University of Messina & University of Catania

DEPARTMENT OF ECONOMICS

Ph.D. Course in Economics, Management and Statistics

XXXI Cycle

ESSAYS ON CREDIT QUALITY AND MACROECONOMIC ENVIRONMENT: NON-LINEAR MODELS AND FORECASTING

Candidate:
Maria Grazia Fallanca

Supervisor:
Prof. Edoardo Otranto
Co-Supervisor:
Prof. Antonio Fabio Forgione

A.Y. 2017/2018

Acknowledgements

Primarily, a special mention goes to my enthusiastic supervisor Edoardo Otranto. I would like to thank him deeply for taking me on as a student, for his guidance throughout the Ph.D., for the generosity with his time and for his valuable insights; they improved my work considerably. Without his continuous academic support and constructive advices, I would not have been able to accomplish the Ph.D. It was a privilege and honor for me to benefit of his relevant talented scientific knowledge and rare human qualities. I could not have chosen a better supervisor!

I also would like to thank my co-supervisor Fabio Forgione who sparked my initial interest in research when supervising my graduation theses. I am thankful for his encouragement unquestionably determinant for the realization of the thesis. He ceaselessly supported me with my work and reassured me when I was down, believing always in me.

Many thanks to my friends and colleagues of the Department for the enjoyable and friendly working atmosphere as well as for the leisure time.
I am much obliged especially to Alessandra Insana and Elena Pagliaro for their companionship, prompt help, and support whenever needed in difficult moments. With them, I have shared moments of deep anxiety but also of great excitement. They have become good friends over the years and I look forward to continue our relationships.

Last, but not least, I would like to thank to my family, my parents Rosario and Giovanna and my brother Tonino for all they have done for me and for being proud of me every day. Without them, I would not have come this far. I dedicate the thesis to them.

You see things; and you say "Why?"

But I dream things that never were; and I say "Why not"

George Bernard Shaw

General Overview of the Dissertation

In all countries, the banking system takes a central role in the economy being the key engine of growth. Financial crises, especially the turmoil of 2007 have underscored that, independently of the stage of development, all countries are susceptible to shocks and understanding the channels between credit risk and business cycle is crucial for evaluating banks robustness and system stability.
Actually, every weakness of the sector and the deterioration of bank's loans quality arises episodes of costly banking system distress, correlated economic crisis, and damage growth prospects. As soon as banks stop functioning normally, they cannot longer provide credit to the economy. Therefore, a prompt recovery is impossible without the awareness and a good understanding about the drivers of credit risk necessary to implement the suitable resolution instruments.
Generally, linear models have depicted the relation between credit quality and macro environment, but the strong dependence of bank loan default on the economic cycle, subject to changes in regime, could suggest nonlinear approaches as more appropriate.

In this regard, the broad object of this thesis is to extend and apply nonlinear models to this relation by developing three self-contained chapters

Chapter 1 lays the groundwork for the next two by reviewing non-performing loans literature along with the literature on Markov Switching and variants model, discussing properties and estimation procedure as well.
The work outlines a critical presentation of the extant ways to estimate the nexus between economic activity and asset quality, individuating some critical lacuna as regard the methodological procedures.
Findings suggest that the existing literature considers diverse sample, geographical area and time framework but fail to address regime shifts in the data and nonlinearities. Specifically, it recognizes the shortcomings of linear methods and a slightly lower number of applications deviated from those approaches. Although a growing empirical literature is using threshold models, there lack empirical works on the determinant of bad debts that employ the Markov regime-switching framework suggested as an appealing research tool to such demand and as a cross-validation method for the robustness of the results.

Chapter 2 focuses on the validity of the Markov switching framework in USA.
It attempts to model and forecast three different kinds of bank loan default, detecting non-linearity and asymmetries in their relations with macro variables by the adoption of a Markov Switching approach.
By comparing this specification with a classical linear model, empirical results lend support for the validity of the non-linear model in capturing the presence of regimes and asymmetries, changing in correspondence of the major recession periods spanning from 1987 to 2017.
Moreover, it gives evidence of a clear outperformance of the Markov switching concerning the linear counterpart, both in modelling and forecasting.

Finally, chapter 3 studies the evolution of correlation between total US delinquency loans and macro variables using the Dynamic Conditional Correlation model (DCC). To our knowledge, it makes the first attempt to provide a time varying analysis of correlation in our area of research. Results document that the dynamic correlation does not increase greatly during financial turmoil except for few variables and in lagged terms. Likewise, it details how correlations change over different time of crises (more specifically the Saving and Loans crisis in 1989 and later, the Dotcom in early 2000s and the Subprime crisis in 2007-08).

TABLE OF CONTENTS

Chapter 1

Modeling Credit Quality: Review of the Theory and International Experiences

Abstract

Exposure to macroeconomic shocks strongly affects the performance and profitability of the banking system besides threatens bank capital and solvency.
In consideration of the critical role of banks and financial markets in the economy, recognizing the pivotal macro variables is valuable for banks as well as regulators.
Sound banking conditions, indeed, are relevant for both economic growth and financial stability given that functioning banks endure financial shocks and facilitate the intermediation of flow of funds from unproductive to productive parties in the economy.
Empirical researches on the effects that macro variables cause on banks performance indicators cannot leave previous literature out of consideration.
In this section, we review the drivers of nonperforming loans (NPLs henceforth) that have been discussed in the actual field of research papers, attempting to highlight achievements and limitations as well as proposing possible applications and implications for future works.
We also include an econometric review of the literature about nonlinear time series models, not considered in the non-performing loans literature, which, in our view, have exceptional potential for modeling and forecasting the dynamics of NPLs.

Keywords: NPLs, credit risk, determinants, econometric modeling, forecast.

1. Objective of the study

The current work aims to examine, conceptually and methodologically, the impact of macroeconomic indicators on credit risk amount, measured in term of NPLs and similar default rates.
The latter have always characterized the banking system but have become subject of renewed interest among investors and workers in the sector especially in recent years.
After the surge of the 2008 financial crisis and the subsequent economic decline, the deterioration of the borrowers' creditworthiness increased and the phenomenon exploded reaching substantial levels.
At the end of 2016, the average NPLs ratio around the world was the 7,07 percent of all banks loans with the highest value in San Marino, Greece and Ukraine (43,38%, 36,30%, and 30,47%) and the lowest in Canada and Hong Kong (0,60% and 0,85)[1].
Regarding intermediary business, credit risk, or that of counterparty default, is the leading risk category because it is linked strictly with the health of the banking industry and the performance of the whole economy. Indeed, credit risk and NPLs reveal the soundness of banking sector by mirroring the quality of the loan portfolio of banks and, in aggregate terms, the quality of the overall sector of a country or region. Besides, the real economy and the financial sector are interrelated closely as a healthy economy and its development depend on the contribution of the banking sector.
Seen from a worldwide perspective, one of the critical causes of financial vulnerability and economic stagnation is represented by the percentage of bank's NPLs on their total assets, both in developing countries and in developed ones, as literature shows.
Quirici (1996) well stated that the non-pathological functioning of an economy's financial system is influenced by the quality of bank loans at both the micro and macroeconomic levels. At the microeconomic level because the presence of a significant amount of NPL seems to be the effect and the cause of inefficiencies. Loans or better revenues generated from interests constitute the primary source of income and therefore, banks that handle qualitatively good loans can operate more profitable than those who do not act in similar circumstances being able to repay the deposit amount and to retain their financial balance, critical to guarantee solidity and future viability. At the macroeconomic level, instead, the highest quality of loans supplied is reflected principally on the financial system stability itself, as financial intermediaries do not operate in isolation. They are part of a strongly interconnected system, in which the crisis of an individual bank or group of banks can spread through interbank exposures due to a domino effect, and thus trigger a systemic crisis.
Modern history provided not few examples of how failures of financial intermediaries have been caused increasingly by a reduction of the lending standards with subsequent excessive and uncontrolled exposures (against overestimated resources or nothing) to creditors scarcely deserving and in extraordinarily difficult circumstances. For instance, the case of the no documentation mortgages and the case of the notorious NINJA loans, whose acronym stands for no income, no job or asset. Both the global financial crisis with the fall of the American "giants" and the later great recession that collapsed the economy of the whole world are an example of how the credit risk occupies undoubtedly a prominent role among the various drivers of the financial crisis. They symbolize how banking crises are often born from a massive accumulation of nonperforming loans, inefficiencies in the identification, measurement, disposition, and management of risks.
For this reason, the focus on credit quality is empirically justified by the size of the problem around the world and the cause for NPLs should be given due consideration to avoid risk proliferation on the banking sector. A deep awareness of the main reasons of these problems will help to minimize and control defaults but, generally, to promote economic recovery and a sustainable growth giving some insights on how tailor macroeconomic policies and regulatory counteractions.

[1] Source: World Bank.

Moreover, in spite the efforts in the post-crisis period across supervisory institutions to establish immediate actions intended to combat and mitigate the issue that have contributed in the decline[2] significantly, NPLs levels remain higher, with differences across time and from one country to another and debtor categories.
As an illustration, the graph below depicts the international NPLs state of play in an aggregate manner.

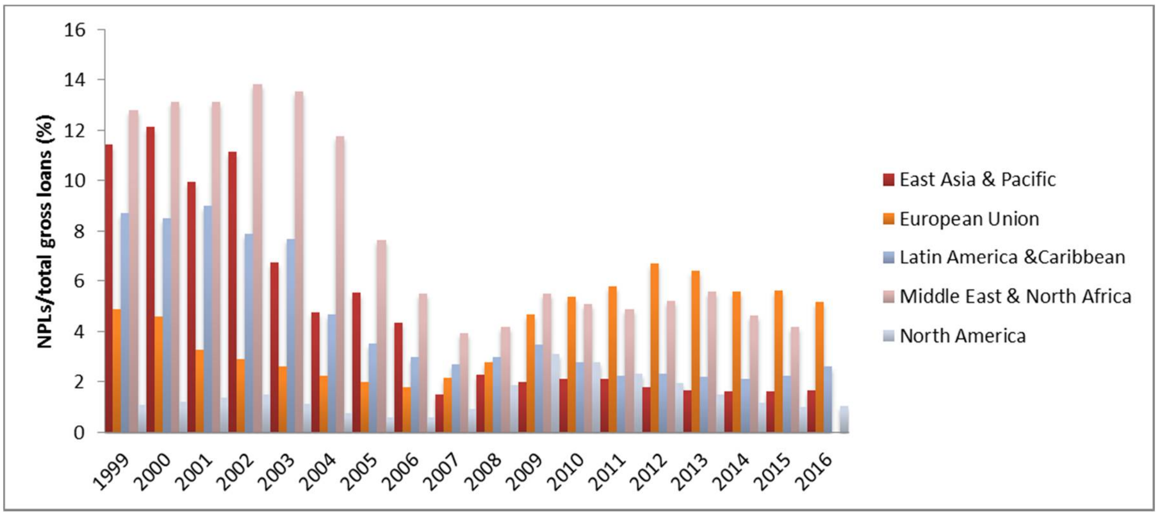


Figure 1-1 NPLs evolution. Source: World Bank

What emerges is a different course of NPLs, with Europe currently most embroiled in the phenomenon[3] even due to its policymakers delay in addressing banks problems and due to the European financial system largely dominated by banks.
At the end of 2016, the European bank balance sheets held € 1062 billion of gross NPLs, about the 5,4% of the total loans[4], with an high dispersion across countries from 1% of Switzerland to 47% of Greece and Cyprus passing through the 10%-20% amount recorded by Portugal, Italy, Ireland, Hungary, Bulgaria, Romania and Croatia. As stressed by the EBA in the indicated report, the Euro area NPLs ratio, although improved compared to the 6.5% of December 2014, remains historically elevated concerning other countries.
By making a comparison between Europe, the United States and East Asia, we can notice that between 2007 and 2009, the ratio rapidly increased in both Europe and the United States. Then, in the period between 2010 and 2015, it can be observed a trend change in the United States alone, which together with East Asia at the end of the reference period have the NPL ratio below 2%[5].
However, the common backdrop of a financial crisis represents similarities between ex-ante and ex-post 2007 paths: the South East Asian crisis at first and the subprime crisis later.
The former had minimal impact on Western economies and was characterized by a fast recovery in the aftermath, the latter, on the contrary, required a prolonged resolution and greatly affected many countries throughout the world, especially the U.S and Europe.

Keep in mind that, this section aims to scrutinize the macro key determinants of loan portfolio credit risk, analyzing scientific publications.

---

[2] Think about insolvency reforms, the strengthening of surveillance, capital controls or the construction of dedicated bad banks where isolate toxic loans from the ongoing activity.
[3] The evolution of Europe aggregate NPLs displays a downward tendency and a significant improvement until 2007 when reversed brusquely.
[4] Source: European Banking Authority (EBA), Risk Assessment of the European Banking System, December 2016.
[5] According the World Bank data, the NPLs ratio for the United States was 1, 5% and 1, 3 in December 2015 and 2016. The East Asia registered 1,7 point percent in 2015.

2. Literary review

There is a well-established frame of international literature that has addressed the problem of bank credit quality due to either the importance for the soundness of banks and the observed link between the rise of NPLs or current and past crises[6].

Academic studies started from few years of the last century and vast and diverse papers (both theoretical and empirical ones) have explored the linkage among NPLs and bank-specific factors or microeconomic or idiosyncratic variables, as well as the relation between NPLs and macroeconomic aggregate. Despite there seems to be an increasing awareness that both micro and macroeconomic factors can affect nonperforming loans, literature focused initially on either micro or macro determinants, more recently on both. First, among factors cited as significant determinants, the size of the bank, the composition of the loan portfolio, the profitability, measures of the power of diversification, capital adequacy, and bank's balance sheet data are well known. See e.g. Berger and DeYoung (1997), Salas and Saurina (2002), Hu *et alt.* (2004), Stern (2004), Fofack (2005), Khemraj and Pasha (2009), Greenidge and Grosvenor (2010), Louzis et alt. (2011), Klein (2013), Suryanto (2015), Chaibi and Ftiti (2015).

Generally, they express NPLs as a result of bank misbehavior or peculiar weakness. The macroeconomic or external variables that are likely to disturb the borrowers' capacity to repay loans are general situational factors in which the banking sector operates, like indicators of the financial market and indexes of the direction in which the economy moves. They include the gross domestic product and the rate of industrial production, competition between banks, regulation, the trend of share prices, the employment rate, the stock index, the interest rate, and so on. See e.g. Keeton and Morris (1987), De Lis *at alt.* (2000), Gambera (2000), Arpa *et alt.* (2001), Salas and Saurina (2002), Ranjan and Dhal (2003), Fofack (2005), Hoggarth *et alt.* (2005), Quagliariello (2006), Filosa (2007), Marcucci and Quagliariello (2008), Louzis *et alt.* (2011), Castro (2012), Saba (2012), Nkusu (2011), Klein (2013), Beck *et alt.* (2013), (Filip, 2015), Kjosevski & Petkovski (2016).

