013	2008-2012	X
Ireland	2007-2012	2008-2012	X
Israel	(1977)-1983		X
Italy	2008-2013	2008-2012	X
Japan	1992-2001	1997-2001	X
Korea, Republic	1985-1988, 1997-2000	1997-1998	X
Kuwait*		1982-1985	X
Netherlands	2008-2013	2008-2012	X
New Zealand	1987-1990		
Singapore			X
Switzerland	2008-2009	2008-2012	X
United Kingdom	2007-2013	2007-2012	X
United States	2007-2010	1988, 2007-2012	X
Uruguay	1981-1985, 2002-2005	1981-1985, 2002-2005	
Venezuela, RB	1993-1996, 2009-2010	1994-1998	X

*Countries excluded from the main dependent variable (Rogoff and Reinhart). RS: Reduced Sample.

Appendix 2. World-wide perspective of the countries involved in the exercises clustered by GDP per capita (in current US $). Dotted ones take part just to the tests employing Laeven and Valencia (2012) dating procedure.



Appendix 3. IIA hypothesis tests for multinomial logit. Hausman test (a) and SUE-based test (b).

a)

	Coefficients (b) partial	(B) full	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
l2gdp	.0199852	.0148814	.0051038	.0060069
linf	.0264092	.0165648	.0098444	.0049207
ldep	-.0158911	-.0065073	-.0093837	.0043361
lmtor	.0234174	.0231946	.0002228	.0016109
lcgdp	.008269	.0064023	.0018668	.0018751
lliq	.0068858	.0069794	-.0000936	.0002714
lnop	-.0054645	-.0070395	.001575	.0016007
ltot	.0098177	.0071496	.0026681	.0029566
_cons	-4.682749	-4.61265	-.0700989	.0774328

```
                 b = consistent under Ho and Ha; obtained from mlogit
  B = inconsistent under Ha, efficient under Ho; obtained from mlogit

Test:  Ho:  difference in coefficients not systematic

              chi2(9) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                      =        2.08
            Prob>chi2 =      0.9901
            (V_b-V_B is not positive definite)
```

b)

```
. test [full_1 = nosecond_1], cons

 ( 1)  [full_1]L2.GDPgrowth - [nosecond_1]L2.GDPgrowth = 0
 ( 2)  [full_1]L.Inflation - [nosecond_1]L.Inflation = 0
 ( 3)  [full_1]L.Depreciation - [nosecond_1]L.Depreciation = 0
 ( 4)  [full_1]L.CreditGDP - [nosecond_1]L.CreditGDP = 0
 ( 5)  [full_1]L.Liquidity - [nosecond_1]L.Liquidity = 0
 ( 6)  [full_1]L.TermsofTrade - [nosecond_1]L.TermsofTrade = 0
 ( 7)  [full_1]L.NOP - [nosecond_1]L.NOP = 0
 ( 8)  [full_1]L.M2Reserves - [nosecond_1]L.M2Reserves = 0
 ( 9)  [full_1]_cons - [nosecond_1]_cons = 0

           chi2(  9) =   12.09
         Prob > chi2 =    0.2081
```

```
. test [full_2 = nofirst_2], cons accum

 ( 1)  [full_1]L2.GDPgrowth - [nosecond_1]L2.GDPgrowth = 0
 ( 2)  [full_1]L.Inflation - [nosecond_1]L.Inflation = 0
 ( 3)  [full_1]L.Depreciation - [nosecond_1]L.Depreciation = 0
 ( 4)  [full_1]L.CreditGDP - [nosecond_1]L.CreditGDP = 0
 ( 5)  [full_1]L.Liquidity - [nosecond_1]L.Liquidity = 0
 ( 6)  [full_1]L.TermsofTrade - [nosecond_1]L.TermsofTrade = 0
 ( 7)  [full_1]L.NOP - [nosecond_1]L.NOP = 0
 ( 8)  [full_1]L.M2Reserves - [nosecond_1]L.M2Reserves = 0
 ( 9)  [full_1]_cons - [nosecond_1]_cons = 0
 (10)  [full_2]L2.GDPgrowth - [nofirst_2]L2.GDPgrowth = 0
 (11)  [full_2]L.Inflation - [nofirst_2]L.Inflation = 0
 (12)  [full_2]L.Depreciation - [nofirst_2]L.Depreciation = 0
 (13)  [full_2]L.CreditGDP - [nofirst_2]L.CreditGDP = 0
 (14)  [full_2]L.Liquidity - [nofirst_2]L.Liquidity = 0
 (15)  [full_2]L.TermsofTrade - [nofirst_2]L.TermsofTrade = 0
 (16)  [full_2]L.NOP - [nofirst_2]L.NOP = 0
 (17)  [full_2]L.M2Reserves - [nofirst_2]L.M2Reserves = 0
 (18)  [full_2]_cons - [nofirst_2]_cons = 0

           chi2( 18) =   23.60
         Prob > chi2 =    0.1686
```

