

University of Neuchâtel

Institute of Economic Research

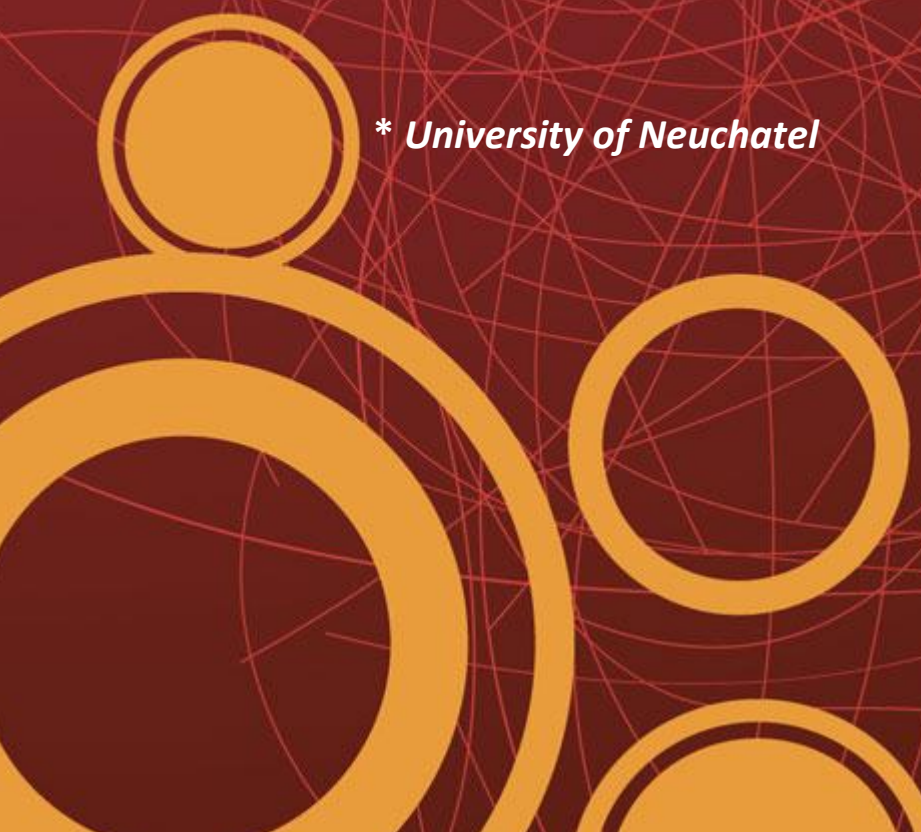
IRENE, Working paper 13-02



Optimal Weights and Stress Banking Indexes

*Stefano Puddu**

** University of Neuchâtel*



unine

UNIVERSITÉ DE
NEUCHÂTEL

Institut de
recherches économiques

Optimal Weights and Stress Banking Indexes*

Stefano Puddu[†]

This Version: **March 2012**

Abstract

The goal of this paper is to provide alternative approaches to generate indexes in order to assess banking distress. Specifically, we focus on two groups of indexes that are based on the signalling approach and on the zero inflated Poisson models. The results show that the indexes based on these approaches perform better than those constructed by using the variance-equal and the factor analysis methods. Specifically, they are better at capturing relevant events, signalling distress episodes and forecasting properties. The importance of this study is two-fold: first, we contribute extra information that can be useful for forecasting banking system soundness in the aim of preventing future financial crises; second we provide alternative methods for measuring banking distress.

Keywords Stress-banking indexes, Signalling approach, Limited dependent variable methods **JEL Classification** C16, C25, G21, G33, G34

*The author is grateful to Luigi Infante, Pierre Monnin, Florian Pelgrin and Pascal St. Amour for their comments, and would like to thank the participants at the Research Days 2008 seminars at HEC-Lausanne, and at the XXXIII Simposio de Análisis Económico -Zaragoza 2008-.

[†]HEC Lausanne, Switzerland, e-mail: stefano.puddu@unine.ch

1 Introduction

The financial turmoil that affected the world economy in the last five years called for a more precise view and more accurate analysis of the quality of the banking sector. The features of the financial crisis, how rapidly it propagated, and its dimensions increase the importance of synthetic distress banking indicators. Their relevance in assessing banking distress levels is three-fold. banking indexes are useful because they provide a quick picture of the soundness of the banking sector. Moreover, they could be employed in forecasting exercises to measure future banking distress levels, allowing the monetary authorities to respond in a timely manner to future potential financial troubles. Finally, they could be a useful tool for verifying the effects of policy implementations, such as the capital requirement rules. Within other contexts apart from the banking system, it is also common for financial analysts and economist to use an index instead of only one variable to describe a particular economic phenomenon. The Consumer Confidence Index, the Dow Jones and the Consumer Price Index are just three examples of the extensive application of this particular tool in economic and financial analyses.

The two main methods employed thus far to compute indexes are the variance-equal (VE) approach and the factor analysis (FA) approach. However, both methods come with several drawbacks and limitations. The VE ascribes to all the variables the same weight without taking into account any prior economic knowledge. The FA weights are sensitive to the existence of missing values and to the set of variables employed. Moreover, it may generate multiple solutions, so that this method cannot always be employed.

The aim of this study is to provide alternative approaches for creating indexes in order to measure banking distress levels. Specifically, we focus on two groups of indexes: the first group is a modification of the signalling approach (MSA) employed by Kaminsky and Reinhart (1999)¹, and the second has been constructed by using zero inflated Poisson estimations.

¹Goldstein, Kaminsky and Reinhart (2000) provide an exhaustive analysis of the signalling approach.

The MSA exploits the ability of the variables of interest to identify crisis periods and to avoid false alarms. The higher the precision of the variables in detecting these two states is, the larger the weight that they receive. The second approach exploits zero inflated Poisson estimations: the estimated values as well as the correspondent standard errors are employed to generate a set weights to ascribe to the variables included in the regression analysis. We compare the properties of these groups of indexes with those of the indexes based on the variance-equal and factor analyses by taking into account three dimensions: their ability to capture specific events of interest and to signal distress episodes; the stability of the weights ascribed to the variables depending on the period taken into account; and their forecasting features.

The results show that the best indexes are those based on the zero inflated regressions. They perform better at capturing relevant events and signalling distress episodes, and they exhibit the best forecasting properties. Finally, the results also highlight that the VE, the FA and the MSA indexes show similar patterns both in terms of shape and forecasting properties.

This study aims to use these results to accomplish two goals. We wish to contribute to the provision of extra information that analysts can use to forecast banking system soundness and to measure future banking distress. Moreover, this study provides alternative methods for computing banking distress indexes.

The rest of the paper is organized as follows: in section 2, we discuss the previous studies on the same topic. In section 3, we analyse the dataset, while in section 4, we describe the methodologies employed to generate optimal weights. In section 5, we show the results and we compare them with the findings referring to the VE and the FA approaches. Finally, section 6 concludes.

2 Literature Review

There exist different ways to construct synthetic indexes. The variance-equal weights approach ascribes the same weight to every variable within a given set. This feature is, at the same time, its main advantage and caveat. Specifically, the variance-equal approach implies that all of the variables are equally important, even if this is not always true. Hanschel and Monnin (2005) generate an index for the Swiss economy, based on this approach, by merging different types of variables. Specifically, they focus on market prices, aggregate balance-sheet data, non-public information and other structural data. Banking distress values larger than zero imply that the banking sector is experiencing a level of distress larger than the average. Data refers to the period between 1987 and 2002, for an overall of 16 annual observations. The index identifies three periods where banking distress level is above average. These periods correspond to Swiss economic downturns. Their index is therefore able to fit the main economic events of the Swiss economy for the period analysed.

Alternatively, the factor analysis exploits the total variance generated by the variables of interest. In this way, it is possible to extract a sequence of weighted linear combinations among them. The first vector of weights explains the majority of the common variance. Thereafter, if some unexplained variance is left, a second vector of weights is computed. These computations continue until all of the variance is explained. As documented by Hanschel and Monnin (2005), one limitation of this approach is that it can be used only if all the variables react to the same set of shocks. Factor analysis can also be sensitive to the existence of missing values, and it is not recommended if only a small number of variables are taken into account. Moreover, the factor analysis may lead to multiple solutions, giving the results no economic meaning. Illing and Liu (2003) use the variance-equal method and the factor analysis approach to generate financial distress indexes for the Canadian financial system. They focus on banking variables, and indicators referring to debt, equity and foreign exchange rates. The period covered goes from 1979 to 2003. The observation are quarterly based. They

find that their indexes fit the financial and economic events of the Canadian economy.

Morales and Estrada (2010) construct a continuous distress index for the Colombian banking sector by merging information related to banks' profitability, liquidity and their probability of default. The period analysed ranges from January 1995 to November 2008, observations are monthly based, for an overall of 167 periods. The weights are ascribed by using the variance-equal weight, principal components and general count data models². The findings suggest that the index is able to correctly report the level of stress in the banking sector.

Using a sample period between 1995 and 2011, for an overall of 17 yearly observations, for the German banking system, Jahn and Kick (2012) build an index for measuring individual bank distress by merging individual probability of default, credit spreads and a measure reflecting the value of the banking sector in the stock market. The approach employed is based on a partial proportional odds model (PPOM) that uses as a dependent variable banks risk profile³. The weakness of this approach is that only a limited amount of potential weights combinations can be tested (Jahn and Kick only test 36 combinations).

The approaches employed in previous contributions highlight several weaknesses. In this paper, we propose two different methods, which are able to overcome some limitations of the above mentioned methods. On the one hand, the MSA combines all the information provided by a set of variables by taking into account the ability of the variables to identify crisis periods and to avoid false alarms. On the other hand, the zero inflated Poisson estimations exploits the estimated values as well as the correspondent standard errors to generate a set weights to ascribe to the variables included in the regression analysis.

The methods presented in this paper have different elements of novelty with respect to those employed in previous contributions.

²The main approaches employed in Morales and Estrada (2010) have been at least partially borrowed from a previous version of this paper.

³The risk profile is ranked in four categories A, B, C and D, where A is the best and D the worst grading.

First, we proxy banking crises by exploiting the relationship between bank failures and the quality of the banking sector. The MSA employs the number of bank failures to define crisis and non-crisis periods, whereas the zero inflated Poisson model uses the number of failures as a dependent variable. Although this hypothesis is consistent with that of Kaminsky and Reinhart (1999), who focus on closures, mergers or takeovers due to bank runs, other studies adopt different definitions for qualifying a banking crisis. For instance, Illing and Liu (2003) define the stress period as “the force exerted on economic agents by uncertainty and changing expectations of loss in financial markets and institutions. Moreover, Demirg-Kunt and Detragiache (1998, 1999) claim that a banking system is experiencing a crisis if at least one of the following situations happen: the ratio of non-performing assets to total assets is larger than the 10%; the cost of the rescue operation is larger or equal than 2% of GDP; a large number of bank runs or government emergency measures occur as a consequence of a crisis; a large-scale nationalization as a consequence of banking sector problems. Finally, Jahn and Kick (2012) employ banks’ risk profile to identify banks under financial distress.