The most common approach of the studies involves estimating on historical data the resilience and sensitivity of banks' balance sheets to adverse changes in macro fundamentals. Large-scale researches confirm that macroeconomic variables can influence loan portfolio quality and, whatever the macro cause is, it will reduce severely the banking system's stability; they claim that business cycle does not relieve banking system.

In effect, good economic conditions seem to be consistent and positive correlated with a good credit market and loan quality. Concurrently, there is evidence of a mutual interdependence and therefore about a strong relation between NPLs and economic growth as well. On the theoretical front, Levine and Zervos (1998) summarised the debate about the nexus between banking system and economic development evoking pioneering authors such as Walter Bagehot and J. Schumpeter according to which banking sector and financial services can strongly stimulate innovation and economic improvement by funding productive investments, mobilizing savings and matching supply and demand of funds. Thus, broadly speaking, well-functioning commercial banks hasten economic growth because banks can be described as a "heart which pumps blood flow in the form of capital" to the economic lifeblood (Messai and Jouini, 2013) (Suryanto , 2015), while NPLs and an unstable banking system are likely to hamper economic development. They harm bank ability to lend to the real economy and to resist risks.

Better, regarding the implications, high levels of NPLs affect adversely a bank's profitability because of provisions to restore balance sheet conditions and an adequate coverage ratio in case these loans need to be written down. In short, even if banks normally set aside stocks of money to compensate potential losses on loans, NPLs require higher provisioning needs that affect ultimately bank expected returns, depressing their amount compared with returns on performing loans. In addition, NPLs reduce bank's profitability as they are "zero yield asset" (Parven, 2011) not contributing to the interest income of bank and its available loan funds for a sustainable intermediation between depositors and debtors.

[6] "The immediate consequence of large amount of NPLs in the banking system is bank failure as well as economic slowdown" (Lata, 2014).

Moving on, NPLs heighten operating and funding costs. Operating costs are caused by the expense of monitoring distressed loans and borrowers, funding costs are due to the fact that lenders are less willing to lend to banks with high NPLs levels. Finally, problem loans restrict the scope of credit supply for new lending to private and, in wider terms, they shrink the financing to the real economy (credit crunch) as they are risky assets that require higher levels of capital and adsorb higher risk weights than assets that are performing (European Parliament , 2016)[7].

As stated above, this chapter addresses the first issue, elaborating the guesswork that macroeconomic factors are key drivers of a degenerating asset quality. First, it describes shortly the NPLs concept; second, it gives a dense critical survey at a global level of the most influential empirical researches and econometric methodological approaches developed in the field. A third section attempts to shed light on the main weakness of previous studies and to suggest directions for future researches, incorporating an econometric examination of suitable models.

2.1 What are the Non-performing loans?

Prior to embarking on the analytical framework, a critical starting point consists of giving the essential working definitions.

Intuitively, a loan is nonperforming when its quality has deteriorated to a varying extent and it has stopped to perform.

For those who are not acquainted with the NPL problems, definitions of NPLs have often been a source of confusion because at what point a bank classifies a loan as non-performing and when it becomes bad debt, it depends on local regulations; an internationally harmonized standard at a practical level is lacking. Variations exist in terms of the classification system, scope, and content, and cross-country comparisons are difficult.

The table below overviews some international definitions.

The one highly accepted and used in banking practice all over the world, which helps to increase cross boarders evaluations invokes the quality differentiation of the loans granted by credit institutions in five descending categories. In accordance with the proposal developed by the Institute of International Finance (IIF), these are divided into standard, watch, substandard, doubtful and loss; and the weakest three typically correspond to NPLs. The quality level of each loan emitted is the highest if it belongs to the "standard" category; it is completely unsatisfactory if it is probably to produce losses for banks (Krueger, 2002).

Generally admitted, the positioning within the previously mentioned classes is based primarily on the criterion of the number of days of delay in payment from the maturity declared in the contract. And in this vein, Parven (2011) citing Choudhury et al. (2002) defined NPLs as “multiclass” concept rather than a “uniclass”, meaning that NPLs can be divided into different categories based on the “length of overdue” of the said loans.

According to one of the most influential sources, the loan is nonperforming when payments of interest and principal are delinquent by 90 days or more (IMF: Bloem and Freeman, 2005). Similarly, as stated by Bholat

---

[7] This refers to the Basel rules on how bank must calculate their regulatory capital: as a percentage of a bank’s credit risk weighted exposures. Weights are risk sensitive and lower rating exposures require higher risk weights than enhanced credit quality in the calculation of the capital adequacy ratio. The credit institution's capital is the "cushion" for potential losses, which protect the credit institution's depositors or other lenders. Basel I adopted as measure of capital ratio the capital (in the numerator) against the bank’s assets (the denominator), where bank capital was broken down by tier 1 capital and tier 2 capital and assets had basic adjustment for their risk level by 4 categories (0%, 20%, 50% and 100%). To be considered sufficiently capitalized under Basel I, a bank had to maintain a capital ratio of 8% (bank’s capital had to be not less than 8% of the value of the bank’s risk-weighted assets.)

Basel II explored the deeper risk-weighting approach with 2 extra risk categories – 35% and 150% and determination of assets by risk utilizing the estimation of credit-rating agencies, such as Standard & Poor’s, Moody’s, etc. Total regulatory capital under Basel II equals Tier 1 capital + Tier 2 capital - deductions (if there are any) + Tier 3 capital (if there is any). To be considered sufficiently capitalized under Basel II, a bank had to maintain a capital ratio of 8%.

Basel III deeper expands Basel II. The primary goal of Basel III is to improve the ability of banks to absorb asset losses without affecting the rest of the economy. In terms of capital regulation, Basel III focuses mainly on the quantity and quality of capital held by banks. The most important part of Basel III is its new definition of regulatory capital, which now includes common equity Tier 1 capital + additional Tier 1 capital + Tier 2 capital - deductions (if there are any). The minimum CAR set at 4.5% Common Equity Tier 1/RWAs; 6% Tier 1 capital/RWAs, and 8.0% total capital/RWAs. To be considered sufficiently capitalized under Basel III, a bank had to maintain a capital ratio of 10.5%.

et al. (2016), a non-performing loan is a financial asset over which a borrower fails to comply with the original contract. Specifically, a loan is nonperforming when a sum of borrowed money has not been giving back for at least 90 and the odds that it will be repaid in full falls below the contracted value carried on bank's balance sheet. However, if the debtor, at some time point, starts making payments again on a nonperforming loan, it becomes a re-performing loan, even if the debtor has not repaid all the missed payments.

Beside this criterion of a change in contract terms[8], qualitative and judgmental characteristics such as the overall financial creditworthiness of the borrowers and portfolio differences represent additional indications and predictors of losses.

Barisitz (2013) makes this point. Comparing the definitions in ten Central, Eastern, and Southeastern Europe (CESEE) countries, he used the quantitative parameter of the number of days of overdue planned payments as the main yardstick to define impaired loans. In any event, he also stressed on the existence of a "well-defined weaknesses" of either the loan or the borrower as an additional characteristic of NPLs, meaning a bank's appraisal of the borrower's economic and financial standing. In this case, the borrower' creditworthiness or ability to honor his or her obligations is not assured, some losses are probable and the 'orderly repayment of the debt is in jeopardy' (IMF: Bloem and Freeman, 2005)[9].

Likely, given that the use of various NPLs definitions and non-identical accounting practices make it difficult to compare the situation in different European Member States, the European banking authority recent uniform definition of NPLs is based on the "past-due" criterion and on the "unlikely-to-pay" criterion (145 of Annex V of the EBA ITS).

Banks have to classify exposure as non-performing, without taking into account the existence of any collateral, when they satisfy either or both of the following criteria: exposures are more than 90 days past-due and they are unlikely to be repaid in full, regardless of the existence of any past-due amount or of the number of days past due.

Hence, criteria refer at both the debtor's past and presumed future performance based on external indicators such as a registered bankruptcy, and on banks' internal judgements that require clearly defined criteria as well.

Table 1-1 A comparison of NPLs cross-country definition. Own elaboration on National Banks information.

Country	NPLs definition according to National Central Bank
Belgium, Czech Republic, Hungary, Latvia, Luxembourg, Romania, Paraguay, Peru, Slovakia, Slovenia	Non-Performing Loans are defined as loans overdue for more than 90 days.
Estonia, Lithuania, Uruguay	Non-Performing Loans are defined as loans overdue for more than 60 days.
Norway Philippines	Non-Performing Loans refer to a loan whose payment of principal and/or interest is 30 days in arrears.
Portugal	Non-Performing Loans are loans not settled within 30 days after the expiration date, as well as future payments if there is a doubt for their collection.
Spain	Delinquent Loans are simply defined as loans in relation to which there is a reasonable doubt regarding full repayment of principal or interest.

[8] Consider, anyway, that threshold depends on national methodology.

[9] See also empirical works: Saba, Kouser and Azeem (2012) defined NPL as one in the position of default or close to default; Filip (2015) referred to as a distorted, undesired and harmful credit relationship, in which there is a daft block of lending resources or losses for the bank. Berger and DeYoung (1997) qualified them as "problem loans" since these can be considered "bad or toxic assets on the bank's books" (Bexley and Nenninger, 2012); Shu (2002) argued efficiently about criticised loans and Parven (2011) about a "byproduct of financial crisis".

United States	Delinquent Loans or Non-Performing Loans are those past due of over 30 days and still accruing interest as well as those in nonaccrual status.
Russian Federation	Non Performing consist of: 1) payment failure at least once for the last 180 days if the overdue period is more than 30 days for legal entities and 60 days for personal loans with a good evaluation of debtor's financial situation; 2) payment failure at least once for the last 180 days if the overdue period is more than 6 days for legal entities and 30 days for personal loans with average evaluation of debtor's financial situation; 3) all loans with bad evaluation of the debtor's financial situation.

2.2. The determinant factor of NPLs

2.1.1. Macroeconomic variables

Concerning the aforementioned link between credit defaults and economic conditions, there is significant empirical evidence of a close negative relationship between NPLs and the growth in real Gross Domestic Product (henceforth GDP), the major proxy of business cycle; a positive with interest rate and unemployment, while evidence on the effect of inflation is mixed.

This section covers a review of the empirical findings.

Following the words of Reddy (2012), GDP is a measure of economic growth in terms of an increase in the size of a nation’s economy. It is a broad measure of an economy’s total value of final goods and services produced within a country's borders in a year and consumed by its final users, not used as an input into other goods, regardless of ownership.

The explanation of the negative correlation between GDP and NPLs that was given by literature lies on the fact that a rise in GDP mirrors positive economic environment favorable for both businesses and households, implying habitually a higher level of income. During beneficial economic conditions, the revenue of households and companies grows and borrowers have enough funds to respect their debts in predetermined deadlines more easily, contributing in turn to lower bank bad loans. Vice versa, a downturn in the economy expands problem loans’ volume, and hence capital requirements because unemployment rises, disposable income diminishes and borrowers have trouble paying their obligations back. In other words, during economic decline, the growth rate of loans goes down, while that of non-performing loans rises due to the worsening financial conditions of borrowers.

One of the earliest and important studies in which we can track down the origin of the idea that macroeconomic aggregators such as GDP can explain changes in defaults rates by influencing the creditworthiness of the obligors is the work of Wilson (1998): “all credits can potentially become bad over time given a particular economic scenario”.

Subsequent important researches include the theoretical contribution of De Lis, Pagés, and Saurina (2000). They focused on the cyclical behavior of bank credit and loan losses in Spain but also considered the provisioning policies over the business cycle[10] as the dependent variable, giving a more complete understanding of credit risk.

[10] Many researches did not focus on just a single measure of asset quality and bank fragility. Even though NPLs is the typical proxy of credit risk, a number of other alternative indicators of loan performance have been used as substitutes such as loan loss provisions (expenses that bank supervisory authorities order banks to store against expected losses), loan charge off (Saba, Kouser and Azeem, 2012), delinquent loan, credit default swap spread, borrower’s default rate. The choice was mainly based on the availability and nature of data, as NPLs usually are available on annual basis.

Specifically, the authors combined the two competing theories about the pro-cyclical (economic and financial cycles commove) and countercyclical banks behavior to represent credit quality dependency from the phase of the cycle.
In their words, there is a strong relationship between the default loans and the economic cycle, which the authors found by considering the distribution of the problem loans ratios during two different positions of the downturn and strong growth. The average quality of borrowers is higher in economic booms when income grows, and poorer at the trough of economic activity because of firms and households' financial distress. Accordingly, banks provision policy is pro-cyclical as banks provision only when NPLs materializes (during unfavorable economic periods and not before), thus magnifying the effects of the negative phase of the business cycle. However, the increase in bad debts as a consequence of recession alone cannot be justified but, it is grounded in unselective and inadequate bank's lending policies built up during business cycle upturns as well. That is to say that both the reliance on the macro environment and individual institution's credit policy affect asset quality, and the GDP has a two-tier impact, one related to a lift in loan growth usually associated with greater GDP and the other related to borrower's higher income.
Indeed, as expansionary phases continue and borrowers are temporary liquid, banks are involved in a strong fierce competition resulting in credit growth and credit risk mistakes. To acquire market share and expand loan portfolio, banks tend to relax their credit selection criteria (bank lending is pro-cyclical), charge lower interest rates, show greater risk appetite and optimistic expectations, they grant under guaranteed credits and lower quality debtors, engage excessive risk-taking activities, reduce the provision for future losses.
All this overheating, keep saying De Lis, Pagés and Saurina (2000), implies adverse repercussions like higher credit risk and financial imbalances that accumulate during periods of economic booms when banks loosen credit standards but they only materialize during period of recessions, causing the growth of nonperforming loans and loan losses[11] that again reinforce economic cycles. The effect can only be weakened by an increment of provisions also at the time of booms (countercyclical view of provisions) [12].
Much of subsequent empirical works tested the accuracy of these conclusions using a variety of methodological approaches and considering other variables in addition to GDP. It is understood that typically, empirical studies go hand in hand with methodological concerns.
On this premise lies the works of Salas and Saurina (2002), an obligatory reference of credit risk literature. Authors shaped the problem loans in Spain and estimated the impact of the cycle and the related loan growth policies using a panel data of commercial and savings banks from 1985 to 1997.
The former indicator, measured through the current and one year lagged GDP growth rates, had a negative and significant impact on banks portfolio of both type of banks; with the current impact much more considerable, suggesting therefore that macro shocks quickly spread to bank's balance sheets. The latter showed a significant and positive hit with a lag of three years, supporting the idea that, as a result of the business cycle, periods of loan growth precede periods of high loan losses: asset quality deteriorates in response to positive loan growth due to lax credit standards applied during the boom period.