Appendix 4. Data sources.

Variable	**Description**	**Source**
Logistic Regression Analysis		
SBCRR (dependent)	Discrete dependent variable which takes on value of 1 during banking crisis episodes and 0 otherwise. When the model is multinomial the crisis years other than the first are set at 2.	Rogoff and Reinhart (2010)
SBCLV (dependent)		Laeven and Valencia (2012)
GDP growth	Annual percentage growth rate of GDP at market prices based on constant local currency.	World Bank WDI
Inflation	Inflation as measured by the annual growth rate of the GDP implicit deflator.	World Bank WDI
Depreciation	Annual change in the official exchange rate determined by national authorities or to the rate determined in the legally sanctioned exchange market.	World Bank WDI
M2-to-Reserves	Ratio of money and quasi money M2 to total reserves under the control of monetary authorities.	World Bank WDI
Credit-to-GDP growth	Annual growth of the domestic credit provided by financial sector as a percentage of GDP.	World Bank WDI
Liquidity	Financial resources provided to the private sector by domestic money banks as a share of total deposits.	World Bank Global Financial Development Database
Net Open Position	Ratio of net foreign assets to GDP.	World Bank WDI
Terms of Trade change	Annual growth rate of the ratio between exports and Imports.	World Bank WDI
Leverage	Ratio of bank capital and reserves to total assets.	World Bank Global Financial Development Database
Real Interest Rate	Lending interest rate adjusted for inflation as measured by the GDP deflator.	World Bank WDI; Oxford Economics (Datastream)
Export Concentration	Herfindahl-Hirschman index of concentration of merchandise import and export on a country base (it ranges from 0 to 1).	UNCTAD (2015)
Kaopen	Known as the Chinn-Ito index is an indicator of a country's degree of capital account openness.	Chinn, Menzie D. and Hiro Ito (2006)
Conflicts	Dummy variable that takes value 1 in case of direct involvement in an armed conflict (a minimum of 25 battle-related deaths per year) and 0 otherwise.	UCDP/PRIO Armed Conflict Dataset (2015)

Explicit	Dummy variable that takes the value of 1 to highlight the presence of an explicit deposit insurance scheme.	Deposit Insurance Database (2013)
Linear Regression Analysis		
GDPmean	GDP per capita in US $ (current value)	World Bank WDI
Latitude	Ranges from 0 to 1 (south to north)	Beck, Demirgüç-Kunt, Levine (2000)
OECD	Binary	OECD
Conc_Beck	Fraction of assets held by the three largest banks.	Beck, Demirgüç-Kunt, Levine (2000)
Conc_Caprio	Share of deposits held by the five largest banks.	Barth, Caprio, and Levine (2001) - Survey of Bank Regulation and Supervision
Conc_3	Fraction of assets held by the three largest banks.	BankScope
Conc_5	Fraction of assets held by the three largest banks.	BankScope
RCC	Total credit booms years per country based on the variation of real credit per capita from its long term trend.	Marco Arena, Serpil Bouza, Era Dabla-Norris, Kerstin Gerling, and Lamin Njie (2015)
Qcluster	Weighted means of economic, political and legal institutional quality indexes (1-5)	Kunčič, A. (2014). Institutional Quality Dataset, Journal of Institutional Economics
Gov_Quality	Voice and accountability, political stability, government effectiveness, regulatory quality, rule of law, and corruption. From 0 to 1, from worse to better governance.	Kaufman, Kraay and Zoido-Lobaton (1999)
frac_ent	Number of entry applications denied as a fraction of the total number of applications received from domestic and foreign entities.	Barth, Caprio, and Levine (2001) - Survey of Bank Regulation and Supervision
restrict	Sum of four measures that indicate the degree of restriction of bank activities in the securities, insurance and real estate markets and ownership and control of nonfinancial firms. From 1 (unrestricted) to 4 (prohibited).	Barth, Caprio, and Levine (2001) - Survey of Bank Regulation and Supervision
rr	Ratio of reserves required to be held by banks.	Barth, Caprio, and Levine (2001) - Survey of Bank Regulation and Supervision
bfree	Indicator of relative openness of banking and financial system.	Index of Economic Freedom (Heritage Foundation)
KAOPEN	Known as the Chinn-Ito index is an indicator of a country's degree of capital account openness.	Chinn, Menzie D. and Hiro Ito (2006)
BCreditGDP	Ratio of domestic credit to private sector provided by banks as a percentage of GDP.	World Bank WDI
HHI	The Herfindahl-Hirschmann Index (Product HHI), is a measure of the degree of product concentration on a country basis.	UNCTAD Statistics
CurrencyCount	Total years of currency crisis experienced by a country during the period 1980-2010.	Rogoff and Reinhart (2010)