Second, the MSA solve two main weaknesses of the traditional signalling approach. On the one hand, the traditional signalling approach wastes a relevant part of the information available. Specifically, despite a large amount of information that the method can potentially employ, only a marginal amount is effectively employed because, depending on predefined criteria, the only variable chosen as an indicator of banking distress is the most precise. The MSA solves the previous weakness by ascribing weights to the variables in accordance with their relative precision in detecting crisis periods, employing all the information provided by the variables. On the other hand, the traditionally signalling approach, employs exogenous thresholds to define which of the variables is the best in predicting banking crises. In the MSA the threshold are endogenously determined, once a definition of crisis is chosen.

Third, in the determination of the weights, alternatively to other contributions, we do not impose any ex-ante knowledge about the sign to ascribe to the variables or the threshold level

to define crisis periods. Alternatively, we let the data “talk”. This is true when the MSA as well as the zero inflated Poisson estimations are employed. In this way, we do not incur in the problems of Jahn and Kick (2012), in which only a limited set of weights combinations is tested, or in those related to the variance-equal approach, in which all the variables receive the same weights and the indicators are merged by using a priori economic knowledge.

Finally, the variables employed to construct the indexes strictly refer to the banking sector, disregarding those variables that refer to the stock exchange market or to the macroeconomic system so that potential spurious effects are neutralized.

3 The Dataset

Our dataset is based on quarterly US-level data and covers the period from 1984 to 2007, with 95 observations overall⁴. We employ information referring to the aggregate commercial banking sector, and the dataset has been generated by collecting information from the Federal Reserve of St. Louis and from the Federal Deposit Insurance Corporation (FDIC)⁵.

The dataset includes six variables: the return on assets (ROA), the net interest margin (NIM), the ratio of net loan losses to average total loans (LSTL), the ratio of non-performing loans to total loans (NPTL), the ratio of loan loss reserve to total loans (LLRTL) and finally, the number of the commercial banks failed (FAILS). All the variables except for FAILS have been normalized to standard normal distributions with zero mean and unit standard deviation⁶.

ROA refers to the profitability of the banks: a low level of ROA should be a signal for a low level of profitability in the banking sector. The lower the ROA is, the larger the banking distress level. NIM is a proxy for bank profits; however, its impact on banking distress is

⁴In Figure 8, Section D of the Appendix, we report the indexes based on updated version of the dataset, expanded until 2011:Q3. The results do not change.

⁵The complete description of the variables is provided in Table 3, Section B of the Appendix.

⁶Table 4, Section B of the Appendix, reports the main descriptive statistics of the variables included in our dataset before proceeding to standardization.

ambiguous⁷. On the one hand, a high NIM implies a high level of profits, and therefore a low level of banking distress. Low level of competition in the banking sector may increase NIM and therefore the level of banking distress decreases⁸. On the other hand, the net interest margin may be also affected by capital requirements enforced by the regulatory authorities. More demanding capital requirements impact NIM negatively, but at the same time, they make the banking system more robust to financial distress. The ratio of net loan losses to average total loans (LSTL) and the ratio of non-performing loans to total loans (NPTL) refer to the current quality of the banking system: the higher the LSTL and the NPTL are, the larger the banking distress. The loan loss reserve over total loans (LLRTL) is a forward looking proxy of the quality of the banking sector. Banks have to set aside capital for precautionary reasons in order to counterbalance potential future loan losses. In several studies⁹, it has been shown that LLRTL has a countercyclical behaviour, amplifying economic boom and bust periods. This implies that LLRTL increases during recessions. Increase in the LLRTL, may lead to a reduction of the credit available for firms, that, in turn, affects the level of output and thus the worthiness of the borrowers. In this case, banking distress may increase.

Previous intuitions are indirectly confirmed by looking at the pairwise correlation matrix reported in Table 4, Section B of the Appendix. Specifically, the results highlight a negative correlation between the quarterly number of bank failures (the proxy employed for identifying banks distress) and ROA (-.77), while it shows a positive correlation with the rest of the series. The correlations values with LSTL, NPTL and LLRTL are quite high (.49, .8 and .486, respectively), while the correlation coefficient with NIM is around .15.

Moreover, the results show a strong negative correlation between the ROA and the variables

⁷See Ho and Saunders (1981), Allen (1988), Zarruk and Madura (1992), Wong (1996), Saunders and Schumacher (2000), Maudos and Fernandez de Guevara (2007).

⁸Competition in banking sector may force banks to decrease lending standards so that the banking distress may increase.

⁹See, among others, Marcucci and Quagliariello (2008) and Balla and McKella (2009).

that measure the quality of the banking system: LSTL (-.47), NPTL (-.85) and LLRTL (-.42). ROA exhibits a negative correlation (-.15) with the net interest margin. NIM is positively correlated with all the other variables. Net interest margin's correlation is relatively weak with respect to LSTL (.26) and to NPTL (.33), while it is stronger with respect to LLRTL (around .70). The correlations between the loan losses over average total loans and LLRTL and NPTL are .63 and .70, respectively. Finally, LLRTL and NPTL show strong positive correlation, reporting a coefficient around .60.

In our dataset, we include in addition the quarterly absolute number of banking failures, FAIL. It is a proxy of the financial banking distress. Specifically, the higher the number of banking failures is, the lower the quality in the banking sector and, therefore, the larger the index of banking distress. The maximum quarterly number of failures is 99, and the peak is scored in 1988:Q2. Quarters with zero failures, the mode of the series, are observed 25 times. The yearly number of the bank failures is recorded in 1988, with 280 overall failures. Moreover, during the period between 1984 and 1993, the yearly failures are always larger than 40. Finally, since 1994, failures drastically decreases to 11 failures per year at most. The quarterly number of bank failures is employed as a benchmark in the baseline analysis of this study. Moreover, alternative proxies of banking distress have been employed in the robustness checks leading to similar results.

Figure 1.a shows the quarterly number of banking failures together with a measure that takes into account at the same time the quarterly fraction of banks that failed and its asset volume counterpart. The banks and assets failed measure (BAF) increases in both terms. The figure shows that FAILS and BAF perfectly overlap, after a rescaling process. Therefore, from a qualitative view point, the two series provide the same information, even if the quarterly number of failures allows us to employ econometric techniques that otherwise could not be used¹⁰.

¹⁰For a qualitative comparison of the indexes obtained using the two measures, see Table 5, Section B of the Appendix.

Figure 1: Failures

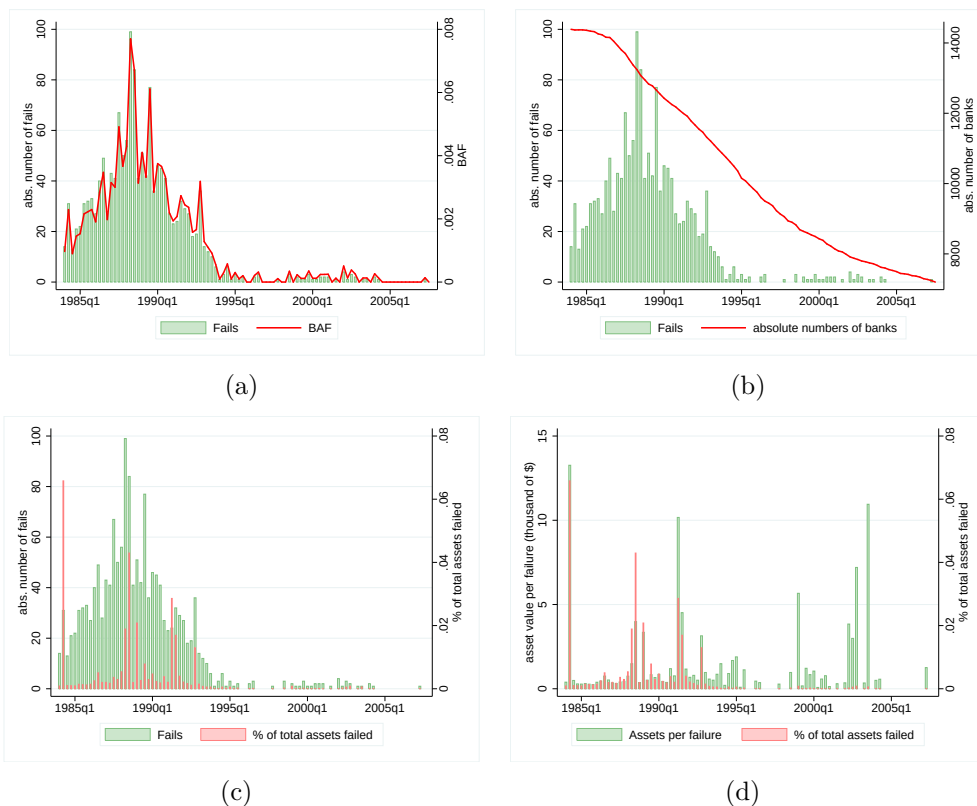


Figure 1.b reports the quarterly number of bank failures together with the evolution of the number of banks over time. At the beginning of the period, the commercial banks in the US numbered more than 14000, while at the end of the period analysed, this number decreased to less than 8000. The pattern of the banking failure seems not to be affected by the evolution of the number of banks over time. Figure 1.c displays the quarterly number of bank failures together with the correspondent value percentage of the failed assets over total assets. Apart from the outlier before 1985, there is a good fit between the two series. Finally, Figure 1.d shows the behaviour of the value percentage of failed assets to total assets and the per failed bank assets value. High values of the latter could be due to a small number of big banks (in terms of assets value) involved.

4 The Methodologies

This section studies the modified signalling approach and the zero inflated method, by the analysis of their strengths and potential drawbacks.

4.1 Modified signalling approach

The modified signalling approach is an improvement of the method employed by Kaminsky (1998) and Kaminsky and Reinhart (1999). This method is based on the ability of a particular indicator to detect crisis periods and to distinguish them from non-crisis periods. Crisis and non-crisis periods are defined by endogenous threshold levels chosen such that the mistakes produced by each variable in detecting the two periods are minimized. The main weakness of the signalling approach is that despite a large amount of information that the method can potentially employ, only a marginal amount is effectively used. This is due to the fact only the variable that minimizes the errors of non detecting a crisis or giving a false alarm is selected. The modified signalling approach attempts to fill this gap.

Specifically, by using the number of banking failures as a proxy for banking distress to establish crisis periods, we define different failures benchmarks as reported in Table 5, Section B of the Appendix. We focus on three main cases. Based on the first criterion, F_0 , a quarter is classified as “crisis quarter” if there is at least one bank fail. According to the other criteria, $F_{10}(F_{40})$, all the quarters of a year in which there are at least ten(forty) banks failures are classified as “crisis quarters”.