[11] The forerunner of all empirical works, Keeton (1999) examined the impact of easy credit standard and credit growth on loan delinquencies in the US commercial banks, asserting early a positive relation between loan growth and loan losses due to a shift in the supply of bank credit. Loan growth resulting from a superior willingness of banks to lend and associating with a falling in credit standards leads the likelihood of loan losses to rise.
[12] Bikker and Metzemakers (2002) provided a meticulous explanation about the two alternative views about the impact of GDP growth. "The common view is that an economic upswing and rising incomes indicate improving conditions for firms and reduce the likelihood of loan defaults, whereas a recession will have the opposite effect. Banks are expected to reflect this feature in their decisions by lowering provisions during an economic boom and increasing them during a downturn… According to this common view, the banks' provisioning behavior is pro-cyclical, meaning that it reinforces the current development of the business cycle. However, an alternative, countercyclical view states that credit risk is build up in a boom and materializes in a downturn. The favorable conditions of an economic expansion could lead to an excessive increase in credit lending and a less critical assessment of creditworthiness. The countercyclical view associates this with higher risks and the build-up of financial imbalances that increase the likelihood of economic contraction. According to this view, provisions should be positively correlated with the business cycle, for banks should recognize this cyclical pattern of credit risk and build up loan loss reserves in good times to be drawn on in bad times". (Pages 4-5).

Similarly, in Italy, Quagliariello (2006) estimated the cyclical pattern of the Italian bank borrowers' new bad debts and loss provision over the period from 1985 to 2002, employing both static fixed-effects and dynamic[13] models as econometric methodology:

$$y_{it} = \alpha_i + \beta x_{it-j} + \gamma d_{t-j} + u_i + \epsilon_{it} \qquad static\ fixed\ model\ (1)$$

$$y_{it} = \alpha_i + \beta x_{it-j} + \gamma \delta_{t-j} + \mu y_{i,t-j} + u_i + \epsilon_{it} \qquad dynamic\ model\ (2)$$

where x_{it-j} and d_{t-j} represents the $it-j$ observation of bank-specific variables and macroeconomic indicators, $i = 1, \dots, N$ and $t = 1, \dots, T$ the observations and the time, j the lag length depending on the variable; u_i the individual-specific, time-invariant effects fixed over time and $\epsilon_{it}$the error term; $y_{i,t-j}$is the lagged dependent variable added among regressors.

Quagliariello (2006) included GDP, interest rates (10 years T-bond), the evolution of the stock exchange index, the spread between loans and deposit rates as explanatory variables and as main indicators of business cycle and credit markets conditions, respectively. He considered explanatory variables contemporaneous and one lag (two for GDP) value to understand the length of time with which these variables have an impact damaging credit quality.

Again, the author confirmed the backward-looking feature of provisions (provisions are not dynamic) and ascribed the pro-cyclical trend to reasons like "disaster myopia[14], over-optimism and herd behaviors[15][16]". Specifically, following the author, at the beginning of an expansionary phase of the economy, banks overlook future losses, as they are overconfident about investment projects and the ability to recover loans, expect lower nonperforming assets, relax lending standards and cut provisions for prospective losses, deepening the impact of the business cycle. Consistently, bad debts behavior is cyclical, decreasing in good macroeconomic times and increasing during downturns. Author found a negative delayed effect of 1 and 2 years of GDP changes that increases in bank's size (suggesting that large banks are more affected by the fluctuations of the real economy) and a positive sign of interest rates in both models, the static and the dynamic, that implied its significance in affecting debtor's ability to pay back debts. Spread indicator, on the contrary, resulted insignificant in the first model, positive, and significant in the last one in accordance with the debt burden theory. Finally, observations suggested the asymmetric effects of macro shocks during recession and expansion phases: "in recessionary phases creditworthiness seems to deteriorate more heavily than it improves in expansions"[17]; and including two dummy variables, down and up, to interact with the GDP growth, he obtained the estimation.

---

[13] This approach rotates around the Arellano-Bond estimator. It is a difference-generalized method of the moments (GMM difference) based on first differences of the regression equation rather than instrumental variables estimation. In fact, the inclusion of lagged values of the dependent variable as regressors in a dynamic model contravenes the moment condition of exogeneity (error terms have to be uncorrelated with each regressors $E(X'u) = 0$) and gives inconsistent and unbiased results. In the Arellano method, first difference removes fixed effects and lags are used as instruments. Theoretically, the involvement of the dependent variable lagged value is justified by the fact that in economics, many relationships are dynamic and problem loans are not instantly written off but continue to be in balance sheet for a long time.

[14] See (Guttentag and Herring, 1984) pages 1365: "The tendency for subjective probabilities to fall below actual probabilities during periods in which no major shocks occur". According to this hypothesis, banks suffering and lending policy errors can be attributed to bank's manager difficulties in measuring the time aspect of risk so that they underestimate the probability of adverse future outcomes during good times but they overestimate it when economic environment deteriorates. This is based on personal judgment and on the idea that the longer the period without a shock the lower the subjective probability attributed to it.

[15] Herding behavior is the imitation of other investor behaviors. It means a situation where decision makers disregard their own private information and are influenced by other decisions. Specifically, here it refers at the imitation of bank's competitors lending behavior that causes concentration of lending and recessions.

[16] See (Gonzales, 2009) for a clear explanation about the definition of pro cyclicality and its reasons.

[17] In this paper appears the issue about the asymmetric effect of business cycle over different regimes and it has been proposed as a further possible theme of research.

This result was coherent with researches conducted in the Greek banking sector by Louzis *et al.* (2011) who used dynamic panel data methods to make a comparison between different type of loan classes (consumer, business and mortgage loans) and, on the other hand, with the multi-country comparative analysis of Castro (2012) and Škarica (2014).

Louzis *et alt.* (2011), in fact, demonstrated that, for all loan categories, macroeconomic variables (GDP, unemployment, interest rates and public debt) motivate NPLs. The clearly different quantitative influence of macro variables among various types of loans indicate the sensitivity of each category: NPLs of consumer loans are more susceptible to change of the real growth and lending rates, NPLs of business loans are more susceptible to change of unemployment rate, while NPLs of mortgage loans are less receptive to change on macro conditions.

Castro (2012) used the same dynamic panel data approach of Quagliariello (2006) to focus on the GIPSI group of countries (Greece, Ireland, Portugal, Spain and Italy), extensively beaten by adverse financial conditions, high levels of public deficits and debts. Findings concluded the significant impact of macroeconomic variables, stating that credit risk increases significantly with a rise in unemployment, interest rate, exchange rate and credit growth, but even with a decrease in GDP, in share and house price indices.

Škarica (2014) extended Castro's analysis on the determinants of the NPLs of European emerging markets such as Bulgaria, Croatia, Czech Republic, Hungary, Latvia, Romania, and Slovakia. Using fixed effects estimators, he confirmed, once more, the results obtained from previous empirical studies regarding the cyclical nature of NPLs according to which economic slowdown is the basic reason of high levels of NPLs. Conversely, the effects of the financial sector through the share price index resulted not significant. It seemed not surprising and justified by the general economic conditions of each country characterized by underdeveloped financial markets and a high level of foreign currency loans. Beck, Jakubik, and Piloiu (2013) underlined that a drop in stock prices indices negatively impacts just the credit quality of countries with large stock markets relative to the size of the economy. Briefly, Beck *et alt.* (2013) too, showed interest in the temporal dynamics and in the uneven amount of non-performing loans in various contexts. They used a quite vast panel data set of 75 countries, both advanced and emerging economies, and many macro factors, from standards determinants of NPLs like changes in the economic activity to exchange rate depreciation and the performance of the stock market of a particular country. Implementing at first a static and later a dynamic estimation method to measure the effect of indicators and capture the persistence of the NPL growth, econometric results corroborated GDP growth as the main driver of bank asset quality for all the selected economies. Secondly, the impact of exchange rate resulted mixed: in countries with specific vulnerabilities, especially with a high degree of lending in foreign currencies, a fall in nominal exchange interest rate could be a determinant as it represents a depreciation of the domestic currency and would cause an increase of NPLs growth rate, rising debts service costs.

Other researches focused on the impact of the financial crisis on financial soundness indicators like asset quality and capital adequacy. Kasselaki and Tagkalakis (2013) revealed that a fall in the output gap in severe crises time leads to a rise of NPLs and bank's provisions more pronounced than those recorded in an economic environment just worsened. A dynamic panel data generalized method of the moments (GMM difference) for twenty advanced OECD economies estimated outcomes, in the period 1997-2009. Kjosevski & Petkovski (2017) choose the same estimation method to focus on Baltic States banks, with quite similar conclusions: bank specific, macroeconomic factors but also previous period amount of NPLs affect banks' asset quality.

In the same spirit, country-specific researches outside of Europe reaffirmed the association between macroeconomic variables and loan portfolio quality.

Ranjan and Dhal (2003), for instance, dealt with the Indian context to study the behavior of borrowers as regards the loan refunds in response to bank's lending terms of credit (for instance, the interest rate charged to different borrowers and business activities), bank size and macroeconomic indicators. Similar to previous

results, the author's panel regression analysis reported a significant impact of favorable macroeconomic conditions, measured again by the lagged growth rate of GDP, on NPLs.
Ranjan and Dhal (2003) too agree that good economic performances involve greater risks taking and loan losses from lenders viewpoint but at the same time, they reduce financial distresses from borrower's perspective, underlying in this way that borrowers and lenders react differently to business cycle states. "The expectation of higher growth reflecting favorable macroeconomic and business conditions has a negative influence on NPAs, suggesting that increased economic activity leads to lower financial distress of borrowers and thus, lower NPAs for banks" (page 111). Nevertheless, from the alternative prospect and considering how a change in the macro environment influences the bank's term of lending, GDP growth rate has a positive effect on NPA. The consequence is that the direction of macroeconomic shocks to loan losses is the result of the combination of debtors and bankers reactions.
Meanwhile, proceeding with the Extra-European background, Espinoza and Prasad (2010) investigated the presence of sudden unfavorable macroeconomic shocks in 80 banks of the Gulf Cooperation Council (GCC) region and the feedback effect of increasing NPLs on economic activity using a VAR model after performing the dynamic panel estimation. Again, we find an application of fixed effects, difference GMM and System GMM models to a set of macro and micro variables among which non-oil GDP growth and interest rate resulted in the most significant. Similarly, Ghosh (2015) used both fixed effects and GMM estimations using annual data for 51 US states for the period 1984–2013 to demonstrate that higher state real GDP, real personal income growth rates and changes in state housing price index reduce NPLs, while inflation, state unemployment rates, and US public debt significantly increase NPLs.
Later, in 2017, he used the same methodological approach to explore the issue at disaggregate and sector-specific level. Among the recent evidence that applied panel data analysis, Memdani (2017) and Filip (2017). The former focused on the Indian banking sector and investigated how the impact of macro-factors varies with respect to the bank ownership structure, revealing a significant effect of inflation against the formation of bad loans in public sector banks. The latter scrutinized the evolution of bad loans in some representative Central and Eastern Europe countries, founding a positive linkage with unemployment and the dummy crisis, and negative correlation with real GDP and inflation.

The above front of literature used principally panel data on bank-specific and macroeconomic variables. This means that the macroeconomic indicators have been included mainly as control variables and thus treated as exogenous. In a different manner, an additional strand of literature has relaxed the exogeneity of macro elements and it has employed the vector autoregressive (VAR) methodology to take into account the simultaneity and the feedback effects between business cycle and bank's balance sheets as well as the interdependencies between multiple macro time series:

$$y_t = c + \sum_{j=1}^{p} \Phi_j y_{t-j} + \epsilon_t \tag{3}$$

where y_t is a vector of endogenous variables also including the nonperforming loans ratio or similar default rates, c is a constant vector, Φ_j are coefficient matrices, ϵ_t is a vector of disturbance terms.
Keeton (1999) was the first one to use a vector autoregression model to analyze loan delinquency. However, using data from 1982 to 1996, he concentrated mainly on the relationship between easy credit standard (anyway connected, as we explained, with economic expansion phases) and the increase in delinquency rates, showing as an exceptional rapid loan growth contributed to higher loan losses in certain states in the US.
To our knowledge, the inclusion of macroeconomic variables in this methodology takes place a few years later. In this way, Fofack (2005) tried to understand whether the causes of NPLs were among the major fuels of the banking and economic crises which affected the Sub Saharan African countries in the 1990s.