SovereignCount	Total years of sovereign debt crisis experienced by a country during the period 1980-2010.	Rogoff and Reinhart (2010)
ConflictCount	Total years of armed conflict experienced by the country.	UNCDP/PRIO Armed Conflict Dataset
Foreign	Percentage of banking system assets in banks that are 50 % or more foreign owned.	Barth, Caprio, and Levine (2001) - Survey of Bank Regulation and Supervision
State	Percentage of banking system assets in banks that are 50 % or more state owned.	Barth, Caprio, and Levine (2001) - Survey of Bank Regulation and Supervision
Muslim	Variables that capture the percentage of population that is Muslim.	Beck, Demirgüç-Kunt, Levine (2000)
Catholic	Variables that capture the percentage of population that is Catholic.	Beck, Demirgüç-Kunt, Levine (2000)

References

Abiad A., Balakrishnan R., Brooks P, Leigh D., Tytell I., 2009, "What's the damage? Medium-term output dynamics after banking crises", IMF Working Paper No. 09/245.

Alessi L. et al., 2015, "Comparing different early warning systems: Results from a horse race competition among members of the Macro-prudential Research Network", Munich Personal RePEc Archive, Paper No.62194.

Allen F., Gale D., 2000, "Financial Contagion", The Journal of Political Economy, Vol. 108, pp. 1-33.

Arena M., Bouza S., Dabla-Norris E., Gerling K., Njie L., 2015, " Credit Booms and Macroeconomic Dynamics: Stylized Facts and Lessons for Low-Income Countries", IMF Working Paper No. 15/11.

Arteta C., Eichengreen B., 2000, "Banking Crises in Emerging Markets: Presumptions and Evidence", Center for International and Development Economics Research, UC Berkeley.

Barth J.R., Caprio G., Levine R., 2004, "Bank Regulation and Supervision: What Works Best?", Journal of Financial Intermediation, Vol. 13, No. 2, pp. 205-48.

Beck T., Demirgüç-Kunt A., Levine R., 2000, "A New Database on Financial Development and Structure", *World Bank Economic Review* 14, pp. 597-605.

Beck T., Demirgüç-Kunt A., Levine R., 2003, "Bank Concentration and Crises", National Bureau of Economic Research, Working Paper No.9921.

Berg A., 1999, "The Asian crisis: causes, policy responses, and outcomes", IMF Working Paper No. 99/138.

Boissay F., Collard F., Smets F., 2013, "Booms and Systemic Banking Crises", European Central Bank, Working Paper No. 1514.

Boyd J.H., De Nicoló G., 2003, "Bank risk-taking and competition revisited", IMF Working Paper No. 03/114.

Boyd J.H., De Nicoló G., Loukoianova E., 2009, "Bank Crises and Crisis Dating: Theory and Evidence", IMF Working Paper No. 09/141.

Borio C., Lowe P., 2002, "Assessing the risk of banking crises", BIS Quarterly Review (December).

Bussiere M., Fratzscher M., 2006, "Towards a new early warning system of financial crises", Journal of International Money and Finance, No. 25, pp. 953–973.