According to these criteria, the average values of the variables of interest, together with their correspondent signs during the two periods, are computed. Therefore, for each point in time, it is possible to assess the ability of the indicators to detect the two periods of interest, depending on their position with respect to their conditional means. Specifically, for each indicator, we focus on the mistakes generated by not detecting a crisis when a crisis occurs

(the values in absolute value are smaller than the correspondent conditional mean) as well as the mistakes generated by signalling a false alarm (the values in absolute value are bigger than the correspondent conditional mean). By combining the two types of mistakes, a proxy of the precision of each indicator is computed. Specifically,

$$m_i = \frac{1 - P(no - alarm|crisis)}{1 + P(alarm|no - crisis)} \in [0, 1] \quad (1)$$

The precision of the indicator decreases as the two types of mistakes increase. If no mistakes are produced, the precision is equal to the unit. If $P(no - alarm|crisis) = 1$, then the precision is zero, while it could be positive if $P(alarm|no - crisis) = 1$. Therefore, this measure gives more weight to the mistake of not detecting a crisis when there is, indeed, a crisis.

The weights are ascribed such that the higher the relative precision of an indicator is, the higher its weight. It follows that the weights are a function of the criterion that identifies crisis and non-crisis periods, the precision of the variables in detecting the two periods of interest as well as the precision of the other indicators. Specifically:

$$w_i = f(Criterion, m_i, m_{\neq i}) \quad (2)$$

The signs ascribed to the variables are those associated with their conditional means during the crisis period.

Our approach shows some differences with respect to methodology by Kaminsky (1999) and Kaminsky and Reinhart (1999). First, the MSA uses the number of banking failures as a benchmark in order to define crisis periods, while previous contributions define banking crises according to the intervention of the government in financial institutions' bailouts.

Second, we define exogenously the minimum number of failures for identifying a crisis, and based on these benchmarks, we identify specific cut-off points. Contrary to our approach, previous contributions define several cut-off points and then chose the threshold level that maximizes the precision of the indicator.

Finally, the most relevant difference arises from the fact that we employ all the information provided by the analysis of the ability of the variables to detect particular events. This is not the case in previous contributions, in which each variable is defined as a good or a bad instrument for predicting crises depending on its precision.

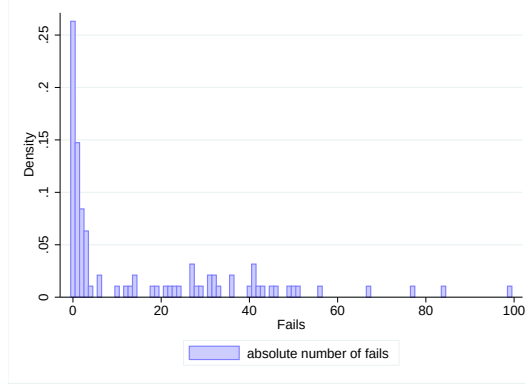
Even if the modified signalling method is intuitive and easy to compute, it has at least two important drawbacks that can affect the results. First, the MSA is unable to measure the magnitude of the errors: it generates the same type of error regardless of the distance of the value with respect to the threshold. Second, the MSA is not able to take into account potential “contagion effects” across the variables; or in other words, it is an approach based on a partial equilibrium analysis. Each indicator affects banking distress in two different ways. On the one hand, there exists a direct impact. On the other hand, there is an indirect impact, depending on the indicator’s ability to affect the position of the other variables in the same period. It follows that if the indicator i has more influence on other variables than the indicator j , then the former should have more relative importance. Unfortunately, the MSA does not measure these effects.

4.2 Econometric approach

Another way to proceed for computing banking stress indexes is to exploit the econometric techniques by assessing the relationship between the per quarter number of banking failures and a set of covariates. As documented in Figure 2, an important feature of the dependent variable, FAILS, is that half of its observations take values between zero and three. Specifically, the outcome “zero” is the mode of the banking failures series with 25 hits. Due

to its particular shape, it is more appropriate to assume a Poisson or a Negative Binomial distribution for characterizing such as the dependent variable.

Figure 2: Density of the absolute number of fails



On the one hand, the “zero” outcomes can be the result of a sampling process: the banking system has a positive probability of performing particularly well. On the other hand, the “zero” outcomes can also be interpreted as structural zeros: the banking system shows zero failures because of its robust structure. If the latter is the case, it could be better to estimate the model by using techniques based on switching regimes, such as the zero inflated Poisson model. This type of model allows us to exploit all the information embedded in the dependent variable. Different models alternative to the zero inflated process can be employed even if they lead to a waste of information¹¹. This argument together with the features of the dependent variable, justify our choice to use the zero inflated Poisson model.

One of the main features of this method is that it is possible to take explicitly into account the amount of “zero” observations characterizing the dependent variable. A first specification defines the elements affecting the zero part of the dependent variable. Moreover, a second specification includes the variables that have an impact on its non-zero part. The following model is estimated where the absolute number of bank failures is the dependent variable:

¹¹FAILS could be transformed into a binary dependent variable and then to estimate a probit or logit model. One way to increase the information available is to transform FAILS into a variable that takes values 1, 2, 3 and 4 (using as thresholds values FAILS quantiles) and to use an ordered logit model.

$$y_t = \begin{cases} \beta Z_{it-1} + \epsilon_{t-1} & \text{if } y_t = 0 \\ \phi X_{it-1} + \eta_{t-1} & \text{if } y_t > 0 \end{cases} \quad (3)$$

A Probit model has been used for the inflated part, while a Poisson model has been employed when the dependent variable takes non-zero outcomes. In the specification *a* (column (1) of Table 7, Section C of the Appendix), the vector *Z* includes the return on assets and net interest margin, two variables reflecting the profitability and the profits of the banking system respectively. Meanwhile, vector *X* includes those variables describing the bank's fragility, such as the net loan losses, the non-performing loans and the loan loss reserve. In the alternative specification *b* (column (6) of Table 7, Section C of the Appendix), the net interest margin has been dropped from the inflated part and added to the vector of the explanatory variables referring to the non-inflated part. This choice is due to the ambiguous impact of the net interest margin on banking distress. All the explanatory variables are lagged by one period (corresponding to one quarter) in order to avoid endogenous issues and reverse causality problems¹².

The estimated coefficients are employed to generate the weights that must be ascribed to the indicators. The weights are assigned such that the most relevant variable obtains the largest weight. In order to measure the relevance of the variables, we construct a ratio that takes into account their estimated impact on the dependent variable and the correspondent precision. Specifically,

$$\chi_i = \frac{\hat{\beta}_i}{1 + se(\hat{\beta}_i)} \in [0, \beta_i] \quad (4)$$

¹²The regression results and the correspondent comments are reported in Table 7, Section C of the Appendix.

The ratio has been constructed so that a more precise estimation brings the ratio closer to the true value of the parameter ($\lim_{se(\hat{\beta}_i) \rightarrow 0} \chi_i = \beta_i$). Inversely, the less precise the estimation, the closer is the ratio to zero ($\lim_{se(\hat{\beta}_i) \rightarrow \infty} \chi_i = 0$).

The impact of the explanatory variables on the dependent variable is measured by the estimated marginal effects or by the Incidence Rate Ratios (IRRs). The former measure allows us to generate flexible weights (FLME) according to a specific parametrization¹³. The latter measure produces fixed weights, and it is defined as the exponential on the estimated coefficient of a particular variable i . It represents, in expected relative terms, how much the dependent variable changes for a unit change of the explanatory variable i , keeping the rest of the variables constant¹⁴. The precision of the estimations is measured by the standard errors of the estimated values.

The weights are ascribed so that when χ_i increases (in relative terms) the correspondent weight increases. Therefore, the generated weights take the following form:

$$\omega_i = f(\hat{\beta}_i, \hat{\beta}_{\neq i}, se(\hat{\beta}_i), se(\hat{\beta}_{\neq i})) \quad (5)$$

4.2.1 Robustness checks

In order to check the robustness of the results to the lag structure of the specification in Table 7, Section C of the Appendix from column (2) to (5) and from column (7) to (10), we report the findings using different lag orders. In particular, we assume that the dependent variable depends on the explanatory variables lagged by two, three, four, and five periods, respectively. The findings suggest that the baseline model results are robust with only some exceptions concerning the loan loss provisions.

Moreover, in order to check the robustness of the results with respect to the variable

¹³The parametrization is based on the values taken by the variables in each period.

¹⁴In order to make the Probit estimates comparable with the Poisson estimates, we proceeded to a rescale process. More details are provided in Section A of the Appendix.

employed for approximating banking distress, we replace the quarterly absolute number of bank failures (FAILS) by a measure that takes in to account both the fraction of bank failures and the value fraction of the assets failed (BAF). The two variables are plotted together in Figure 1.a and, as previously documented, they behave in the same way even if they have a different scale.

The correspondent vector of optimal weights using BAF has been computed according to eq.(4) and eq.(5) after estimating the following linear model:

$$BAF_t = \alpha + \beta_1 ROA_{t-1} + \beta_2 LSTL_{t-1} + \beta_3 NPTL_{t-1} + \beta_4 LLRTL_{t-1} + \beta_5 NIM_{t-1} + \xi_t \quad (6)$$

4.3 Weights and indexes

According to eq.(4) and eq.(5), we compute the vector of the optimal weights. The main findings are reported in Table 1¹⁵.

Table 1: Vector of weights

	F_0 (1)	F_{10} (2)	F_{40} (3)	IRRa (4)	FLMEa (5)	IRRb (6)	FLMEb (7)	BAF (8)
ROA	-.165	-.225	-.295	-.21	-.17	-.19	-.16	-.14
LSTL	.175	.16	.15	-.14	-.2	-.13	-.19	-.14
NPTL	.18	.30	.175	.51	.42	.48	.37	.35
LLRTL	.20	.18	.22	.04	.06	.09	.12	.21
NIM	.28	.135	.16	.1	.15	-.11	-.16	-.16

The results referring to the MSA weights, columns (1), (2) and (3), are not robust to the criterion employed, even if they are similar to those obtained by employing a variance-equal approach. Specifically, each of the criterion produces a different leading indicator: NIM if F_0 is employed, NPTL if F_{10} is used, and ROA if F_{40} is preferred. The leading indicators show a weight, in absolute value, around .3, while the other variables report, in absolute value, a

¹⁵It is possible to compare the results found using the different methods by using average values for the FLME weights.

weight between .16 and .225 depending on the case. The correspondent indexes exhibit similar patterns, even if the three criteria show different weight vectors. This fact suggests that the signs ascribed to the variables, which in the three cases are the same, play a crucial role in the determination of the shapes of the indexes.

Focusing on the results based on the zero inflated regressions, from column (4) to column (7), the main finding refers to the leading role played by NTPL. Its weight reports values included between .37 and .5. The second important result is about the ROA, LSTL and the NIM. In both specifications (*a* and *b*), these three variables play a crucial role. In absolute value, their weights are always larger or equal to .1 and smaller or equal to .21. Finally, a minor role is played by LLRTL. It reports in absolute value a weight bounded between .04 and .12.