Referring to all banks, public and privately-owned, he questioned these causes using correlation and causality analysis with a subset of macroeconomic variables such as GDP per capita, inflation, interest rates, changes in the real exchange rate, interest rate spread and broad money supply (M2).
In line with other studies, his work revealed that GDP best explains the dynamic of NPLs. Indeed, the result that a fall in the per capita income is related to an increasing amount of NPLs is consistent for most countries. Moreover, he showed a binary causality relation between real GDP per capita and NPLs: economic contraction deteriorates assets quality but, in other direction, accumulation of NPLs and banking crisis stoke economic decline. In addition, inflation and real interest rate are Granger-causal of nonperforming except for few countries.
The causal analysis was extended further by supposing variations in the given micro and macro variables over a prediction range and considering two alternative scenarios. The first case consisted in considering an improved macroeconomic framework in which GDP was assumed to grow higher than the average over the reference period, in the second scenario was assumed a lower per capita GDP over the prediction range.
The Granger Causality research results were validated by an unbalanced panel model to forecast of the potential impact of macroeconomic variables on the dynamics of these loans.
Work shows a significant role of GDP per capita, real interest rates, broad money supply (M2), changes in the real effective exchange rate in the rise of impaired loans for the entire panel of countries, while other macroeconomic and bank-specific variables are significant for a subset of them. The effects of the variables on nonperforming loans have been assessed, either sequentially, taking each variable at the time, or jointly, allowing a simultaneous variation in the set of banking and macroeconomic variables over the prediction range.
In like manner, Hoggarth, Sorensen and Zicchino (2005) accounted the link between macro variables (the UK quarterly output gap, retail and house price, inflation, the nominal short-term interest rate and the real exchange rate) and loan aggregate write-offs (data are since the late 1980s) and sector losses (corporate and household losses, where data are since the early 1990s). Results were in favor of a significant negative and long lasting (up to six quarters forward) relation between the output gap and write off in the aggregate case. Nonetheless, maybe the different estimation period considered implied a more and much smaller sensitivity of the corporate and household sectors, respectively, to output gap with respect to the aggregate write off.
Furthermore, the studies of Filosa (2007), and Marcucci and Quagliariello (2008) for the Italian banking system adopted the VAR methodology. Marcucci and Quagliariello (2008) referred to a reduced VAR to prove the cyclical pattern of default rates, falling in good macroeconomic times and increasing during downturns: evidence resulted robust to different measures of the output gap and convincing for both the household and corporate sector.
Filosa (2007) estimated three different indicators of bank's soundness, adopting three alternative VAR models. In the first one, the proxy of banks' health was given by the ratio of the flow of new defaults to outstanding performing loans. In the second VAR, he used the ratio between the stock of nonperforming loans and outstanding performing loans. In the third model, the indicator of bank' soundness was represented by the interest margins-to-loans ratio. In contrast to the literature, his results demonstrated a weakly pro-cyclical behavior of NPLs to macroeconomic shock (output gap and inflation are chosen as two macro indicators) scenarios and to business cycle changing conditions in all models. Namely, he demonstrated the inverse correlation between default rate and external variables, consistent with the view of procyclicality but he found evidence of a significantly different behavior and intensity of the relationship across cycles in the first model, and no evidence of procyclicality when focused on the NPLs to loans ratio as a measure of distress. Moreover, the VAR supported the presence of significant feedbacks between banking conditions and the real economy: positive/negative shocks in the bad loans indicators reduce/rise real activity and inflation.
On the other hand, among recent evidences, Nkusu (2011) and Klein (2013) used both panel regression and panel vector autoregressive (PVAR) model as complementary approaches to evaluate, respectively, 26 advanced economies and 16 Central, Eastern and South Eastern Europe countries. Nkusu's panel VAR confirmed the feedback effect between NPL and macroeconomic determinants, suggesting that the

deterioration in the macroeconomic environment, like slower growth, higher unemployment or falling asset prices, is associated with debt service problems, reflected into rising NPLs, otherwise in the case of the favorable macroeconomic environment.
Results of Klein panel estimation suggested that NPLs are indeed affected by both macroeconomic and bank-level factors while the question of the responsiveness of real economy to NPLs provided by panel VAR analysis proved the existence of powerful macro-financial connections. An increase of one percentage point in real GDP leads to a decline slightly less than proportional in NPLs in the subsequent year, and an increase in impaired loans has a significant and negative effect on credit, real GDP growth, unemployment, and inflation in the periods ahead.
Latterly, Sandica and Dudian (2017) developed a VAR to obtain the evolution of companies' NPLs and describe the NPLs impulse response functions to shocks in the unemployment rate, output gap, and credit cycle. Results are in line with similar studies: in Romania, economic growth and NPLs are related negatively, while unemployment and credit cycle are linked positively with their evolution.

A further strand of literature is the one that uses, from a methodological perspective, a single equation time series approach. It allows checking just causes of vulnerability over time, contrary to panel data analysis that also assess the role of country or bank specific variables and capture the state-specific effects and the unobservable differences between countries.
Marcellino (2007) validated the employment of linear regression in many applied works defining linear time series as "a good benchmark for theoretical models" whose superiority can be scarcely beaten if they are carefully detailed (page 29).
In a linear regression model non-performing loans at time t is simply a linear function of regressors:

$$y_t = \alpha + X_t'\beta + \varepsilon_i \tag{4}$$

where α is a constant term, ε_i are unobserved errors and X is a $n \times k$ matrix of regressors or explanatory variables (the cited macroeconomic and bank specific factors determinants of NPLs); β is a $k \times 1$ vector of all unknown regression coefficients.
As an example, the aforementioned Keeton and Morris (1987) used a simple regression to deal with the American financial system and to propose similar and other explanations for the causes behind diversity in bank's loan performance. Using NPLs net of charge-offs as the primary measure of loan losses and a sample of almost 2500 insured commercial banks over the period 1979-1985, authors ascribed loan vulnerabilities to local economic conditions along with the lacking sufficient performance of particular sectors like agriculture and energy. "Banks with heavier loan losses are located in areas with worse economic conditions… and losses have tended to be higher in areas dependent on agriculture and energy". However, additional non-macro factors emerged as an explanation of loan losses: bank's credit management, portfolio diversification, and bank's risk appetite.
Alike, Arpa, Giulini, Ittner, Pauer (2001) made the case that credit risk is affected by macroeconomic development using a simple regression model. In making the comment "banks' health reproduces the health of their borrowers, which in turn reflects the health of the economy as a whole" (pages 92), authors, in fact, insisted that the macroeconomic dimension of credit risk cannot be ignored and whichever disturbance in the economy influences the stability of the banking system. The observed variable, risk provision of Austrian banks for the 1990s, was regressed on the real GDP growth rates, the developments of real estate price and real interest rates. In particular, estimation results suggested bank's pro-cyclical behavior according to which banks rise risk provisions in times of declining real GDP growth rates, while banks reduce them at booms. The model also reported a rise in risk provisions along with a rise in inflation (expected as increasing inflation usually rises risks) and a decrease in risk with real estate prices (unexpected because normally a boost in real

estate prices should increase the value of mortgages and reduce the likelihood of loan losses and therefore provisioning).

In line with the expectations derived from the other studies, Shu (2002) focused upon the sensitivity of Hong Kong's bad loans to macroeconomic indicators over the period 1995-2002. His model established a link between bad loan ratio to total lending and economic growth, unemployment, inflation and interest rates. While unemployment resulted insignificant, with regard to GDP and inflation effects, the expected signs are confirmed negative; the positive sign of interest rate implies that effects of less loan demand and higher borrowing costs (that cause more defaults) have control over bank's propensity to lend more and expand credit. Using a similar single equation time series model, Kalirai and Scheicher (2002) also modeled the resilience of the Austrian banking system to the external scenario. To portray the state of the economy, real GDP growth rate, industrial production, consumer price inflation, money growth and stock market indices were used as explanatory variables. According to the estimates, GDP did not enter significantly into the regression but the industrial production index had the most predictive power along with money growth and financial market indicators.

In a 2009 study, Khemraj and Pasha (2009) analyzed the sensitivity of NPLs to macroeconomic and bank-specific factors in the Guyanese banking sector. Three macro factors (the growth in real GDP, the annual inflation, and the real effective exchange rate) have been used, both contemporaneously and with one-year lag in order to capture the effect of time on the variables selected, considering that adverse shocks may not affect immediately on the loan portfolios of banks. The dataset consisted of six commercial banks over the period 1994-2004 investigated by a simple linear regression. The correlation analysis showed a significant negative relationship between GDP growth and NPLs, meaning that betterment in the real economy and a change in real income are likely to imply an immediate discount of the commercial banks NPLs. The same negative association has been found between inflation and the ratio of NPLs to total loans, contrary to other studies, and a positive link between the ratio of NPLs and real effective exchange.

Following the methodological rut, more recently, in Italy, Bofondi and Ropele (2011) depicted the development of NPLs ratios for the two groups of households and firms over the period 1990-2010. The study has been conducted separately because the capability to repay loans may be conditioned by different macroeconomic determinants. Among the indicators tested appear, as usual, the GDP (the annual growth rate) and the unemployment rate as key factors of the state of the economy, the cost of borrowing and the burden of debt. Bofondi and Ropele (2011) also noted the dependence of loan quality on the business cycle; both firms and households assets improve as economic conditions become more buoyant. Particularly, regarding households, for instance, NPLs are related inversely with GDP and the house prices, while there is a positive association with unemployment and short-term nominal interest rates.

Saba, Kouser, and Azeem (2012) observed the asset quality of the United States banking sector, using a time series dataset from 1984 to 2010. He provided, once more, a strongest relationship between Real GDP per capita and NPL rate, suggesting that US banks should consider that factor while issuing loans.

Among recent works focused on Europe situation, Shingjergji and Shingjergji (2013) tried to understand the ongoing growth of the non-performing loans and the role of the macroeconomic situation in Albania. The magnitude of GDP has been tested by a simple regression model on data from 2002 to 2012 but, the finding betrays the international evidence about a negative association of NPLs with GDP growth rate so the increasing amount of bad loans is not a proof of the not easy economic conditions of businesses and customers.

Other researches on single country include that of Filip (2015) who studied correlations among macro factors and NPLs on Romanian and EU data processing an OLS regression. Not departing dramatically from the findings of similar studies for other countries, he revealed a significant negative correlation between the variation of NPLs ratio and the real GDP growth, unemployment rate and inflation rate for both Romania and EU econometric models. Recently, Adeola and Ikpesu (2017) offered a unique investigation of Nigeria but reversed such results showing that GDP has a positive relationship with NPLs but it doesn't influence or determine their amount in Nigeria. On the contrary, the remaining variables, inflation, exchange rate, lending rate, money supply and unemployment maintained the expected positive signs.

Summarizing, the empirical evidence asserts convincingly an association between favorable macroeconomic environment conditions and borrower capability in loan repayment.

It is no strange to indicate real GDP growth as the main driver of NPLs and the major proxy of the business cycle. Specifically, concerning its impact, the negative relationship between NPL and GDP growth rates is consistent with the literature. While few researchers found no relationship between real GDP and loan portfolio quality, mostly found that credit risk turns out to be higher when the business cycle reverses downwards, reflecting the increased riskiness of the credit portfolio and the danger of a credit crunch. GDP growth, instead, translates into higher income that improves the ability of borrowers to pay back the debts.

There is also some evidence about the fact that credit risk is built up in boom phases and materializes in downturns, as intermediaries do not make enough provision in good macroeconomic times. Consequently, when economic conditions turn, loan losses come out, provisions increase, profitability and credit supply decrease, amplifying the effects of the recession and delaying economy potential recovery.

Moreover, worthy to say, results mostly display similar symptoms of credit risk; they do not vary with respect the different methodologies and across countries, reflecting similarities in banking systems; some differences may depend on the nature of the economy of each country and on the coinciding effects of these variables.

However, what written must not convince that the gross domestic product is the only relevant macroeconomic variable in a study on problem loans. Better, economic activity is not able to clarify completely the development of non-performing loans across countries and over time.

Macroeconomic turbulences of almost any type can negatively damage bank balance sheets and borrowers' debt-servicing capacity, and if the shock is large enough it can threaten the solvency of large parts of the banking system. Aside from GDP, it seems plausible to include other macro variables such as unemployment, inflation, exchange and interest rates, market volatility, industrial production. They may provide additional information regarding the impact of macroeconomic conditions on household and firms. Exactly, for example, bad economic conditions raise unemployment that influences negatively the incomes of households and increases the debt burden because people lose their cash flow streams and they cannot repay debts. In the same way, unemployment affects negatively the production and the sales of firms due to the lower consumption of goods during high unemployment time.

Following the words of Nkusu (2011), the impact of inflation on the borrower ability to repay debts is not clear, it is a complex multi-sided process. On one hand, higher inflation can make debt easier to pay back, reducing the real value of the outstanding loan, but on the other hand, can also reduce the real income of borrowers when their salaries are fixed.

There is an empirical evidence of a positive relationship between unemployment and NPLs in Gambera (2000), Baboucek and Jancar (2005), Bofondi e Ropele (2011), Nkusu (2011), to name just a few. The association between inflation and impaired loans has been investigated by Fofack (2005), Khemarj and Pasha (2009), Nkusu (2011), to name others.

Continuing with the other factors, the relationship between the exchange rate and problem loans is confused and indeterminate.

A devaluation of this variable is expected to improve the debt servicing capabilities of export-oriented firms and lower the NPL ratio. It means a decrease in the local currency that make the goods and services produced in that country relatively less expensive. This strengthens the competitiveness of export-oriented firms and affects positively their ability to service their debt (Castro, 2012).

On the other hand, the relationship reverses if debtors are mainly borrowing in foreign currencies but have income in local currencies as the depreciation of domestic currency rises the repayment amount and causes loan losses.

The depreciation in the exchange rate it is expected to make worse the position of borrowers because they cannot afford to repay less than borrowed initially, therefore enlarging loan defaults and losses.
To conclude with the variable list, industrial production indicator is a picture of the changes in the output level of an economy that helps to determine the nation's strength and ability to produce goods domestically and it helps to define turning points in the business cycle. Such as index heads the GDP growth and it is expected to reduce loan losses. Similarly, the cyclical trend of the economy is led by developed and buoyant stock market; booming market indices reflect a positive outlook on firms' profitability and the nation's health, lowering the probability of loan default. Higher interest rates, finally, reproduce the cost of borrowing and are connected with higher loan losses since the greater the cost of borrowing the less the probability of firms and households to service their debt.