Caggiano G., Calice P., Leonida L., 2014, "Early warning systems and systemic banking crises in low income countries: A multinomial approach", Journal of Banking and Finance, No. 47, pp.258-269.

Calomiris C. W., 1989, "Deposit Insurance: Lessons from the Record", Federal Reserve Bank of Chicago, Economic Perspectives pp. 10-30.

Caprio G., Klingebiel D., 1997, "Bank Insolvency: Bad Luck, Bad Policy or Bad Banking?", The International Bank for Reconstruction and Development, World Bank.

Caprio G., Klingebiel D., 2003, "Episodes of Systemic and Borderline Financial Crises", World Bank.

Caprio G., Martinez-Peria M.S., 2000, "Avoiding Disaster: Policies to Reduce the Risk of Banking Crises", Egyptian Center for Economic Studies.

Chaudron R, de Haan J., 2014, "Identifying and dating systemic banking crises using incidence and size of bank failures", De Nederlandsche Bank, Working Paper No. 406.

Chinn M. D., Hiro I., 2006, "What Matters for Financial Development? Capital Controls, Institutions, and Interactions", Journal of Development Economics, Volume 81, Issue 1, pp. 163-192 (October).

Čihák M., Schaeck K., 2007, "How Well Do Aggregate Bank Ratios Identify Banking Problems?", IMF Working Paper, No. 07/275.

Davis E.P., Karim D., 2008, "Comparing early warning systems for banking crises", Journal of Financial Stability, No.4, pp. 89-120.

Davis E.P., Karim D., Liadze I., 2010, "Banking Crisis in Asia and Latin America – A Single Pattern for Emerging Market Economies?", NIESL and Brunel University (London), Working Paper No. 9-09.

Demirgüç-Kunt A., Detragiache E., 1997, “The Determinants of Banking Crises: Evidence from Industrial and Developing Countries”, Policy Research Working Paper No.1828, The World Bank and International Monetary Fund.

Demirgüç-Kunt A., Detragiache E., 2000, “Monitoring Banking Sector Fragility: A Multivariate Logit Approach”, World Bank Economic Review, Vol. 14, No. 2, pp. 287–307.

Demirgüç-Kunt A., Detragiache E., 2002, “Does deposit insurance increase banking system stability? An empirical investigation”, Journal of Monetary Economics, No. 49, pp. 1373-1406.

Demirgüç-Kunt A., Detragiache E., 2005, “Cross-Country Empirical Studies of Systemic Bank Distress: A Survey”, IMF Working Paper No.05/96.

Demirgüç-Kunt A., Levine R., Hong-Ghi M., 1998, “Opening to Foreign Banks: Issues of Stability, Efficiency and Growth”, in The Implications of Globalization of World Financial Markets, Bank of Korea.

Demirgüç-Kunt A., Kane E., Laeven L., 2013, "Deposit Insurance Database", Policy Research Working Paper No. 6934, World Bank.

Diamond D., Dybvig P., 1983, “Bank Runs, Deposit Insurance, and Liquidity”, The Journal of Political Economy, Vol. 91, No. 3, pp. 401-419.

Eichengreen B., Rose A.K., 1998, “Contagious Currency Crises: Channels of Conveyance”, forthcoming in Changes in Exchange Rates in Rapidly Developing Countries.

Ergungor O.E., Thomson J.B., 2005, “Systemic Banking Crises”, Federal Reserve of Cleveland.

Friedman M., Schwartz A.J., 1963, “A Monetary History of the United States, 1867–1960”, Princeton University Press for NBER.

Gaytán A., Johnson C.A, 2002, “A Review of the Literature on Early Warning Systems for Banking Crises”, Central Bank of Chile, Working Paper No. 183.

Gavin M., Hausmann R., 1996, “The Roots of Banking Crises: The Macroeconomic Context”, Inter-American Development Bank, Working Paper No. 318.

González-Hermosillo B., 1999, “Determinants of Ex-Ante Banking System Distress: A Macro-Micro Empirical Exploration of Some Recent Episodes”, IMF Working Paper No. 99/33.

Gorton G., 1988, “Banking Panics and Business Cycles”, Oxford Economic Papers, New Series Vol. 40, No. 4, pp. 751-781.