The MSA and the zero inflated approach lead to different results also in terms of the signs associated with the variables. The LSTL enters in the indexes with the negative sign according to the econometric approach, while NIM shows a negative sign in the alternative econometric specification (*b*).

The weights obtained by employing BAF as a measure of bank failures are similar in sign and in magnitude to those based on the FLMEb. Therefore, the results suggest that the findings are robust to the variables employed for describing banking distress periods. In terms of variables' weights variability depending to the method chosen, LSTL turns out to be the most stable, while NPTL shows the highest variability.

The indexes are generated according to the results of Table 1. Indexes are easy to interpret. They are generated by combining indicators that have a zero mean and a unit standard error. It follows that index levels larger than zero refer to periods of banking distress that is higher than average, while negative values are associated with a banking situation that is better than average.

In the baseline analysis, due to the similarity in results between the three indexes based

on MSA method, we focus on the F_{40} index. Moreover, we also report the IRRa and IRRb as well as the FLMEa and FLMEb indexes, according to the zero inflated regressions. These indexes have been compared with those obtained by using the variance-equal approach and the factor analysis.

5 Ability, Stability and Forecasting Properties of the Indexes

In order to measure the performance of the indexes we look at three dimensions. First, we assess their ability to capture specific events of interest, such as recessions, financial crises and stock exchange crashes. Moreover, we also analyse how the indexes behave in correspondence with high levels of the value fraction of the assets failed or large values of the per failed bank assets value. Second, we check the stability of the weights ascribed to the variables. In particular, we test their robustness with respect to the sample periods chosen. Finally, we study their forecasting properties, testing as well their robustness properties with respect to the model's specifications.

5.1 Ability to detect events of interest

We measure the ability of the indexes to signal distressing situations by taking into account several economic and financial events. Specifically, we focus on regional US economic downturns (1986-92 in the South-Western states, 1991-92 in the North-Eastern region, and 1992-93 in California) and on some events characterizing the US stock market (the October 1987 crash, the October 1997 mini-crash and the March 2000 dot-com bubble crash). Moreover, we also focus on the two recessions that hit the US economy between July 1990 and March 1991 and from March to November 2001. Finally, we analyse the behaviour of the indexes in 2006 and 2007, just before the beginning of the sub-prime crisis. Meanwhile, we

test the capability of the indexes to mimic the main patterns of relevant banking distress indicators.

In order to assess the performance of the indexes with respect to this dimension, we focus on the magnitude of the banking distress reported as well as on their trend patterns in correspondence of specific relevant economic and financial episodes.

Figures 3 and 4, Section D of the Appendix, report the main findings. Specifically, in Figure 3, each of the indexes has been plotted together with the VE index and the quarterly number of banking failures. The results highlight several interesting elements. First, the MSA and the VE indexes (Figure 3.a) show similar features in terms of shape and the level of distress they can report. The same is true for the index based on FA estimations (Figure 3.b), even if there are some differences in terms of its magnitude. The indexes, based on the zero inflated regressions, show relevant differences with respect to the VE index benchmark in terms of their shape and banking distress levels they report. Specifically, the main differences arise at the beginning and at the end of the period analysed and during the period included between the first recession and the end of the regional crisis. The VE, MSA and the FA indexes exhibit a substantial dispersion around their mean values (between .78 and .96). Furthermore, the indexes based on the zero inflated regressions are more centred around their average values (zero) with standard errors between .42 (FLMEb) and .68 (IRRa).

All the indexes correctly capture the several regional US economic crisis crises occurring between 1986 and 1993, as they report values of banking distress that are larger than average. However, indexes based on the zero inflated regressions highlight more inflated distress levels before the beginning of the regional crises, suggesting that they lead the episodes studied. In addition, the first economic recession is well represented by all the indexes in terms of distress level (higher than the average) and trend. The best results refer to the indexes based on the zero inflated regressions. The more heterogeneous results refer to the second economic recession. First, none of the indexes reports a distress level higher than average. In terms

of the distress trend highlighted by the indexes, the best result is achieved by the VE, MSA and the FA indexes, while the IRR (figure 3.c, 3.d) and FLME (figure 3.e, 3.f) indexes are unable to capture this event. Finally, at the beginning of the real estate burst, none of the indexes reports levels of banking distress higher than average. The best result is achieved through the FLMEb index, which exhibits a level of distress that is closer to the average value. The VE, MSA and FA indexes show an increasing pattern, although they report a distress level that is one standard deviation smaller than the average value. The IRRa, IRRb and the FLMEa indexes show intermediate results.

All the indexes lead the 1987 stock exchange crash, showing distress values above the average. However, the results related to the mini-crash of 1997 are more heterogeneous. On the one hand, all of the indexes report a distress level below the average, while on the other hand, only the IRRb and FLMEb indexes show a measurable increasing trend of distress level. When analysing the behaviour of the indexes during the dot-com bubble bust in 2000, the best results are achieved by the FA, IRRb and the FLMEb indexes, because they show a measurable increasing trend, while the other indexes provide less informative results.

Our analysis of the ability of the indexes to signal distress levels adequate to describe high levels of the value fraction of the assets failed or large values of the per failed bank assets value, highlights that the best indexes are those based on econometric estimations. More precisely, as shown in Figure 3 all the indexes fit well the banking failures series, reporting higher levels of banking distress when failures are higher. This result is not surprising, given the fact that the weights of these indexes have been obtained by using the number of banking failures as the indicator of banking distress. The results related to the VE and FA approaches are instead more surprising, because they have been constructed without taking into account any banking distress benchmark.

Moreover, the IRR and FLME indexes perform better if we take into account the ability of the indexes to signal distress episodes according to the per failed banks assets value in

thousand of dollars or to the value fraction of the failed assets. Figure 4, Section D of the Appendix highlights that this is true for both the shape of the indexes and for the magnitude of the distress level reported.

After combining previous findings and witnessing how they operate together, it follows that from a general view point, the FLMEb is the best index to detect the main economic and financial events that occurred during the period analysed. The FLMEb is also the best index to signal high levels of the value fraction of the assets failed or large values of the per failed bank assets value. Moreover, it seems that all the indexes capture events related to the banking sector better than they can capture other events. This fact could be due to the variables employed that strictly refer to the banking sector.

5.2 Weights stability

The indexes can be classified depending on the features of their weights. On the one hand, the VE, MSA, FA, IRRa and IRRb indexes are characterized by fixed weights, while the FLMEa and FLMEb are based on flexible weights, specifically, on weights that change at each point in time.

In this section, we test the stability properties of the weights of those indexes that can be affected by arbitrary sample period choices. We focus on the MSA, FA, IRRa and the IRRb indexes. We exclude the VE index because, by definition, its weights are invariant to time period, while for the opposite reason, we exclude the indexes based on flexible weights. In order to test the weights stability we define two periods. We use the period between 1984:Q1 and 1998:Q4 to compute the vector of optimal weights; we thereafter employ the weights for generating the indexes by using the period out of the sample, going from 1999:Q1 to 2007:Q3.

The weights and therefore the index is considered stable over time if the out-of-sample index lies between the confidence interval at 95% bounds constructed around the index

generated using the entire sample period. The results are reported in Figure 5, Section D of the Appendix.

Figure 5 shows that the best indexes in terms of weights stability are the MSA (figure 5.a) and the IRRa (Figure 5.c). In both cases, the out-sample index perfectly overlap the index computed by taking into account the entire sample period.

Moreover, the FA (Figure 5.b) and the IRRb (Figure 5.d) indexes highlight unstable weights. In both cases, the out-sample indexes lie outside the confidence interval bounds.

5.3 Forecasting

Monetary authorities could anticipate situations and take actions in order to prevent banking crises by forecasting banking distress. In order to deal with this issue we follow Early Warning Systems models (EWSs) based on a macro approach. Specifically, the following variables have been used in order to explain the variability of the banking quality indicator: the percentage difference between the cyclical component and the trend component of the gross domestic product (GAP), the inflation rate (CPI), the prices of the shares of the companies traded on national or foreign stock exchanges (SP), the percentage of the credit-income ratio (CI) and the Median Sales Prices of New Homes Sold in the United States (MDHP). GAP has been included to take into account the overall economic condition, CPI and SP help in understanding the prices dynamics, CI is a proxy for the banking system credit behaviour and MDHP captures the features of the real estate market. Based on previous studies on this topic¹⁶, we estimate the following model:

$$Stress_i = \alpha + \beta_1 GAP_{t-1} + \beta_2 CPI_{t-1} + \beta_3 CI_{t-1} + \beta_4 MDHP_{t-1} + \beta_5 SP_{t-1} + \epsilon_t \quad (7)$$

¹⁶See Hanschel and Monnin (2005), among others.

Table 8, Section C of the Appendix, reports the main results and a detailed analysis of the expected signs. In the majority of the cases, the variable referring to the macroeconomic conditions (GAP) shows the expected negative sign, and the corresponding coefficient is statistically significant. The same is true for the CPI; in all of the specifications, it shows the expected sign and the coefficient is statistically significant.

The findings referring to the other variables are puzzling: depending on the index, the results change drastically. This can be due to the fact that the results reflect the approaches employed for aggregating the variables included in the index. Depending on the sign ascribed to each variable, the impact of the explanatory variables on the dependent variable changes significantly. From a general point of view, the results are robust to the modification of the lag structure characterizing the specification (Table 9, Section C of the Appendix), even if some important changes occur. More precisely, after employing the explanatory variables lagged by four periods, the macroeconomic variable (GAP) is no longer statistically significant, while the results referring to the share prices (SP) show a positive and significant impact on the dependent variable. The findings about the rest of the variables do not change, after employing the explanatory variables lagged by four periods.

In order to implement the forecasting part, an in-out-sample fixed window analysis has been employed. First, we estimate the fitted values using the full sample. Moreover, we split the sample in two parts. The coefficients are estimated using the first 90 observations (from 1984:Q1 to 2006:Q2), while the rest of the sample is used for the forecasting part. The results are reported in Figure 6 and Figure 7, Section D of the Appendix.

We check the quality of the forecast part has been checked by comparing the root-mean-square error (RMSE) obtained by the regression analysis. The smaller the RMSE is, the better is the forecast computed. We check the robustness of the results with respect to the lag structure of the model by employing the covariates lagged by four periods. The main findings are reported in Table 2. The best results are those of the FLMEb, regardless the lag

structure employed. Moreover, the MSA and VE indexes are the least performing in terms of forecasting properties.

The three elements that this paper employs to assess the performance of the indexes proposed by the current study are the indexes' ability to detect relevant events, their weight stability properties and their forecasting features. Our results suggest that the best indexes to use to assess banking distress levels is the FLMEb, based on the zero inflated Poisson regressions. This index performs better with respect to the three features we identified for qualifying the index quality. Moreover, the results also highlight that the VE and the FA indexes show similar patterns, both in terms of shape and in forecasting properties. One relevant role in the performance of the indexes seems to be played by the sign that is ascribed to the variables in the aggregation process.