2.2.2. Bank specific factors

Along with macroeconomic factors, copious studies included internal bank-specific variables as signal and proximate cause of risky lending. This confirms the impact on the intermediary performance of the interaction between general economic conditions and banks' management decisions to regulate loan portfolio such as leverage, observation of capital adequacy ratio, targeting size, return on assets and so on.
Contributions belonging to this group include the preliminary works of Berger and DeYoung (1997) who focused on profit efficiency indicators and financial capital, and De Young and Hasan (1998) who took into consideration banks stage of life as structural characteristics.
The latter demonstrated that bad loans are related inversely to the age of the intermediaries, meaning that the *de novo* American banks are more likely than the *established* ones to underperform. This result was demonstrated to be valid in the US market for a period of not less than five years, while banks with more than five years have characteristics not dissimilar to the mature ones[18].
De Young (1999) later specified the idea according to which latest entries initially have almost non-performing loans as their loan portfolios seem to be composed by borrowers with a strong financial basis. Nevertheless, over time, just after three years, the borrower's quality inevitably declines and emerges the inability of *de novo* banks to contain risky loans. "Only after three years, *de novo* bank outperforms one-quarter of the established banks" (page 22).
Moving on the Berger and DeYoung (1997) Granger causality analysis, it suggested a negative bidirectional relation between cost efficiency[19] and problem loans by testing some fundamental interpretative hypotheses about the relation between loan quality and management behavior. Among the equations-explanations, the so-called *bad management*, the *bad luck*, *skimping* and *moral hazard.* The first one ascribed the deterioration of the asset quality to a lack of skills of the management and to poor bank management practices, like scarce client scoring and monitoring activity. "Low levels of cost efficiency Granger cause or predict increases in nonperforming loans, implying that cost-inefficient managers are also poor loan managers" (page 5). The second hypothesis clarified the causality from NPLs to cost efficiency assuming the problem loans dependency on macroeconomic shocks (environmental characteristics) rather than on management choices. The principle of causality is clearly reversed compared to the previous hypothesis: the shock occurs at first and worsens the quality of the loans; this problem affects the efficiency of the intermediaries entailing an increase in credit management costs for the bank; higher unit costs per product impinge efficiency performance.
Finally, the so-called moral hazard conjecture regarded the nexus capitalization-problem loans and Berger and DeYoung (1997) pointed out that capital ratio negatively Granger-causes nonperforming loans ratio: a drop in bank capital ratios generally precedes an increment in nonperforming loans in weakly capitalized banks

[18] "Although profit efficiency improves rapidly over the next two years, the typical de novo bank does not become as profit efficient as similar-size established banks until it is nine years old" (page 585).
[19] Cost efficiency are measure by the ratio between operating expense and operating income.

probably due to their incentives to engage risky lending behavior. Indeed, as insufficiently capitalized banks risk lower loss in terms of capital, managers are particularly risk loving to exploit it if profitability goes up.
The "bad management" hypothesis, according to which reduced cost efficiency fosters an increase in non-performing loans was later supported by Podpiera and Weill (2008), and the negative correlation among capital-adequacy and NPLs confirmed by Salas and Saurina (2002) and Klein (2013), while it did not find explanatory power in Louzis, Vouldis and Metaxas (2011).
Also, Salas and Saurina (2002) and Louzis, Vouldis, and Metaxas (2011) proposed bank size among bank variables and showed that banks size is related negatively to the rate of NPLs both because bigger size allows for more diversification opportunities and because large banks own more means and competencies for assessing loans. In this line of research, Hu, Li, and Chiu (2004)[20] and Ranjan and Dhal (2003). However, there is no consensus in the literature about this relation, as after all already seen for capital adequacy as well. No relationship was found in Greenidge and Grosvenor (2010) and Salas and Saurina (2002), while a positive suggestion was declared by Stern (2004), Khemraj and Pasha (2009) and Suryanto (2015). Stern (2004), for instance, exploited the "too big to fail" hypothesis and talked about the moral hazard of a bank manager to justify his claim. Large banks are encouraged towards leverage, excessively risky and less responsible behavior given that it is known that government protects and bails them out in case of failure. This awareness increases in this way the magnitude of non-performing loans and results in a positive relationship.
Earlier, we pointed out how the empirical evidence relates bank size to diversification opportunities. A further variable used in literature to indicate a portfolio consisting of other sources of income except for loan making was the non interest income, as done by Louzis, Vouldis, and Metaxas (2011) and Hu, Li, and Chiu (2004).
The impact of non-interest income on NPLs appear again to be mixed. It can be correlated negatively to NPLs as banks that can generate efficiently cash flow from other operations reduce the incentive to contract high-risk loans and to finance speculative projects (Ranjan and Dhal, 2003). Better, given that loan portfolio is typically the largest asset and the return on loans constitute the primary source of income, revenues from other operations allow banks to loosen the pressure to create revenues from loans (Hu, Li, and Chiu, 2004). Further, the nexus can be negative but insignificant (Louzis, Vouldis, and Metaxas, 2011) or positive as explained by Stiroh (2004)'s dark side hypothesis of diversification that ascribes it to the possibility to enter unexperienced activities.
Regarding other variables, the bank's profitability measured by return on equity (ROE)[21] relates to NPLs through different sides. The expected sign can be positive as suggested by Chaibi and Ftiti (2015) quoting Rajan (1994) due to the fact that banks renounce of long-term profitability to improve current earnings chasing risky business and lending opportunities. Second, the relation could be even negative as suggested by the bad management hypothesis according to which bank performance mirrors the quality of management as proven again, in a different context, by Louzis, Vouldis and Metaxas (2011), Klein (2013) and Chaibi and Ftiti (2015). Similarly, Fofack (2005) found such negative evidence considering both net interest margin and return on assets as alternative measures of profitability in the Sub –Saharan Africa.
Finally, the role of management related issues of credit quality has been proved by Shehzad, de Haan, Scholtens (2010). They tested the position that ownership concentration improves banking firm performance against the view that ownership concentration does not matter for banks' riskiness for around 500 commercial banks from more than 50 countries averaged over 2005–2007. Focusing on three different definitions of concentrated ownership based on the percentage of shares held by stakeholders (10, 25 and 50%), authors found ownership concentration reduces non-performing loans at least if ownership is above 50% of the shares. Results suggest that when two or three shareholders hold the ownership, the quality of the portfolio of the bank may deteriorate for the reasons related to bargaining problems, conflicts of interest among the controlling and

[20] (Hu, Li and Chiu, 2004)'s theoretical model is also particularly interesting for analyzing the relation among NPLs and ownership structure of commercial banks in Taiwan, revealing that the rate of NPLs decreased as government ownership in a bank increased.
[21] It is defined by the ratio of net income to total equity.

the minority. In contrast, when there is the control of just one owner, the monitoring of the bank's management is more efficient, leading to a lower impaired loans ratio.
Proceeding, Fatimoh (2012) emphasized the role of a good governance as a means to achive financial performance. The higher the nonperforming loan ratio, the lower the corporate governance quality that is a rise in non-performing loans depicts poor quality of governance, which does not provide meaningful and reliable financial report on business practices.
Likewise, Leventis, Dimitropoulos, and Owusu-Ansah (2013) investigated the effect of corporate governance structure on US listed commercial banks credit quality. They recognized a significant association claiming that banks having effective board and audit governance structures tend to identify losses (specifically loan loss provisions are accrued expenses reflecting management's current period assessment of the level of future loan losses of a bank's loan portfolio) more timely than gains.

3. Gaps in empirical contributions

The preceding account, confirming that both bank-level and macroeconomic factors play a role in affecting the banks' asset quality, shows as a wealth of studies adopted increasingly time series or panel models to describe the development of NPLs ratio over the business cycle. The methodological approach differs with respect to the type of data and the kind of study, depending on whether it was a cross-country or an individual country's analysis, and there was the willingness to take into account the dynamism with which factors influence NPAs.
However, as well as all macroeconomic and financial time series exhibit decisive non-linear patterns in their behavior due to large-scale events (transformation in government policy, wars, financial panics, periods of low/high stock market valuations, low/high interest rates or natural disasters) that alter their dynamics, in the same way, the relation between the real and financial sector may be non-linear.
The majority of the abovementioned researches based on the hypothesis that time series are stationary linear processes, failed to regard this potential feature in the estimated relationships so even substantial macro shocks, reason of structural breaks or simply regime shifts, would have a very limited impact and results could be misleading[22].
Figure 1-2 is revealing in this sense, illustrating the USA systemic non-performing ratio and GDP growth rate over the period 1987-2016.
The shaded areas show the recessions in American economy according to the National Bureau of Economic Research (NBER) and data confirm a different behavior of NPLs in different states of nature.
NPLs pick up during the three distinct periods coincided with episodes of economic recession while decaying subsequently in an economic growth phase.
The feature that a variable behaves differently during economic cycle (in our case the different association between NPLs and GDP in expansion and contraction) simultaneously with the different magnitude of GDP impact in each cycle constitute the basis of an important set of non-linear studies.
Several nonlinear models have implemented in the literature to overcome the insufficiencies of linear models, and to picture the observed asymmetric behavior of economic variables. All of them consider the presence of different regimes in the dynamic of the time series, meaning inference on the regime in operation at any point in time and forecast of the possible future switches.

[22] Foglia (2009) stressed the greater aptness and performance of non-linear approaches.

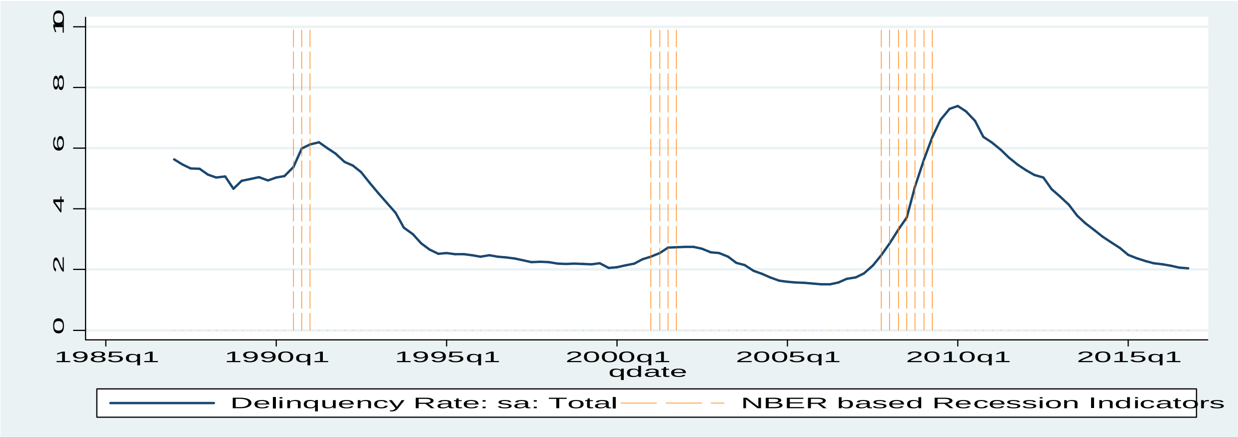


Figure 1-2 USA Delinquency rates amount. Source: Federal Reserve Board data.

3.1. Brief Econometric survey

It was the contribution of Hamilton (1989), Tong (1983), Teräsvirta and Anderson (1992) and others that introduced the possibility to capture parameter variation by regime-switching models, making workable exchange rate, unemployment, volatility and gross domestic product analysis, for instance.
Generally speaking, regime-switching models can be classified into three groups: threshold regime-switching, smooth transition, and Markov regime-switching models.
By now, the properties of these processes are well known.
In particular, suppose that at a certain point in time t_0 there is a significant change in the average level of a series caused by some imperfectly predictable forces[23] that produced changes in the intercept from $a_{0,1}$ to $a_{0,2}$ and involved two different AR specifications for the variable y_t (the first AR applies for $t = 1, \dots, t_{0-1}$ and the other fits the data for $t = t_0, t_{0+1}, \dots T$):

$$y_t = \begin{cases} a_{0,1} + a_{1,1}y_{t-1} + \cdots + a_{p,1}y_{t-p} + \epsilon_t \\ a_{0,2} + a_{1,2}y_{t-1} + \cdots + a_{p,2}y_{t-p} + \epsilon_t \end{cases} \tag{5}$$

a regime switching model is given by the larger equation

$$y_t = a_{0,s_t} + a_{1,s_t}y_{t-1} + \cdots + a_{p,s_t}y_{t-p} + \epsilon_t \tag{6}$$

that encompass the above formulas rather than state that two different autoregressive processes govern data up to t_{0-1} and after that date; s_t is a random variable that may assume the value 1 until $t = t_{0-1}$ and $s_t = 2$ from $t = t_0$ until the end of the sample, T.
The assumptions about the regime-switching variable differentiate Threshold and Markov-switching models. Indeed, a complete picture of the probability law leading the observed data necessitates a probabilistic model of what caused the change from $s_t = 1$ to $s_t = 2$ and how the evolution of the state process is modelled.
The econometric literature proposes precisely the three formulations of threshold models, smooth transition models, and Markov Switching models (hereafter MS).
In a threshold autoregression (TAR) model, the evolution of the state process is supposed triggered by the observed behavior of the level of an economic variable in relation to an unobserved threshold value r[24]. In Markov switching models the behavior of a time series is posit to be disrupted occasionally by shocks that produce different transition states whose regime switch depends on an unobserved series or latent variable

[23] According to Hamilton (1994), the regime is not perfectly predictble, but rather is a random variable.
[24] The number of thresholds is always equal the number of regimes minus one.

such that the prevailing state at any given time can never be determined precisely and it can be only assigned probabilities to different schemes. Smooth transition autoregressive (STAR) models related closely to TAR and simply suppose that a variable that follows a stationary and ergodic process switches smoothly from one state to another, instead of moving abruptly as assumed in threshold autoregressive.

3.1.1. Markov Switching Model

The roots of regime-switching models date back to the earliest applications of Goldfeld and Quandt (1973) who introduced the first Markov-switching model regression, by assuming that the probability of being in the current regime depends on its immediate preceding state through a Markov chain process and introduced the transition probability matrix to capture the state-dependency.

Later, the contribution of Hamilton (1989) popularized regime-switching analysis in econometrics presenting an autoregressive approach and a nonlinear iterative filter.

The general idea of Hamilton algorithm is that the parameter of a time series model depends on a discrete unobservable regime variable $s_t = 1, \dots . N$ that represents the regime in operation at time t.

It means that supposing y_t the variable of interest, for example, the GDP following the literature or credit quality in the context of our thesis, then a two-state Markov-switching model assumes that the conditional expectations of the series of interest are different from recession versus expansion periods (conditional mean model):

$E(y_t) = \mu_1$ if the economy is in expansion and $E(y_t) = \mu_2$ if the economy is in recession, with $\mu_1 > \mu_2$ as it has been stated that the average growth rate tends to be higher in expansions than in recessions.

These two expected values can be formulated in a shorter form, using the following equation form:

$$y_t = \mu_{s_t} + u_t \tag{7}$$

where s_t is the stochastic (uncertain) discrete unobservable latent variable taking on values 1 when the economy is in expansion and two when the economy is in contraction, for all t.

Clearly, these two different expected values are not sufficient to explain the output dynamics; they are not the only drivers of the dynamic behavior of the series. As economic time series present powerful dependence between past observations and there is autocorrelation in the pattern of the series that is hidden in the error u_t, it can be captured by allowing error to follow a general autoregressive process of order p.

Therefore,

$$u_t = \sum_{i=1}^{p} \phi_i u_{t-i} + \varepsilon_t \tag{8}$$

with $\varepsilon_t \sim N(0, \sigma^2)$ a standard independent and identically distributed Gaussian white noise with zero mean and variance σ^2 that may be regime dependent too or constrained to be the same in both regimes as assumed by (Hamilton, 1989) himself.

The replacement of (8) in (7) gives

$$y_t = \mu_{s_t} + \sum_{i=1}^{p} \phi_i u_{t-i} + \varepsilon_t \tag{9}$$

and substituting u_{t-1} by its value defined in (7), we obtain the final real output growth model where the estimated value of the series depends on an unobserved variable s_t, representing the state of the business cycle; or better, it depends not only on the current regime but also on the previous ones. Thus, for each date t, the process depends simultaneously on the mean at time t and at date $t-1$.

$$y_t = \mu_{s_t} + \sum_{i=1}^{p} \phi_i\left(y_{t-i} - \mu_{s_{t-i}}\right) + \epsilon_t$$

(10)

Without considering past regimes, the MS process of Hamilton in the case of an AR (p) process is given by the following equation:

$$y_t = a_{0,s_t} + a_{1,s_t}y_{t-1} + \cdots + a_{p,s_t}y_{t-p} + \epsilon_t \quad \text{with } s_t = 1{,}2$$

(11)

It obviously helps to understand why the model is a simple example of what is known as a piecewise linear model that is, although the model is globally non-linear, each component is a linear model.