Hardy C.D., Pazarbaşioĝlu C., 1999, “Determinants and Leading Indicators of Banking Crises: Further Evidence”, IMF Staff Papers, Vol. 46, No. 3.

Hausman J. A., McFadden D. L., 1984, “Specification tests for the multinomial logit model”, Econometrica 52, pp. 1219–1240.

Honohan P., 1997, "Banking System Failures in Developing and Transition Countries: Diagnosis and Prediction", Bank for International Settlements, Working Paper No.39.

Kaminsky G.L., Graciela L., 1998, “Currency and Banking Crises: A Composite Leading Indicator”, International Finance Discussion Paper Series No. 629, Washington: Board of Governors of the Federal Reserve.

Kaminsky G.L., Reinhart C.M., 1996, “The Twin Crises: The Causes of Banking and Balance of Payments Problems”, International Finance Discussion Paper Series No. 544, Washington: Board of Governors of the Federal Reserve.

Kaminsky G.L., Reinhart C.M., 1999, “The Twin Crises: The Causes of Banking and Balance of Payments Problems”, American Economic Review, Vol. 89, pp. 473–500.

Kaminsky G.L., Goldstein M., Reinhart C.M., 2000, “Early Warning Systems: Empirical Results from the Signal Approach”, Munich Personal RePEc Archive, Paper No. 24577.

Kane E.J., 1989, “The S&L Insurance Mess: How Did It Happen?”, The Journal of Finance, Vol. 44, No. 5.

Kaufman, D., Kraay, A., Lobaton, P.Z., 1999, "Governance Matters", World Bank Policy Research Department Working Paper No. 2196.

Kindleberger P., 1978, “Manias, Panics, and Crashes: A History of Financial Crises”, Basic Books, New York.

Kunčič A., 2014, “Institutional Quality Dataset”, Journal of Institutional Economics.

Laeven L., Valencia F., 2012, “Systemic banking crises database: an update”, IMF Working Paper WP/12/163.

Lambregts E., Ottens D., 2006, “The Roots of Banking Crises in Emerging Market Economies: a panel data approach”, De Nederlandsche Bank, Working Paper No. 84.

Lainà P., Nyholm J., Sarlin P., 2014, "Leading indicators of systemic banking crises: Finland in a panel of EU countries", Bank of Finland Research Discussion Papers No. 14.
Mehrez G., Kaufmann D., 1999, "Transparency, Liberalization and Financial Crises", World Bank.
Mendis C., 1998, "External Shocks and Banking Crises in Small Open Economies: Does the Exchange Rate Regime Matter?", Center for the Study of African Economies (Oxford).
Navajas M.C., Thegeya A., 2013, "Financial Soundness Indicators and Banking Crises", IMF Working Paper WP/13/263.
Pill H., Pradhan M., 1995, "Financial Indicators and Financial Change in Africa and Asia", IMF Working Paper No. 23.
Reinhart C.M., Rogoff K.S., 2008, "Banking Crises: An Equal Opportunity Menace", National Bureau of Economic Research, Working Paper No. 14587.
Reinhart C. M., Rogoff K.S., 2009, "This Time Is Different: Eight Centuries of Financial Folly", Princeton University Press.
Reinhart C. M., Rogoff K.S., 2010, "From Financial Crash to Debt Crisis", NBER Working Paper 15795, Forthcoming in American Economic Review.
Rojas-Suarez L., 1998, "Early Warning Indicators of Banking Crises: What Works for Emerging Markets?", Inter-American Development Bank.
Rojas-Suarez L., Weisbrod S.R., 1996, "Banking Crises in Latin America: Experience and Issues", Inter-American Development Bank, Working Paper No.265.
Van den Berg J., Candelon B., Urbain J.P., 2008, "A cautious note on the use of panel models to predict financial crises", Economics Letters 101, pp. 80–83.
Wilms P., Swank J., de Haan J., 2014, "Determinants of the real impact of banking crises: A review and new evidence", De Nederlandsche Bank, Working Paper No. 437.
Yanagitsubo M., 2004, "Determinants affecting severity in financial crises: a successful approach for reintegration into the international financial markets", Journal of International Development Studies 13, pp. 31-49.