Table 2: RMSE

	F_{40}	VE	FA	$IRRa$	$IRRb$	$FLMEa$	$FLMEb$
1 lag	1.0661764	1.1483574	1.1783307	.40568919	.22906947	.34507577	.14145408
4 lags	1.0304408	1.1175789	1.1917365	.3760213	.19408506	.34378766	.07401539

6 Conclusion

The financial turmoil that plummeted the world economy in recent years called for a more precise view and a closer monitoring of the quality of the banking sector. The faster a financial crisis shows up and the bigger its impact on the economic sector is, the more relevant the synthetic measures are for assessing the level of banking distress. Analysts need these tools to gain a quick picture of the soundness of the banking sector; to run more precise forecasting exercises assessing the effect of specific policies on the quality of the banking sector; to improve the quality of the studies about the impact of new capital requirements rules on the banking sector.

In the literature, the two main methods employed to merge the information of several

indicators into a single index are the variance-equal approach and the factor analysis. The former approach has no actual economic support, while the latter is not always applicable. In this paper we proposed two alternative methods, based on the modified signalling and the zero inflated regressions approach for generating synthetic indexes in order to measure the distress level in the banking system. We compared the properties of the indexes obtained by applying the approaches proposed in this study with those of the indexes constructed by using the variance-equal approach and factor analysis approach. This comparative exercise has been set by taking into account three dimensions: the ability of the indexes to capture specific events of interest; the stability of the weights ascribed to the variables depending on the period taken into account; and the forecasting features of the indexes.

The results show that the FLMEb index, based on the zero inflated regressions, is the best among the indexes. It is the best at capturing relevant events and signalling distress levels that correspond to high level of the value fraction of the assets failed or large values of the per failed bank assets value. The FLMEb exhibits, as well, the best forecasting properties. Moreover, the results also highlight that the VE, FA and the MSA indexes show similar patterns both in terms shape and in forecasting properties.

The relevance of this study is two-fold: we contribute to the provision of extra information that can be useful for forecasting banking system soundness and preventing future financial crises. Moreover, this study provides alternative methods for computing banking distress indexes, and we show that these models perform better than those based on the traditional approaches.

References

- [1] L. Allen, “The determinants of bank interest margins: A note,” *Journal of Financial and Quantitative Analysis*, vol. 23, pp. 231–235, June 1988.
- [2] C. Arteta and B. Eichengreen, “Banking crises in emerging markets: Presumptions and evidence,” Center for International and Development Economics Research (CIDER) Working Papers C00-115, University of California at Berkeley, Aug. 2000.
- [3] E. Balla and A. McKenna, “Dynamic provisioning: a countercyclical tool for loan loss reserves,” *Economic Quarterly*, no. Fall, pp. 383–418, 2009.
- [4] S. Carbó-Valverde and F. Rodríguez-Fernández, “The determinants of bank margins in european banking,” *Journal of Banking and Finance*, vol. 31, pp. 2043–2063, July 2007.
- [5] S. Carbó-Valverde, D. Humphrey, and F. Rodríguez-Fernández, “Deregulation, bank competition and regional growth,” *Regional Studies*, vol. 37, pp. 227–237, May 2003.
- [6] A. Demirguc-Kent and E. Detragiache, “Financial liberalization and financial fragility,” Policy Research Working Paper Series 1917, The World Bank, May 1998.
- [7] A. Demirguc-Kunt and E. Detragiache, “Monitoring banking sector fragility : a multivariate logit approach with an application to the 1996-97 banking crises,” Policy Research Working Paper Series 2085, The World Bank, Mar. 1999.
- [8] A. Demirguc-Kunt and E. Detragiache, “The determinants of banking crises in developing and developed countries,” *IMF Staff Papers*, vol. 45, p. 29, March 1998.
- [9] M. Morales and D. Estrada, “A financial stability index for Colombia,” *Annals of Finance*, vol. 6, pp. 555–581, October 2010.

- [10] M. Goldstein, C. Reinhart, and G. Kaminsky, “Assessing financial vulnerability, an early warning system for emerging markets: Introduction,” MPRA Paper 13629, University Library of Munich, Germany, 2000.
- [11] B. González-Hermosillo, “Determinants of ex-ante banking system distress: A macro-micro empirical exploration of some recent episodes,” IMF Working Papers 99/33, International Monetary Fund, Mar. 1999.
- [12] E. Hanschel and P. Monnin, “Measuring and forecasting stress in the banking sector: evidence from Switzerland,” in *Investigating the relationship between the financial and real economy* (B. for International Settlements, ed.), vol. 22 of *BIS Papers chapters*, pp. 431–49, Bank for International Settlements, December 2005.
- [13] T. S. Y. Ho and A. Saunders, “The determinants of bank interest margins: Theory and empirical evidence,” *Journal of Financial and Quantitative Analysis*, vol. 16, pp. 581–600, November 1981.
- [14] M. Illing and Y. Liu, “An index of financial stress for Canada,” Working Papers 03-14, Bank of Canada, 2003.
- [15] N. Jahn and T. Kick, “Determinants of banking system stability: A macro-prudential analysis,” Mimeo, University of Münster, Finance Center Münster, 2012.
- [16] G. L. Kaminsky, “Currency and banking crises - the early warnings of distress,” IMF Working Papers 99/178, International Monetary Fund, Dec. 1999.
- [17] G. L. Kaminsky and C. M. Reinhart, “The twin crises: The causes of banking and balance-of-payments problems,” *American Economic Review*, vol. 89, pp. 473–500, June 1999.

- [18] M. Manove, A. J. Padilla, and M. Pagano, “Collateral versus project screening: A model of lazy banks,” *The RAND Journal of Economics*, vol. 32, no. 4, pp. pp. 726–744, 2001.
- [19] J. Marcucci and M. Quagliariello, “Is bank portfolio riskiness procyclical: Evidence from Italy using a vector autoregression,” *Journal of International Financial Markets, Institutions and Money*, vol. 18, pp. 46–63, February 2008.
- [20] J. Maudos and J. Fernandez de Guevara, “Factors explaining the interest margin in the banking sectors of the European Union,” *Journal of Banking and Finance*, vol. 28, pp. 2259–2281, September 2004.
- [21] M. A. Petersen and R. G. Rajan, “The effect of credit market competition on lending relationships,” *The Quarterly Journal of Economics*, vol. 110, pp. 407–43, May 1995.
- [22] M. Quagliariello, “Banks’ performance over the business cycle: A panel analysis on Italian intermediaries,” Discussion Papers 04/17, Department of Economics, University of York, 2004.
- [23] A. Saunders and L. Schumacher, “The determinants of bank interest rate margins: an international study,” *Journal of International Money and Finance*, vol. 19, pp. 813–832, December 2000.
- [24] S. Shaffer, “Patterns of competition in banking,” *Journal of Economics and Business*, vol. 56, no. 4, pp. 287–313, 2004.
- [25] K. P. Wong, “On the determinants of bank interest margins under credit and interest rate risks,” *Journal of Banking and Finance*, vol. 21, no. 2, pp. 251 – 271, 1997.
- [26] E. R. Zarruk and J. Madura, “Optimal bank interest margin under capital regulation and deposit insurance,” *Journal of Financial and Quantitative Analysis*, vol. 27, pp. 143–149, March 1992.

Appendices

A Marginal effects in the Probit model

Let us assume to estimate the following Probit model:

$$y_i = X\beta + \epsilon \quad (8)$$

with

$$y_t = \begin{cases} 1 & \text{if } y^* > 0 \\ 0 & \text{if } y^* \leq 0 \end{cases} \quad (9)$$

In this case, the probability that $y = 0$, given a set of values of the explanatory variables, $P(y = 0 \mid X)$ is given by the following formula:

$$P(y = 0 \mid X) \equiv 1 - G(X\beta) \quad (10)$$

where $G(X\beta)$ is the cdf of the error term ϵ . Let us assume that the only explanatory variable is the ROA and a constant. Moreover, let us define by ρ the ratio between the $P(y = 0 \mid X)$ computed at $ROA = 1$ and the same probability at $ROA = 0$. Specifically, we have:

$$\rho = \frac{P(y = 0 \mid ROA = 1)}{P(y = 0 \mid ROA = 0)} \quad (11)$$

The ratio ρ is a function of the parametrization and of the estimated coefficient $\hat{\beta}_{ROA}$ and $\hat{\alpha}$. That is $\rho = f(ROA, \hat{\beta}_{ROA}, \hat{\alpha})$. Suppose that the ratio $\rho = .33$, this means that the probability of having $y = 0$ increases by .33% when increasing by one unit the explanatory variable while keeping the rest constant. In our analysis, the ratio ρ has been used to compare the results in the no-inflated part to the IRR results that refer to the inflated part of the regression. The correspondent χ_i , for the explanatory variables included in the Probit part of the Zero inflated process, is equal to the ρ_i , relative to the variable i , over one plus its standard error

$$\chi_i = \frac{\rho(\hat{\beta}_{ROA})}{1 - se(\rho(\hat{\beta}_{ROA}))} \quad (12)$$

B Tables

Table 3: Variables description

Variable	Definition	Source
Return on Assets (ROA)	Fiscal year's earnings divided by its total assets (%)	Fed. St. Luis
Net Loan Losses over Average Total Loans (LSTL)	Defaulted loans	Fed. St. Luis
Non-performing Loans over Total Loans (NPTL)	Any commercial loans that are more than 90 days overdue	Fed. St. Luis
Loan Loss Reserve over Total Loans (LLRTL)	Reserves for those assets at banks whose allowance for loan and lease losses exceeds their NPTL	Fed. St. Luis
Net Interest Margin (NIM)	The dollar difference between interest income and interest expenses	Fed. St. Luis
Number of bank failures (FAILS)	Number of commercial banks failed	FDIC
Gross Domestic Product (GAP)	the % difference between the cyclical and the trend component of the GDP	OECD
Inflation rate excluding food and energy (CPI)	% change with respect to the same quarter of the previous year	OECD
Credit-income ratio (%) (CI)	Household financial obligations as a % of Disposable Personal Income	Fed. St. Luis
Median and Average Sales Prices of New Homes Sold in United States (MDHP)	% change with respect to the same quarter of the previous year	Census.gov
Prices of common shares of companies traded on national or foreign stock exchanges (SP)	% change with respect to the same quarter of the previous year	OECD

Table 4: Descriptive statistics and correlations

Variable	Obs.	Mean	St. Dev.	Skewness	Kurtosis	Min	Max		ROA.	LSTL	NPTL	LLRTL	NIM	FAILS
ROA	95	1.04	.349	-1.31	4.92	-4.04	1.05	ROA	1					
LSTL	95	.78	.2944	.919	3.16	-1.47	2.83	LSTL	-.47	1				
NPTL	95	1.97	1.088	.41	1.49	-1.17	1.77	NPTL	-.855	.7	1			
LLRTL	95	1.94	.5	.177	1.81	-1.56	1.61	LLRTL	-.42	.63	.57	1		
NIM	95	4.02	.31	-.074	2.99	-2.19	2.86	NIM	-.15	.26	.33	.68	1	
FAILS	95	15.73	21.62	1.55	5.17	0	99	FAILS	-.77	.49	.8	.486	.153	1

Notes: In the first seven columns, for all the series, the number of observations, the mean, the standard error, the skewness and kurtosis, the minimum and maximum values are reported. In the last six columns the correlations between the variables are reported.