Moreover, the whole specification of the Markov Switching needs an assumption about the law of motion for the unobservable variable to estimate the model. Better, if it is known the probabilistic description of how the economy changes from one regime to another, we will have a complete picture of the dynamics of y_t.

In the opinion of Hamilton (1989), s_t is the realization of a discrete[25], first order[26], irreducible and ergodic[27] Markov chain (it evolves as a Markov chain)[28]. This assumption implies that the transition probabilities exhibit the so-called Markov property:

$$p_{ij} = P(X_{t+1} = j|X_t = i) =$$
$$= P(X_{t+1} = j|X_t = i) = P\left(X_{t+1} = j|X_t = i, X_{t-1} = i_{t-1}, X_{t-2} = i_{t-2}, \dots X_0 = i_0\right)$$

(12)

where the probabilities (p_{ij}) with i,j=1,2 are called transition probabilities of moving from one state to another.

The Markov property is a certain of conditional probability statement saying that the future states of the process depend only upon the present state, not on the sequence of events that preceded it.

It simply specifies that the transition probability is not affected by the past of the process: if X_t is the current state and we want to predict X_{t+1}, the probability of the state at time $t + 1$, given the entire past history of the system, doesn't depend on the past.

Saying in other words, only the most recent information matters and right the current state is enough to determine all the distribution of the future; the prediction of the future cannot be improved by adding any knowledge of the past. All information about the particular path followed to reach a state and how we get there can be ignored[29]; the past has no bearing in making prediction about the future and what we have to take into account is just where we are right now: "the probability of a change in regime depends on the past only through the value of the most recent regime" Hamilton (2014).

Naturally, it doesn't mean that the past lacks information about the future behavior as well as it doesn't imply that the past is independent of the future. We are just saying that past and future are conditionally independent given the present, meaning that the past does affect the future through the present value.

---

25 The state variable can assume only finite number of regimes.

26 As specified below, Markov chain is a first order because the only immediate preceding period can influence the current state.

27 An ergodic Markov chain is a covariance stationary process. Better, following (Hamilton , 1994) it is an irreducible Markov chain with one of the eigenvalue of the transition matrix P equal to 1 and the other eigenvalues inside the unit circle. A Markov chain is called irreducible if there is only a single communicating class in the state space. This communicating class exists if every state j is accessible from every state i within finite time. Saying that simply, chain is irreducible if it possible (with positive probability) to go from anywhere to anywhere.

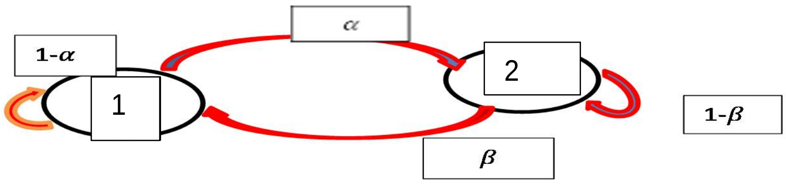
28 A Markov chain is an example of stochastic process (it is a probability model used to describe phenomena evolving over time or space) that satisfies the Markov property.

29 Markov process is not path dependent. Actually, models such as the autoregressive moving average (ARMA) and the generalized autoregressive conditional heteroskedastic (GARCH) models exhibit path dependence because the dependent variable at date t (the error term in ARMA models, the conditional variance in GARCH) depends on the entire path of states that have been followed until that date.

The transition probabilities can be written as a matrix, the transition matrix P whose rows sum to 1 as, at any given time, the variable must be in one of the states. Columns may or may not sum to 1.
So, $p_{i1} + p_{i2} = 1$.
Consider credit risk in the context of our thesis. Non perfroming loans or delinquency rates at any time can be higher (state 1) or lower (state 2).
Supposing that over a quarter period:

1. Delinquency rates change from lower to higher with probability $\beta \epsilon (0,1)$
2. Delinquency rates change from higher to lower with probability $\alpha \epsilon (0,1)$

In terms of a Markov model, we have $S = (1,2)$. When $s_t = 1$ the economy is in downturn leading to an observed growth of NPLs; when $s_t = 2$, the economy is in the upturn.
We can represent the Markovian dynamics as in the following graph:



and write the transition probability matrix form as $P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$.

(13)

More generally, the transition matrix for a two-state Markov chain takes this form

$$P = \begin{bmatrix} P(s_t = 1|s_{t-1} = 1) & P(s_t = 2|s_{t-1} = 1) \\ P(s_t = 1|s_{t-1} = 2) & P(s_t = 2|s_{t-1} = 2) \end{bmatrix} = P \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = P \begin{bmatrix} p_{11} & 1-p_{11} \\ 1-p_{22} & p_{22} \end{bmatrix} \tag{14}$$

where p_{11} and p_{22} denote the probability of being in regime one, given that the system was in regime one during the previous period, and the probability of being in regime two, given that the system was in regime two during the previous period, respectively. Exactly these two parameters are indicative of the persistence of recession and expansion, referring to the probability of an expansion after having been in expansion in the previous period and to the probability that we are in a recession after having been in recession in the previous period. If the probability $p_{11} = 1$ ($p_{22} = 2$), the state 1 (2) is said to be an absorbing state (once the process enters to regime 1 there is no possibility of returning to state 2) and the Markov chain is reducibile.
Furthermore, p_{12} defines the probability that the series will change from state 1 in period t -1 to state 2 in period t, and p_{21} defines the probability of a shift from state 2 to state 1 between times t -1 and t. The equivalence $p_{12} = 1 - p_{11}$, instead, comes from the fact that starting from regime one, we can only switch to either regime two or reamin in one.
Every probability p_{ij} is one of the model parameters that should be estimated; does not require any prior information and it is deduced from the data. Other parameters to estimate include the autoregressive coefficient ϕ_i, the constant in the two regimes μ_{s_t} and the error variance σ^2.
Hamilton's Markov-switching auto-regression was widely applied in the following literature and remains one of the most popular and valuable nonlinear time series tools to analyze economic and financial time series. However, his aforementioned paper spawned several further research in a number of directions, and the estimation techniques adjusted to include these developments.
For example, Diebold, Lee, and Weinbach (1994) developed a first extension since he pointed out that "although the popularity of Hamilton's model is well deserved, it nevertheless incorporates a potentially severely binding constraint, the constancy of state transition probabilities".

Authors, differently, noted an explicit lack of flexibility of transition probabilities assumption in explaining changes in regimes since it does not include the driving variables behind the switching process as well as the possibility that the transition probabilities are driven in part by the past evolution of the process. The consequence is that transition probabilities are time-invariant over the cycle (fixed transition probabilities[30]). It means that even if the expected duration of expansions and recessions can differ with the underlying strength of the economy, they are forced to be constant over time and the probability that a recession at time $t-1$ is followed by a recession at time t remains the same and it is independent of the duration of the contraction state. Different recent works have extended the Hamilton model to allow for time-varying (TVTP).
These models preserve the Markov structure for the unobserved states but incorporate additional covariates. Some of them suppose the transition probabilities to change according to information variables such as the duration dependence[31] on the number of periods the system has been in a particular state (how long the economy remained within the same regime or is duration dependent)[32], other refer to strictly exogenous variables like observable economic fundamentals[33]. In both cases, the implication is a not constant expected duration of each regime and the transition probabilities are expressed as a function of z_t, the indicator variable[34] which governs the evolution of the unobserved regimes (Filardo, 1994).

$$P(S_t = s_t|S_{t-1} = s_{t-1}, z_t) = P \begin{bmatrix} p_{00}(z_t) & p_{01}(z_t) \\ p_{10}(z_t) & p_{11}(z_t) \end{bmatrix} = \begin{bmatrix} p_{00}(z_t) & 1-p_{00}(z_t) \\ 1-p_{11}(z_t) & p_{11}(z_t) \end{bmatrix} \tag{15}$$

Filardo himself popularized, moreover, the logistic parametrization of the transition probabilities so that they are always between zero and one. For the two-state case, the resulting functional form is

$$\begin{bmatrix} \frac{\exp(\theta_0 + \theta_1 z_{t-1})}{1+\exp(\theta_0 + \theta_1 x_{t-1})} & 1-\frac{\exp(\theta_0 + \theta_1 z_{t-1})}{1+\exp(\theta_0 + \theta_1 x_{t-1})} \\ 1-\frac{\exp(\theta_2 + \theta_3 z_{t-1})}{1+\exp(\theta_2 + \theta_3 x_{t-1})} & \frac{\exp(\theta_2 + \theta_3 z_{t-1})}{1+\exp(\theta_2 + \theta_3 x_{t-1})} \end{bmatrix} \tag{16}$$

Likely, in Diebold, Lee, and Weinbach (1994) the transition probabilities evolve as logistic functions of $\beta' x_{t-1}$, where the vector x_{t-1} contains economic variables that have an effect on the state transition probabilities and β are the parameter of the transition matrix, then the probability of a change in regime is modeled as:

$$\begin{bmatrix} P(s_t = 0|s_{t-1} = 0, x_{t-1}\beta_0) & P(s_t = 1|s_{t-1} = 0, x_{t-1}\beta_0) \\ P(s_t = 0|s_{t-1} = 1, x_{t-1}\beta_1) & P(s_t = 1|s_{t-1} = 1, x_{t-1}\beta_1) \end{bmatrix} = P \begin{bmatrix} p_{00} & p_{01} = 1-p_{00} \\ p_{10} = 1-p_{11} & p_{11} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{\exp(x_{t-1}'\beta_0)}{1+\exp(x_{t-1}'\beta_0)} & 1-\frac{\exp(x_{t-1}'\beta_0)}{1+\exp(x_{t-1}'\beta_0)} \\ 1-\frac{\exp(x_{t-1}'\beta_1)}{1+\exp(x_{t-1}'\beta_1)} & \frac{\exp(x_{t-1}'\beta_1)}{1+\exp(x_{t-1}'\beta_1)} \end{bmatrix} \tag{17}$$

Filardo's methodology is well suited for NPLs investigation and its sensitiveness to macroeconomic shocks.

---

[30] For this reason, Markov chain are also defined as time homogeneous, meaning that the transition matrix is constant over time and does not change for all pairs of i and j.

[31] Durland and McCurdy (1994).

[32] The evidence is that the longer recession is, the more likely it is the probability of remaining in recession to end soon but the same it is not true for expansions.

[33] Diebold, Lee and Weinbach (1994).

[34] In Durland and McCurdy (1994), as insatance, probabilities are parameterized as functions of the inferred current state and the associated number of periods the system has been in the current state $P(s_t = i|s_{t-1} = i, D_{t-1} = d) = P_{ii}$ and $P(s_t = j|s_{t-1} = i, D_{t-1} = d) = P_{ij} = (1-P_{ii})$.

Specifically, for our purposes we may consider a linear regression with regime switching coefficients:

$$y_t = \beta_0(s_t) + \sum_{i=1}^{n} \beta_i x_{it}(s_t) + \varepsilon_t \tag{18}$$

$$NPLs_t = \beta_0(s_t) + \beta_1 GDP_t(s_t) + \beta_2 UNP_t(s_t) + \beta_3 CPI_t(s_t) + \cdots \beta_n NPLs_{t-1}(s_t) + \varepsilon_t \tag{19}$$

where s_t is the state or regime index and x_{it-1} the exogenous explanatory variables measured at time t, which may include lagged values of y.

The law of motion of s_t still assumes that the probability of s_t depends on s_{t-1} plus a vector of exogenous variables z_t at time $t-1$ (as our objective will be forecast) that may include one or more elements of X'_t with a logistic specification. Formally:

$$\mathrm{P}(\mathrm{s_t} = 1|\mathrm{s_{t-1}} = 1, \mathrm{z_t}) = p(z_t) \tag{20}$$

$$\mathrm{P}(\mathrm{s_t} = 2|\mathrm{s_{t-1}} = 1, \mathrm{z_t}) = 1 - p(z_t) \tag{21}$$

$$\mathrm{P}(\mathrm{s_t} = 2|\mathrm{s_{t-1}} = 2, \mathrm{z_t}) = q(z_t) \tag{22}$$

$$\mathrm{P}(\mathrm{s_t} = 1|\mathrm{s_{t-1}} = 2, \mathrm{z_t}) = 1 - q(z_t) \tag{23}$$

Considering just the effect of GDP growth rate, we will have:

$$p_{11}(GDP_{t-1}) = \frac{\exp(\theta_0 + \theta_1 \mathrm{GDP_{t-1}})}{1 + \exp(\theta_0 + \theta_1 \mathrm{GDP_{t-1}})} \tag{24}$$

The extensions of the basic framework discussed above are primary a modification of the state transition probability functions. Others allow for multiple regimes, the incorporation of the switching mechanism into conditional variance models (ARCH and GARCH of (Engle, 1982) and Bollerslev (1986)) or consider the kind multivariate form (MS-VAR), and panel of univariate and multivariate MS.

Despite this flexibility, regime-switching models generally follow the same structure.

MS-VAR, for instance, is a generalization of the univariate autoregressive one deeply discussed in Krolzig (1997) who defined them as simply VAR models with switching parameters depending on an unobservable regime variable s_t, that hence became a function of the state. The model can be written as

$$y_t = \mu(s_t) + \Phi(L) y_{t-1} + \epsilon_t \text{ with } \epsilon_t \sim N(0, \sigma_{st}^2) \tag{25}$$

and y_t a $k \times 1$ vector of random variables, μ the vector of constant terms, $\Phi(L)$ the polynomial lag operator of order p and ϵ_t the stochastic error term. Literature suggests that parameters can be combined such that only some of them are conditional to the regime switch, resulting in a variety of specifications. Krolzig (2003) established a unique notation: 1) Markov-switching mean MS(M), 2) Markov-switching intercept MS(I), 3) Markov-switching autoregessive parameters MS(A), 4) Markov switching heteroscedasticity MS(H).

The same logic applies to panel VAR that is a standard VAR plus a cross-country dimension (Billio, Casarin, Ravazzolo, Van Dijk, 2016).

Switches in the conditional variance models mean time-varying variance with an ARCH process whose parameters themselves are subject to change. Indeed, as stated above volatility is not mean reverting and persistent for a long time; it is also affected by sudden jumps and structural breaks in the economy. Bauwens, Preminger, and Rombouts (2010) claimed that estimating GARCH model on data displaying a structural break yields misspecifications, bias estimates, integrated GARCH, poor predictions and only models whose parameters vary within different regimes are appropriate for modeling volatility.