Table 5: Criteria features

Criteria	F_0	F_{10}	F_{40}
Number of failures	0	10	40
Frequency	Quarterly	Yearly	Yearly
“Crisis quarters” (%)	73%	50%	42%

Notes: F_0 : A quarter is classified as “crisis quarters” if there is at least one bank fail; $F_{10}(F_{40})$: All the quarters of a year in which there are at least ten(forty) banks failures are classified as “crisis quarters”.

Table 6: MSA averages and bounds

	Crisis			Non-crisis		
Averages	F_0	F_{10}	F_{40}	F_0	F_{10}	F_{40}
ROA	-.255**	-.67***	-.94***	.714***	.691***	.685***
LSTL	.234**	.60***	.704***	-.655***	-.613***	-.512***
NPTL	.33***	.85***	1.1***	-.924***	-.87***	-.806***
LLRTL	.244**	.508***	.54***	-.683***	-.519***	-.39***
NIM	.249**	.445***	.327***	-.698***	-.455***	-.237**
	Crisis			Non-crisis		
Bounds	F_0	F_{10}	F_{40}	F_0	F_{10}	F_{40}
ROA	.0676	-.304	-.568	.61	.61	.6
LSTL	-.0827	.224	.287	-.38	-.43	-.279
NPTL	.03308	.6	.941	-.83	-.81	-.72
LLRTL	-.06507	.105	.041	-.37	-.31	-.17
NIM	-.009	.16	.044	-.13	-.07	.17

Notes: *** = $p < .01$, ** = $p < .05$, * = $p < .1$. Null hypothesis: the parameter equals zero. In case of a variable with positive (negative) average crisis values, we report only the lower (upper) confidence interval bound. The opposite is true for the non-crisis periods. In the upper part of Table 6, according to the criteria (F_0 , F_{10} and F_{40}), the conditional means of the standardized variables are reported. The return on assets (ROA) shows the expected sign with an average value during the crisis below the overall period average. The opposite is true for the net interest margin, which shows a positive mean during the crisis. The variables referring to the quality of the banking sector (LSTL, NPTL and LLRTL) show values above the overall period average during the crisis. In the majority of the cases, the gap between the observed value and its overall period average increases as the criterion for defining a crisis becomes more conservative. All indicators show conditional means that are consistent with the economic theory. The results are also robust to small changes of the criteria conditions. In the lower part of Table 6 the thresholds for measuring the ability of the variables in detecting crisis and non-crisis periods are reported. More precisely, the thresholds are defined as the upper or lower bounds of a confidence interval at 99% around the conditional means during the crisis and non-crisis periods.

C Regression results

Table 7: Regressions results

	a	L2	L3	L4	L5	b	L2	L3	L4	L5
Inflated part	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ROA	2.084*** (.474)	2.087*** (.767)	2.063*** (.733)	1.981*** (.695)	2.362*** (.826)	2.853*** (.667)	2.824*** (.901)	2.772*** (.878)	2.514*** (.800)	3.026*** (1.002)
NIM	-.685** (.332)	-.623* (.337)	-.617* (.337)	-.469 (.330)	-.422 (.335)					
Const	-2.033*** (.134)	-1.995*** (.562)	-1.879*** (.538)	-1.811*** (.514)	-1.955*** (.607)	-2.335*** (.44)	-2.307*** (.669)	2.207*** (.0585)	-2.234*** (.654)	2.270*** (.0567)
Obs.	25	25	25	25	25	25	25	25	25	25
Not inflated part										
LSTL	-.275*** (.0979)	-.363*** (.0358)	-.306*** (.0363)	-.374*** (.0376)	-.298*** (.0386)	-.267*** (.0898)	-.352*** (.0352)	-.294*** (.0356)	-.347*** (.0367)	-.271*** (.0377)
NPTL	1.286*** (.122)	1.383*** (.0524)	1.344*** (.0519)	1.466*** (.0552)	1.449*** (.0556)	1.188*** (.110)	1.279*** (.0519)	1.231*** (.0504)	1.333*** (.0522)	1.340*** (.0529)
LLRTL	.0742 (.0891)	.0674** (.0296)	-.0177 (.0298)	-.0521* (.0297)	-.186*** (.0304)	.189* (.0998)	.170*** (.0343)	.112*** (.0344)	.101*** (.0338)	-.0378 (.0344)
NIM						-.334** (.137)	-.312*** (.0559)	-.401*** (.0576)	-.523*** (.0615)	-.522*** (.0650)
Const	2.156*** (.325)	2.087*** (.0592)	2.138*** (.0584)	2.040*** (.0633)	2.058*** (.0639)	2.268*** (.126)	-2.073*** (.601)	2.188*** (.0598)	-2.387*** (.757)	2.181*** (.0611)
Obs.	69	68	67	66	65	69	68	67	66	65
Vuong/Lnalpha	2.24**	2.131**	2.04**	1.852**	2.121**	2.41***	2.5***	2.52***	2.464***	2.599***
Tot obs.	94	93	92	91	90	94	93	92	91	90

Notes: *** = $p < .01$, ** = $p < .05$, * = $p < .1$. For both estimations, the Vuong test confirms that the zero inflated model is preferred to the classical Poisson model. When focusing on the inflated part of baseline estimation (a), column (1), ROA is statistically significant and with the expected sign. The larger the ROA, the higher the probability of having zero failures. The coefficient for the NIM is statistically significant, and it has a negative sign: a higher NIM implies a lower probability of zero failures. The leading role in the non-inflated part is played by NPTL. Its coefficient is statistically significant, reporting the expected positive sign. A higher level of non-performing loans positively affects the probability of having a non-zero number of failures. The coefficient referring to the loan loss reserve is not statistically significant, even though it reports the expected sign. Finally, the net loan losses highlights statistically significant results, reporting a sign that is opposite to the expected, but consistent with the correlation coefficient. Specifically, the estimated coefficient is negative, implying that the larger the LSTL, the lower the probability of failures. The results referring to the alternative specification (b), column (6), support the finding about the key role played by the non-performing loans. The results related to ROA are also confirmed: the higher the return on assets, the lower is the probability of default. The LSTL results are similar in the two specifications. The estimated coefficient for the LLRTL is statistically significant at 10%: higher LLRTL implies a higher probability of failures. Finally, the findings about the net interest margin are negative and statistically significant: the higher the NIM, the lower the probability of default. It turns out that the NIM results obtained in the two specifications are contradictory.

Table 8: Regression results, forecasts baseline

	MSA (1)	VE (2)	FA (3)	IRRa (4)	IRRb (5)	FLMEa (6)	FLMEb (7)
GAP_1	-.110* (.048)	-.118* (.049)	-.143** (.054)	-.085** (.031)	-.057 (.029)	-.057 (.032)	-.003 (.026)
CPI_1	.565*** (.044)	.531*** (.044)	.718*** (.051)	.546*** (.029)	.510*** (.025)	.395*** (.030)	.360*** (.021)
CI_1	-.087 (.099)	-.124 (.105)	-.066 (.117)	-.091 (.050)	.102** (.035)	-.246*** (.054)	.140*** (.029)
$MDHP_1$	.006 (.012)	.000 (.012)	.009 (.014)	.021** (.007)	.024*** (.006)	.024** (.008)	.024*** (.005)
SP_1	.001 (.003)	-.000 (.003)	-.002 (.004)	.004 (.002)	.003 (.002)	.003 (.002)	.003 (.002)
Constant	-.335 (1.719)	.450 (1.811)	-1.183 (2.017)	-.321 (.923)	-3.568*** (.659)	2.827** (.985)	-3.817*** (.542)
Obs.	94	94	94	94	94	94	94
R ²	.676	.647	.701	.850	.844	.789	.786

Notes: *** = $p < .01$, ** = $p < .05$, * = $p < .1$.

Table 9: Regression results, forecasts robustness

	MSA (1)	VE (2)	FA (3)	IRRa (4)	IRRb (5)	FLMEa (6)	FLMEb (7)
GAP_4	.021 (.035)	.012 (.036)	.007 (.042)	-.010 (.026)	.028 (.027)	-.058* (.026)	.024 (.025)
CPI_4	.649*** (.037)	.634*** (.035)	.830*** (.040)	.571*** (.028)	.498*** (.027)	.430*** (.029)	.331*** (.024)
CI_4	.103 (.096)	.089 (.102)	.203 (.109)	.015 (.044)	.173*** (.035)	-.171*** (.050)	.155*** (.035)
$MDHP_4$	-.006 (.010)	-.013 (.010)	-.005 (.012)	.017** (.006)	.022** (.006)	.022** (.007)	.028*** (.006)
SP_4	.007* (.003)	.005 (.003)	.006 (.003)	.007** (.002)	.006* (.002)	.006* (.002)	.004* (.002)
Constant	-3.911* (1.643)	-3.557* (1.730)	-6.226** (1.867)	-2.283** (.796)	-4.806*** (.645)	1.355 (.893)	-4.034*** (.649)
Obs.	91	91	91	91	91	91	91
R ²	.764	.753	.796	.858	.818	.804	.725

Notes: *** = $p < .01$, ** = $p < .05$, * = $p < .1$.

The gross domestic product gap is expected to negatively affect the stress banking index: a higher GAP implies a lower level of defaults. This leads to a reduction of the non-performing loans, a decrease in loan loss provisions and increases in ROA. The higher the GAP is, the lower the banking stress level. GAP expected sign is negative.

A higher level of the consumption prices index has detrimental effects on household consumption. Moreover, a higher level of CPI may generate a response of the monetary authority through an increase in the interest rate. Therefore, the price of assets can decrease, the households' net worth decreases, the non-performing loans increase and then the level of stress rises. An unexpected increase in the interest rate in response to high inflation rate may also have another effect on the stress level working through the cost-of capital channel: lending conditions become more demanding, leading to a contraction in the credit supply. Consequently, the level of stress in the banking system increases as the price level increases. The expected sign for the CPI is positive.

The credit to income ratio is defined as the ratio between household financial obligations and disposable personal income. Usually, a growing economy shows an increasing level in the credit ratio. However, if credits grow too fast with respect to disposal income, the credit risk in the banking sector could be increasing. This means that banks are decreasing their lending standards, and borrowers of lower quality may have access to the credit market. This implies that the banking stress level increases as the credit ratio deviates from its long-term trend. As a consequence, its expected sign is positive.