Cai (1994), Hamilton and Susmel (1994) were the first to combine the two approaches into a Switching-ARCH model (SWARCH), that is they apply Hamilton's idea of regime switching parameter into an ARCH feature, defining the conditional variance equation as:

$$h_t = \sum_{i=1}^{k} \left(\alpha_0^i + \sum_{j=1}^{q} \alpha_j^i \epsilon_{t-j}^2 \right) \tag{26}$$

The generalization into a GARCH needs when shocks have a persistent effect on volatility and it is quite difficult:

$$h_t = \alpha_0^i + \alpha_1^i \epsilon_{t-1}^2 + \beta_1^i h_{t-1} \tag{27}$$

Problem arises as h_t is path dependent or better, as the conditional variance h_t depends on h_{t-1}, then h_t is a function of the entire history of the state variable depending not only on s_t but also on s_{t-1}: h_{t-1} dependence on h_{t-2} implies h_t is affected by the value of s_{t-2} , and so on.
Different estimation techniques were proposed to avoid such state dependence. Gray (1996) model was among the first to circumvent the problem and generalizing a switching regression GARCH. To have no longer path dependence, he suggested computing h_t as weighted sums of $h_{i,t}$ with the weights being the prediction probabilities$P(s_t = i|\Phi_{t-1})$.
That is

$$h_{it} = c_i + \sum_{j=1}^{q} a_{ij} z_{t-j}^2 + \sum_{j=1}^{p} b_{ij} h_{t-j}, \qquad i = 0,1 \tag{28}$$

$$h_t = h_{o,t} P(s_t = 0|\Phi_{t-1}) + h_{1,t} P(s_t = 1|\Phi_{t-1}) \tag{29}$$

where just $h_{0,t-j}$ and $h_{1,t-j}$ have been used to compute h_{t-j} without considering all possible values of the state variables.

3.1.2. Threshold autoregressive models and STAR

We said that when s_t is a random variable, the model is the random switching model of Hamilton. If s_t is assumed as an indicator variable taking the value of 1 or zero depending on whether is greater than a cut off value (threshold), the model is called a threshold autoregressive model.
Developed by Howell Tong (1978), subsequently discussed in details by Tong and Lim (1980) (see Tong, 1983), the TAR is a regime-switching model in which the evolution of the state process is determined by the observed behavior of the level of an economic variable in relation to an unobserved threshold. Better, it is a switching model where the state variable is assumed to be a deterministic function of an observed variable.
TAR approach catches the nonlinear path of time series by splitting the series into two regimes and then modeling them through a linear autoregressive model.
As a simple example:

$$y_t = \begin{cases} \mu_1 + \phi_1 y_{t-1} + u_{1t} \; if \; s_{t-k} < r \\ \mu_2 + \phi_2 y_{t-1} + u_{2t} \; if \; s_{t-k} \geq r \end{cases} \tag{30}$$

s_{t-k} is the state variable and it can be any variable such that y_t shifts from one state to another. It is equal to the current value of the state variable if $k = 0$, it is equal to the immediately preceding value of s if $k = 1$.

This means that the entire series is nonlinear but to ensure its stationarity it is divided into two linear parts that depend on the lag k and threshold r. In the applications where this variable is taken to be a particular lagged value of the process itself, an endogenous variable, regime shifts are said to be "self-exciting TAR" or SETAR: $s_{t-k} = y_{t-k}$.

$$y_t = \begin{cases} \mu_1 + \phi_1 y_{t-1} + u_{1t} \ if \ y_{t-k} < r \\ \mu_2 + \phi_2 y_{t-1} + u_{2t} \ if \ y_{t-k} \geq r \end{cases} \tag{31}$$

This two regimes first order SETAR model can be alternatively written as:

$$y_t = (\mu_1 + \phi_1 y_{t-1})I + (\mu_2 + \phi_2 y_{t-1})(1 - I) + u_t \tag{32}$$

where I is the indicator function or Heaviside indicator function defined as:

$$I = \begin{cases} 1 \ if y_{t-k} < r \\ 0 \ if \ y_{t-k} \geq r \end{cases} \tag{33}$$

Clearly, in one state y_{t-k} deceeds the threshold variable so $I = 1$ and $1 - I = 0$ and y_t follows this autoregressive procees $\mu_1 + \phi_1 y_{t-1} + u_{1t}$. In the other state, $I = 0$, $1 - I = 1$ and y_t follows the second autoregressive process.

The two regime models identified above can be extended to consider more than one lag of the dependent variable used in each regime, or a different number of lags in both regimes or multiple regimes, increasing the number of states to be more than two and determined by a single variable, by multiple variables, or by linear combinations of variables (cointegration case).

A general threshold model that notationally compute more than two regimes and more than one lag can be written by defining m thresholds $r_{o,} r_1 \dots r_m$ such that $-\infty = r_0 < r_1 < \cdots r_m = \infty$ as:

$$y_t = \sum_{j=1}^{m} I_t^{(J)} \left(\phi_o^{(j)} + \sum_{i=1}^{p_j} \phi_i^{(j)} y_{t-i} + u_t^{(j)} \right) =$$

$$= \phi_{o,j} + \phi_{1,j} y_{t-1} + \phi_{2,j} y_{t-2} + \cdots + \phi_{p,j} y_{t-p} + u_t$$

$$\text{if } r_{j-1} \leq y_{t-d} < r_j \tag{34}$$

where $I_t^{(J)}$ is an indicator function taking the value one if the variable is in state J and zero otherwise; $j = 1, \dots m$ and y_{t-i} is the threshold variable with delay parameter d that can also be an exogenous variable z_{t-d}.

When the regimes are driven by own lags of the underlying varibales, again, we talk about self-excting TAR.

In other words, TAR model can be thought as an extension of autoregressive models, allowing for changes in the model parameters according to the values of weakly exogenous threshold variables z_t, assumed to be past values of y.

In the context of the thesis, we could adopt threshold models to investigate the asymmetries in the relationship between credit risk and the macro fundamentals.

To allow for a threshold variable, we may develop a threshold estimation in the regression framework, as developed by Hansen (2000).

Given the starting hypothesis that NPLs are affected by the macro environment and such impact is subject to one or more regime-switches, a general model for this representation can be written as:

$$\begin{cases} y_t = \beta_1 X_t + \varepsilon_t & if\ q_t \le \gamma \\ y_t = \beta_2 X_t + \varepsilon_t & if\ q_t > \gamma \end{cases} \tag{35}$$

where y_t represents the NPLs ratio, X_t the explanatory vector variable set including, for instance GDP, unemployment rate, inflation, interst rate or any other useful regressor; ε_t is the error term, q_t the threshold variable that splits the sample into two regimes depending on the general macroeconomic conditions; γ is the threshold parameter to estimate endogenously, β_1 and β_2 are the different regression slopes that distinguish regimes.

The observable threshold variable q_t can be any independent variable we would like to consider as indicator of macroeconomic conditions. We could use the GDP growth rate to confirm the cyclical hypothesis of credit risk, supporting that the dependence of credit quality on the business cycle is significantly negative only during recessions. Particularly, we could assess whether exist at least one threshold value such that the dependence is nonlinear, we could infer the estimated threshold levels above and below which GDP influences NPLs and establish in which phase of the business cycle such affect is more pronounced and statistically significant. Alternatively, we could also use interest rates and inflation as threshold to find what is the highest cost of borrowing and the increment in the general level of prices that generate higher loan losses.

Finally, a more complex regime model may allow regimes switches to be determined by different threshold variables, GDP, and unemployment, or, once more, market volatility, for instance.

In the TAR model, there are discrete transitions between one regime and another, meaning that they take place in a single period of time. In other cases and for other series this assumption of the sharp threshold is not rational and the shift process may be smooth. The class of threshold autoregressive models known as smooth transition autoregressions (STAR) is an extension of TAR that allows for gradual regime switching.

The STAR was originally introduced by Chan and Tong (1986) and subsequently developed by Teräsvirta and Anderson (1992) by replacing the discrete indicator function in the TAR model with continuous smooth function $\theta = F(z_{t-d}, \gamma, r)$ bounded between 0 and 1 to define and guarantee gradual regime switch:

$$y_t = (\mu_1 + \phi_1 y_{t-1})\theta + (\mu_2 + \phi_2 y_{t-2})(1 - \theta) + u_t \tag{36}$$

A common choice of the continuous transition function is the logistic function:

$$\theta = F(y_{t-d}, \gamma, r) = \frac{1}{1 + \exp[-\gamma(y_{t-d} - r)]} \tag{37}$$

It leads to the so-called Logistic Smooth Transition Autoregressive (LSTAR) model.

The threshold variable y_{t-d} could also be an external threshold variable z_{t-d}; the parameter γ instead determines the smoothness of the change in the value of the logistic function or saying in other words, it is a measure of the speed of the regime switch and thus the transition from one regime to the other. A small value of γ denotes that the transition happens gradually, large value signifies that it occurs rapidly so for la arge value of γ LSTAR model approximates the TAR.

Presuming that $\gamma > 0$, for $y_{t-d} - r \to -\infty$, $\theta \to 1$, while for $y_{t-d} - r \to \infty$, $\theta \to 0$. Therefore, the transition function change between 0 and 1 according as the value of y_{t-d} changes realtive to a threshold.

An additional alternative approach to the LSTAR model is the exponential STAR (ESTAR). Specifically, in this case, is employed an exponential smooth function to model the switch:

$$\theta = F(y_{t-d}, \gamma, r) = 1 - \exp[-\gamma(y_{t-d} - r)^2] \tag{38}$$

3.2. Econometric survey and gaps in empirical contributions continuation

Perhaps business cycle is the best example of switching application in economics since periods of economic expansion and recession display asymmetrical behavior and economic variables change dramatically in business cycle phases. According to Kuan (2002), as for instance, linear models are not appropriate to catch GDP growth rates as they generally float around higher levels and they are more persistent during expansion but, they display lower and less persistent levels during contractions. Similarly, Bruno and Otranto (2008) cited Neftci (1982) described periods of high growth as longer and smoother than periods of declining, meaning that it is more likely that an expansionary quarter is followed by another quarter of positive growth while recession, on the contrary, is more short-lived.

Not surprising, therefore, considerable studies belong to business cycle researches that have adopted a non-linear approach to model, identify phases and classify each observation into one of two regimes[35]. Works are based on the common intuition that there is a "relation between the concepts of changes in cyclical phase and change in regime" and transitions matrix can help to describe how phases are related (Anas, Billio, Ferrara, Lo Duca, 2007).

Indeed, (Hamilton, 1989) first applied a Markov Switching (MS) approach to postwar US real GNP data spanning from the second quarter of 1952 to the last quarter of 1984. He fitted an MS(2)-AR(4) process with switching mean in order to demonstrate that the smoothed probabilities of his model were well able to track the corresponding conventional dates of recession provided by the NBER.

Along his lines, Huang, Kuan, Lin (1998) and Huang (1999) applied univariate Markov switching model to identify turning points in Taiwan GDP; (Bodman and Crosby, 2000) demonstrated the capability of two and three states first order Markov-switching models to fit the GDP data better than linear models in the Canadian economy. Chauvet and Piger (2003) assessed the performance of the Markov-switching model in replicating the NBER's dates through two different datasets, the growth in the quarterly real gross domestic product (GDP) and growth in monthly employment.

To follow, Chaveut and Hamilton (2005) estimated the ability of the multivariate version of the Markov switching model referring always to the USA. More recently, Moradi (2016) is a similar reference regarding developing countries, investigating the Iranian business cycle characteristics via univariate and multivariate Markov-switching and forming reliable chronology.

Krolzig (2003) devoted his attention to the question of cycle synchronization analyzing the case of specific country groups or, more in detail the Euro-zone business cycle. He provided an approach for the construction of turning points chronology and to evaluate the comovements and the synchronization of business cycles among seven euro countries by three generalized Markov switching vector autoregressive models with regime dependent intercepts and variance and using quarterly OECD GDP, OECD industrial production, and Eurostat GDP data, respectively.

Equally, Billio, Casarin, Ravazzolo, Van Dijk (2016) investigated the phenomenon of interconnection between booms and busts, duration and synchronization of regimes in eurozone and US economies. They considered USA and six largest European countries and dating turning points based on monthly industrial production growth rather than GDP by a Bayesian panel VAR (PSV-VAR) with intercept switching parametrization of the autoregressive model and a Gibbs sampling algorithm for posterior inference.

Bruno and Otranto (2008) went beyond the application of one single methodology and compared the performances of different dating procedures, included a MS version, with respect to ISAE Italian business cycle chronology after had reviewed literature proposals and classified the approaches into four categories:

[35] The classification rule is very simple and consists in converting the transition probabilities into turning point dates. Generally, recession turning points are designed in the period when the probability becomes greater than 0.5 and peak otherwise.

parametric and no, indirect versus direct detection of switch points[36]. Similarly, the (Schirwitz, 2009)'s comparative exercise included MS and suggested a consensus cycle dating for Germany.

Differently, Billio, Ferrara, Guégan, and Mazzi (2013) compared the ability of two non-linear parametric models (MS and TAR) to date and detect turning points. Using two different economic datasets, industrial production index and the unemployment rate, they showed that it is not possible to state the superiority of one model over the other but that it is preferable to use them in a complementary manner. Each model, in fact, explains a part of the analysis that the other does not do exactly because models differ in how model the movement between regimes.

Expressly, authors noticed that MS outperforms SETAR as it did not locate any false recession cycles even if encountered difficulties in identifying the double-dip recession of the early 1980s: " the two-regime SETAR model does not miss any recession episodes of the reference dating, but a false signal of recession is emitted between April and November 1977" (page 582).

Among the applications of TAR model to analyze business cycle asymmetry, it should be mentioned the pioneering works of Tiao and Tsay (1991), Tiao and Tsay (1994) and Teräsvirta and Anderson (1992). The former fitted a TAR model, divided quarterly US real GNP into the four regimes of contraction, contraction but improving, non-contraction but declining and expansion and demonstrated the different nature of economic activity over the different stages and a good out of sample forecast in terms of mean square error criterion.

The latter, using industrial production data introduced and estimated a TAR version, the STAR, to demonstrate business cycle asymmetry in 13 OECD countries and Europe.

In a like manner, recently Ferrara and Guégan (2005) applied this model to the Euro-zone industrial production index to detect the dates of troughs for the industrial business cycle referred to as the industrial recessions. Moreover, Osińska, Kufel, Błażejowski and Kufel (2016) examined and compared business cycles of Central and Eastern European economies to the entire EU business cycle using threshold variables, such as consumer price index, short and long interest rates, unemployment rate and an exchange rate vs. the U.S. Dollar.

In addition to its application to GDP dynamics, MS and TAR have also been fortunately implemented in other economics and finance areas, revealing a great versatility.