The Median Sales Prices of New Homes Sold in the United States (MDHP) captures the real estate market impact on the stress index. Given the fact that an important fraction of the households' wealth is in real estate, higher levels in MDHP are expected to positively affect home-owners' wealth. In other words, real estate can be interpreted as collateral. This means that householders' borrowing capacity can increase as MDHP goes up. MDHP is expected to positively affect the quality of the banking system. The same is true for the shares price.

D Figures

Figure 3: Indexes and economic and financial events

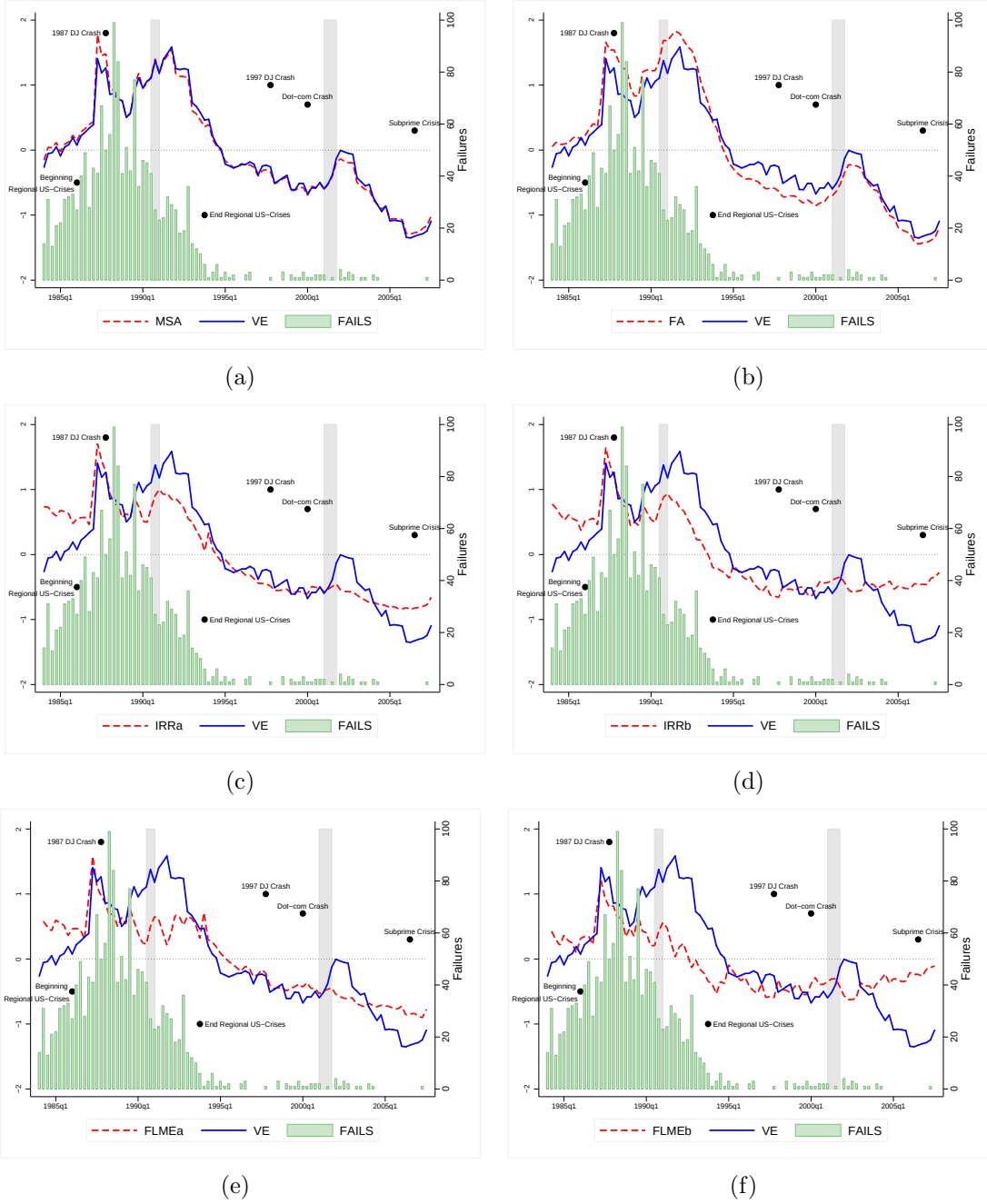
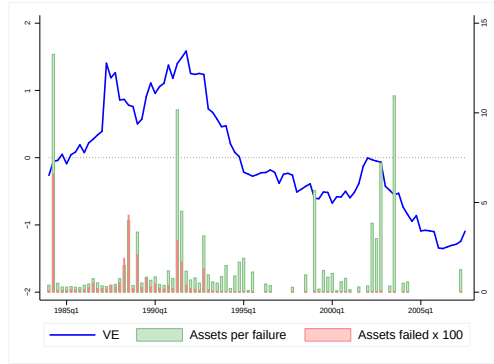


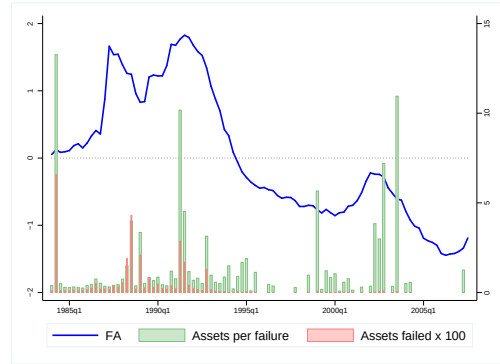
Figure 4: Indexes and banking distress indicators



(a)



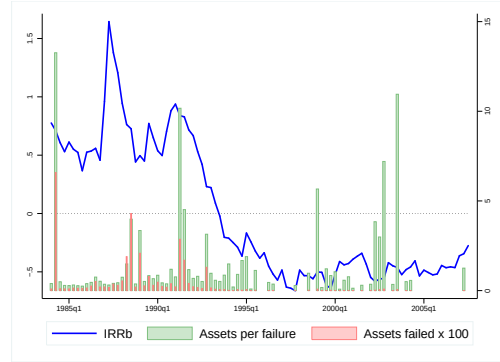
(b)



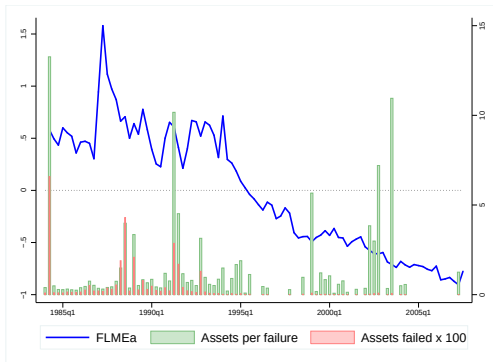
(c)



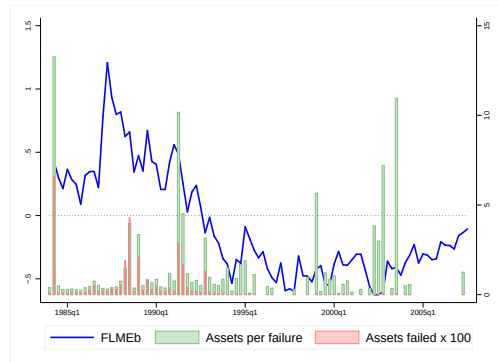
(d)



(e)

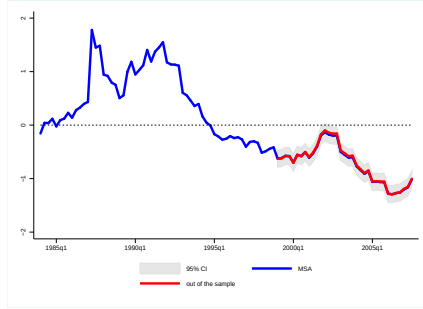


(f)

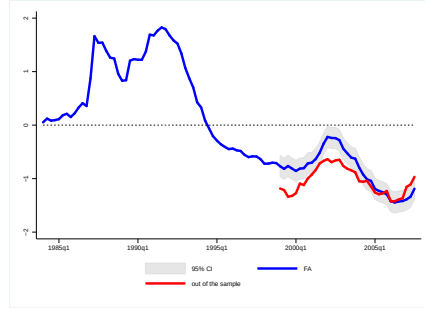


(g)

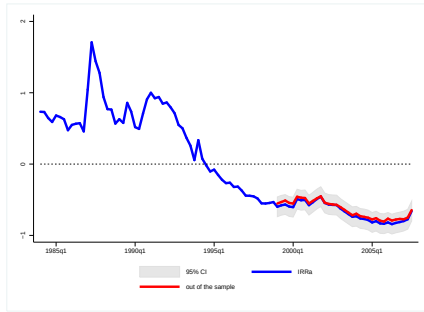
Figure 5: Weights Stability



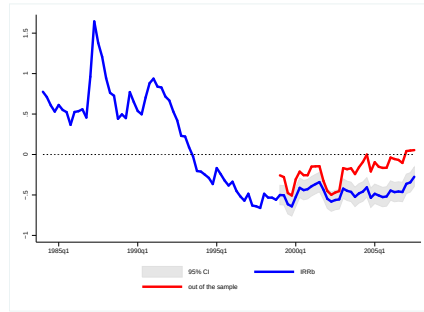
(a)



(b)

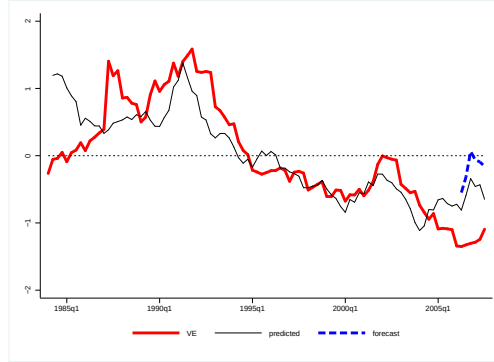


(c)

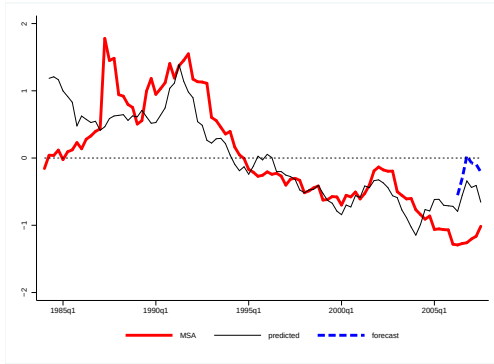


(d)

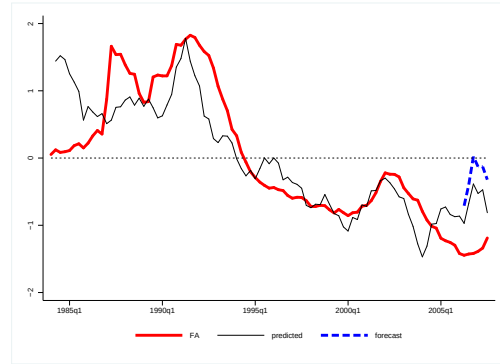
Figure 6: Forecasts baseline



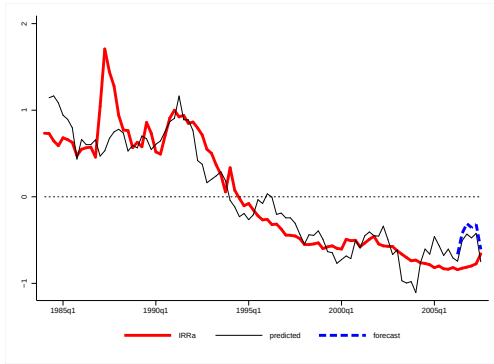
(a)



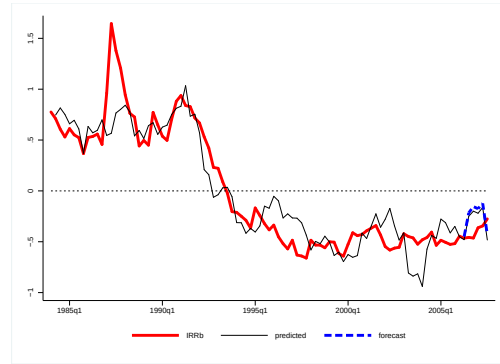
(b)



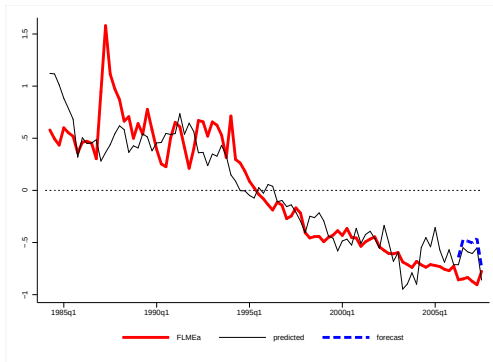
(c)



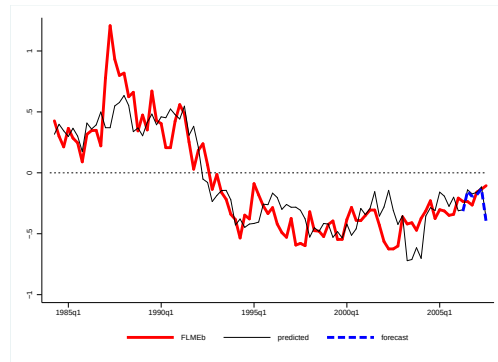
(d)



(e)

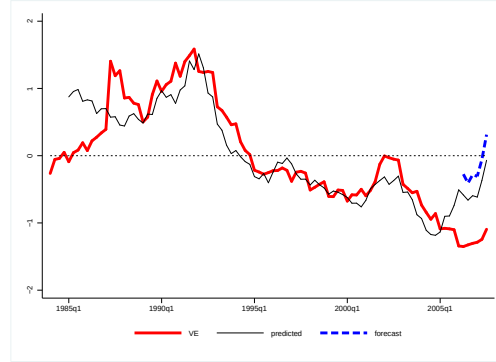


(f)

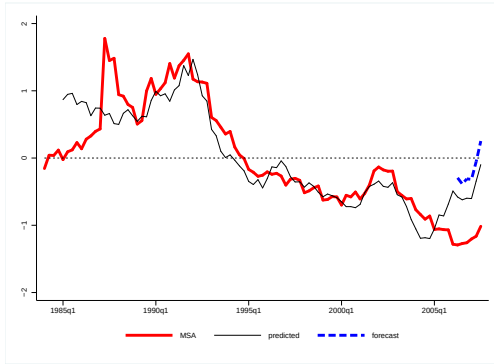


(g)

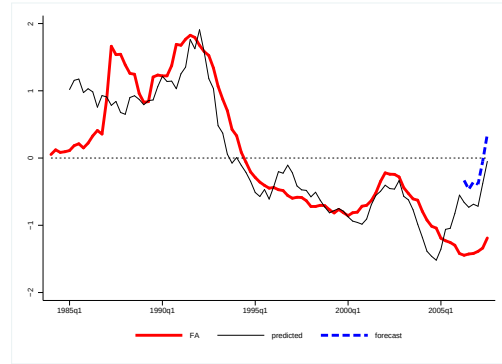
Figure 7: Forecasts robustness



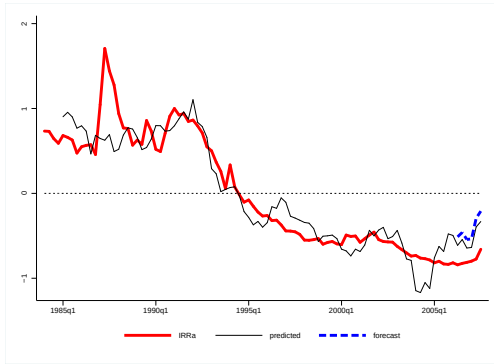
(a)



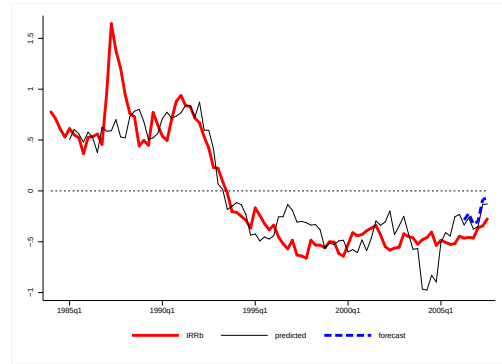
(b)



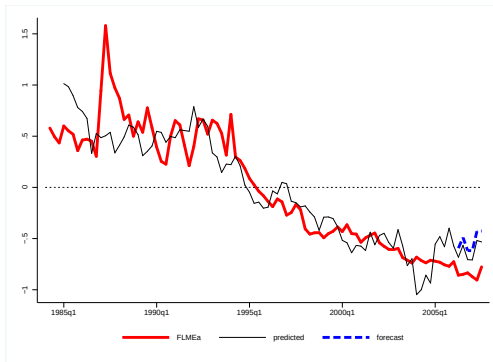
(c)



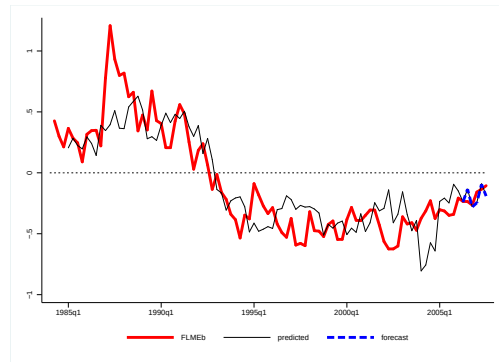
(d)



(e)

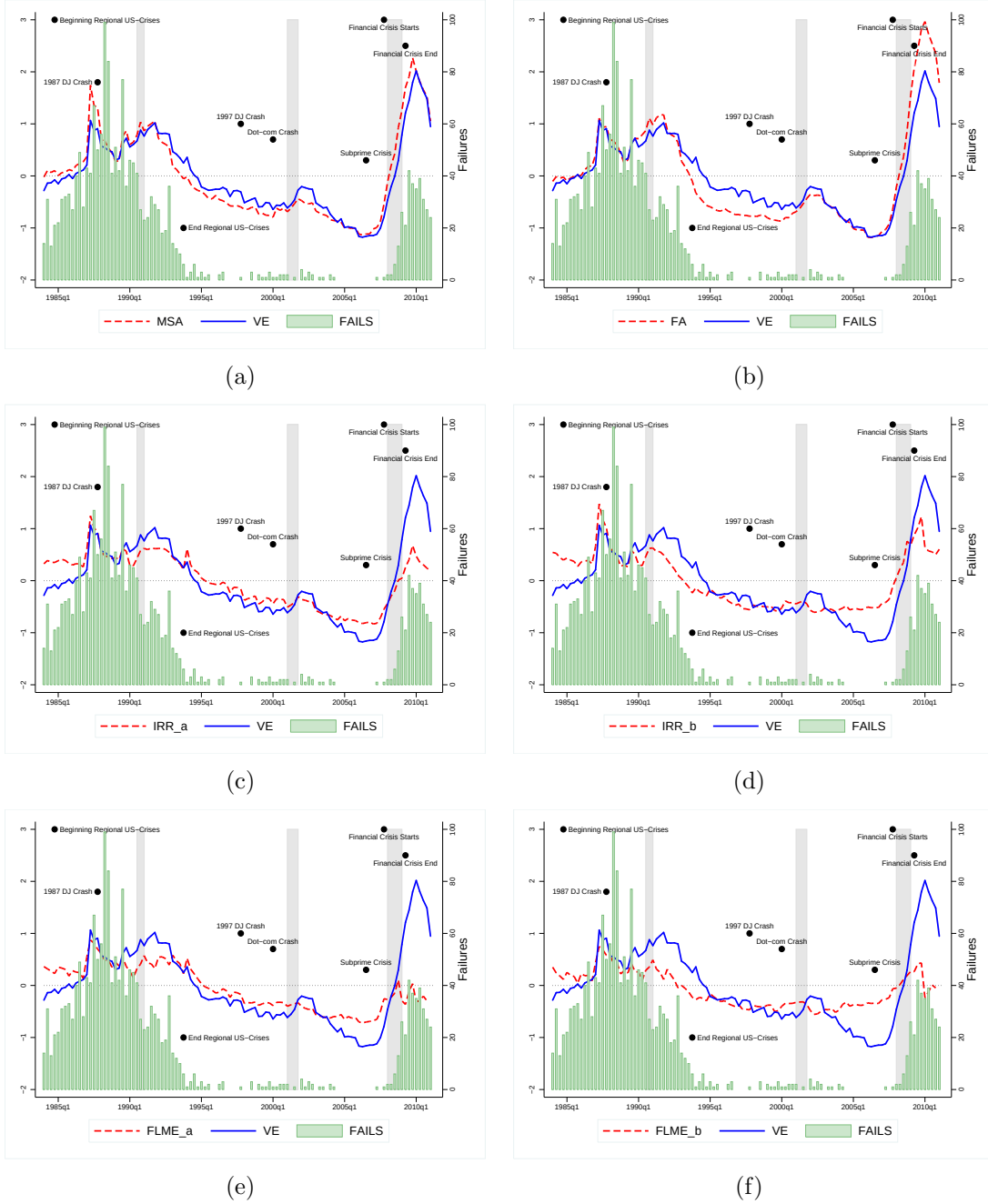


(f)



(g)

Figure 8: Indexes and economic and financial events -Update-



Notes: Due to the fact that the time when the data have been collected was prior to the 2007-2009 financial crisis, in this part of the Appendix, we report the graphs referring to the indexes discussed in this paper by taking into account also the data posterior to 2007:Q4.

The main results of the baseline analyses do not change: the best indicator proves to be the *FLMEb*, even if the last financial crisis has been better captured by the VE, MSE and FA indexes.