Linear models, for instance, cannot explain the tendency of volatility to increase more following large price falls than following price rises of the same magnitude, which is known as leverage effect. Cai (1994) and Hamilton and Susmel (1994) estimated models have suggested the practical relevance of the issue dealing with Treasury bills and a portfolio of stocks traded on the New York Stock Exchange.

Other different MS examples include Engel and Hamilton (1990) who dealt with swings in exchange rates, and Garcia and Perron (1996) who followed an autoregressive specification of order 2 to model shifts in U.S. real interest rate with two different data set from 1961 to 1986 and three possible regimes affecting mean and variance: $\mu(s_t) = \alpha_0 + \alpha_1 S_{1t} + \alpha_2 S_{2t}, \sigma(S_t) = \omega_0 + \omega_1 S_{1t} + \omega_2 S_{2t}$.

Among successive works, Kaufmann (2002) who assessed within a Bayesian framework the time variation of monetary policy through the first difference of the 3- month interest rate and, Guidolin (2012) that surveyed several financial applications included returns, currencies, derivatives, portfolio choice, and asset pricing.

With the same aim, Cao and Tsay (1992) explored the use of TAR in characterizing monthly volatility series of S&P stock returns aware that certain features of the series such as clusters of outliers cannot be described as linear models.

Lately, threshold auto-regression model was used in the estimation of the pass-through effect in the Croatian economy. Posedel (2009) specified a model according to which import prices respond a one percent change in the nominal exchange rate of the German mark in an asymmetric way around a threshold. Below hypothetical threshold, change in nominal exchange rate does not affect prices; above the threshold, the effect on inflation is strong and significant.

[36] Indirect method consists in detecting turning points on different time series according some specific rules and then aggregating; direct is founded on calculating at first a composite indictor based on different economic variables and, subsequently, deriving turning points on the latter.

Continuing, Ordóñez, Sala, and Silva (2011) adapted Smooth transition regression (STR) to prove the effect of shocks in oil prices (transition function) on US unemployment and in the general labor market. And Nguyen and To (2016) found a strong evidence of a nonlinear relationship between foreign direct investment (FDI) and economic growth in Asian countries using, again, through a Threshold Autoregressive (TAR) model and two threshold levels of FDI.

In spite of the number of works focused on the asymmetric effects of the business cycle, few are the empirical evidence about the likelihood that macro conditions influence loan quality differently in the different phases. Among works that, to our knowledge, adopted non-linear approaches to picture the influence of explanatory variables on NPLs, one needs to mention Gasha and Morales (2004). They applied the SETAR model to NPLs and GDP data of Argentina, Australia, Colombia, El Salvador, Peru and the United States, showing that there are tipping points of GDP growth, different from country to country and bank to bank, below which GDP has a significant impact on NPLs. The threshold values for which a downtrend in the economic activity makes NPLs rise more likely result below a certain rate of about 1–2 percent of GDP growth.

We also should mention Marcucci and Quagliarello (2008). Authors did not focus on NPLs in the analysis between credit risk and the business cycle, preferring the default rate as a more accurate indicator of bank's portfolio riskiness and the output gap as a proxy for the business cycle. However, again, they resort to threshold regression models at both the aggregate level and bank level to quantify the magnitude of the business cycle in bad economic times and to test the hypothesis that its impact on NPLs is more severe in the recession than in expansionary phases.

Differently, authors through light on the relationship between the business cycle and credit risk subject to regime switches, stepping forward on this issue and calculating the thresholds at which the model switches from one regime to the other.

Results were estimated on Italian data and performing three models:

1) threshold regression model for the aggregate time series of default rate (the threshold variable can be either the dependent or the independent variable and it allows the existence of two regimes)

$$dr_t = (\beta_{01} + \beta_{11} GAP_{t-1}) I(GAP_{t-1} \leq \gamma) + (\beta_{02} + \beta_{12} GAP_{t-1}) I(GAP_{t-1} > \gamma) + \epsilon_t \tag{39}$$

2) the threshold model for panel data with 2 or more regimes defined over the same single threshold variable[37] (the default rate at first and business cycle regimes later[38])

$$dr_t = \alpha_1 \ln(TA_{it}) + \cdots + \alpha_4 lgr_{it} + \beta_{11} GAP_{t-1} I(dr_{it} \leq \gamma_1) + \beta_{12} GAP_{t-1} I(dr_{it} > \gamma_1) + \epsilon_t \tag{40}$$

and

$$dr_t = \alpha_1 \ln(TA_{it}) + \cdots + \alpha_4 lgr_{it} + \beta_{11} GAP_{t-1} I(GAP_{t-1} \leq \gamma_1) + \beta_{12} GAP_{t-1} I(GAP_{t-1} > \gamma_1) + \epsilon_t \tag{41}$$

(where TA and gr are the total assets and the loan growth rate);

3) and the threshold model for panel data with 4 regimes determined through two different threshold variables (the micro variable default rate and the GAPT, the business cycle indicator)[39]

[37] In this model, observations are divided into 2 regimes, good or bad bank, depending on whether the default rate of the i bank at time t is smaller or larger than the threshold.

[38] In this model, GAPT indicator characterizes the threshold variable and depending on whether it is lower or greater a certain value, the first regime is represented by recession and the second by upturn conditions.

[39] This model allows differentiating between less risky and riskier banks in growing declining conditions.

$$dr_t = \alpha_1 \ln(TA_{it}) + \cdots + \alpha_4 lgr_{it} + \beta_{11} GAP_{t-1} I(dr_{it} \leq \gamma_1) I(GAP_{t-1} \leq \gamma_2) + \beta_{12} GAP_{t-1} I(dr_{it} \leq \gamma_1) I(GAP_{t-1} > \gamma_2) + \beta_{13} GAP_{t-1} I(dr_{it} > \gamma_1) I(GAP_{t-1} \leq \gamma_2) + \beta_{14} GAP_{t-1} I(dr_{it} > \gamma_1) I(GAP_{t-1} > \gamma_2) \epsilon_t \quad (42)$$

As an example of the different impacts, outcomes revealed portfolio quality an important variable in the impact of business cycles. Banks with sound asset quality are less affected by the overall economic conditions than lower ones and specifically, at the aggregate level, there is evidence that during riskier periods and when asset quality is lower, macroeconomic conditions have a negative statistically significant impact on credit risk. By contrast, in favorable periods and when the default rate is below the threshold, the macro impact is almost nonexistent and not significant.

The evidence is confirmed from models with two or more regimes with one threshold variable: both good and bad banks are affected by the business cycle, but the lower the banks' asset quality the stronger the influence of the business cycle on credit risk. Likewise, the relationship is stronger in recessionary conditions rather than in booms when GAPT is used as the threshold variable.

Finally, the four-regime model, in a more comprehensive manner, shows that riskier banks are affected by the business cycle during recessions almost five times higher than what we have during expansion. In addition, the impact of macroeconomic conditions on riskier banks during an economic downturn is more than three times as much as that for sound banks.

In addition, the impact of the business cycle on credit risk for less risky banks is almost the double of that similar banks experiment in booming phases; and the impact of the business cycle on good and banks riskiness during expansionary phases is substantially the same.

A panel threshold NPL ratio-growth model is also the methodological path followed by Mohaddes, Raissi, and Weber (2017). Encouraged by cross-board experiences about the weight of rapid growth for cutting NPLs, the paper analyzed a panel of 17 Italian disparate regions over the period 1997–2014, establishing the necessity of a persistent real GDP growth above 1.2 percent to reduce significantly Italian NPLs ratio (by about 6.5 to 9.5 percent per year).

Anyway, we saw as threshold regression models are just a variant of the non-linear model to take into account non-linearity in time series.

Overall, the current literature on the evolution of credit quality over business cycle switches is only a partial picture.

Other non-linear formulations that characterize the dynamics of a variable include, as reported before, the leading method of the Markov switching model.

Anyway, there is a paucity of related empirical evidence in our interest research's area; the use of the Markov switching model is scarce among researchers.

By contrast, credit risk modeling literature ordinarily employs the Markov Chain approach with transition matrices probabilities to show the evolvement of creditor quality and to quantify the rating migration (the probability one rating class of borrowers migrates to another).

Lu and Kuo (2006), for instance, adopted a Markov chain model to estimate the default probability of thirty one domestic banks in Taiwan, and to compare differences of the default probability of banks whether partecipating in financial holding or not. Results did not find significant difference in credit risk management.

Other studies in the same field have investigated the dependence of rating transition probabilities on the business cycle so taking into account even the above-mentioned asymmetric effects of the economic performance. Just for an example, Nickell, Perraudin, and Varotto (2000) used Moody's data from 1970 to 1997 and an ordered probit approach to compare the dynamics of credit standing across different type of obligatory and stages of the business cycle, founding that the business cycles dimension is the most important in explaining rating changes. In the same way, Bangia et alt. (2002) analyzed S&P rating to test whether transition probabilities develop differently across recessionary and expansionary conditions, concluding that downgrading probabilities are significantly higher during recessions, especially for lower quality credit, while upgrade probabilities remain stable or even decrease during contraction. And on this bases and in line with previous studies, Jones (2005) estimated transition matrices in conditions of aggregate data[40] when individual transitions are not available, finding proofs of transition matrices that change over the business cycle and providing different transition matrices for expansions and contractions.

Conversely, Coe (2002) referred to a bivariate Markov switching model to timing the US financial crisis of the 1930s using montly data on the growth rate of deposit current ratio and the yield spread from 1919 to 1941. Evidences revealed that crisis did not begin with the 1929 stock market crash, but with the first banking panic in late 1930 and continued until the introduction of the Federal Deposit Insurance in January 1934.

A more extensive usage of the methodology is evident only recently in Chevallier (2012) who adopted the MS framework to investigate the interactions between global imbalances, credit spreads, housing markets, macroeconomic indicators, commodities and equities during 1987q1- 2011q1. After a primary investigation of data trough univariate MS models, the linkages between variables have been explored in a multivariate MS environment and findings can be summarized along three main scenarios.

Scenario 1 features the importance of global imbalances in leading to the development of financial crisis and credit market freezes. Surprisingly, global imbalances are not found to be the root of financial crisis but to be statistically impacted by changing market conditions in credit and commodity such as gold and oil. Besides, commodities are found to be negatively related but to statistically impact credit spread.

Scenario two investigates the relationships between global imbalances, depressing conditions in US mortgage and housing markets, and global equity markets. Author documents that rising global imbalances and the uncontrolled development of the US house and mortgage markets have been deeply destabilizing the economy with various shocks affecting subsequently equity markets (the S&P 500).

Finally, scenario three tests the consequences of the financial crisis on the real economy and proves a significant negative impact exerted on industrial production by mortgage, credit spread and housing market variables.

Therefore, despite these applications, and turning now to our investigation concern, the Markov model could be an additional, less used tool to take into account the impact of the macro scenario on expected losses. It will help to consider the possibility that the behavior of financial and economic series change over time, yielding the GDP growth effect (and any other variable) on NPLs per each state and so, two estimates for each period.

For all we know, until to date, this methodology has only been applied to check these dynamics in Barbados (Grosvenor and Guy, 2013). The authors focused on the global banking system and on its six individual bank constituent. Conclusions were in favor of regime switching approach as a methodology to analyze NPLs reaction at GDP growth, inflation and loan growth as well as bank size[41].

The expected negative association between NPLs and GDP growth was obtained for the overall banking system while non-homogeneity resulted among institutions during positive phases. GDP growth resulted more important in a period of low NPLs (regime 0) with respect to high NPLs (regime 1), meaning that in a period of expansion NPLs decrease higher than the rate at which they do during a recession. Regarding the other variable, in regime 1 inflation rate is correlated positively with NPLs and in regime 0 a higher inflation rate leads to lower NPLs; period of loan growth, lastly, are associated positively with lower NPLs (the result of economic boom).

The non-linear procedure was also adopted for predicting NPLs transition probabilities from one state to the next one period ahead.

[40] The different classes of rating A, B, C, and D were associated to the categories of performing, due 30–89 days, nonaccrual plus due 90 days or more, and loss.

[41] Likelihood ratio (LR) tested the presence of non-linearity in the series.

In addition to the paucity of nonlinear analysis in the examination of NPLs, an underestimated area in the field, according to our knowledge, regards the theme of forecasting this ratio and determine relations among macroeconomic factors and the quality of loans in later periods.

In fact, researches have mainly tended to focus on intertemporal trends in economic conditions and risk growth in any given time to deduce a relationship retroactively; they have studied how selected macro shocks have shaped soundness indicators simultaneously or with some delay. The little attempts to predict problem loans preferred, again, the employment of linear models; an exception is Betancourt (1999) and Grosvenor and Guy (2013).

The former indicated and proved the Markov Chain as a reasonable approach for estimating loan losses although the requirement of strong assumption about stationary transition probabilities and homogeneous payment behavior. The author estimated loan losses for a portfolio of single-family residential mortgages simply multiplying a start vector B_0 of mortgages at time 0[42] times the transition matrix P containing the probability a mortgage in one category at time n shifting or remaining in the same category in the following month. The product effect was a forecast vector B_1 of the mortgages distribution at time $t + 1$, and can be generalized to obtain an infinite number of forecasts using the most recently forecasted distribution as a started vector B_0.

Speaking about the linear applications, in his work on loan quality, Gambera (2000) questioned the need for linear model to predict bank's financial conditions and defined "a linear model a very easy forecasting tool" due to the simplicity of estimation, a simple OLS, and due to the flexibility in representing interactions between more variables (the case of VAR model). The investigation of how simple linear multivariate models and bivariate vector-autoregressive systems can be used to predict problem loans ratio shows that the past value of the dependent variable and a number of regional and national macroeconomic variables such as bankruptcy filings, farm income, housing permits and unemployment are often good predictors of USA large banks problem loans ratios.

Subsequently, Babouček and Jančar (2005) investigated the transmission of macro variables on the development of the Czech economy creating a forecasting through the VAR methodology and Chase, Greenidge, Moore, and Worrell (2005) and Greenidge and Grosvenor (2010) forecasted Barbados banking performance. The former employed a simple Ordinary Least Squares and used the Treasury bill rate, the consumer price index, real GDP and a lagged dependent variable as explanatory variables. The latter, in order to evaluate which model performed best, preferred to hark back to the univariate ARIMA model, ARDL multivariate model, and a combined model[43]. Conclusions from the comparison work favored the multivariate suggesting that the impact of macro factors on NPLs level should be considered and accounted for the forecasting models. Nevertheless, in all cases, the combination model outperformed both models; showing small forecast errors, it revealed to be the most preferable alternative.

[42] The model development assumes at any month n, a mortgage can be classified into one of the following categories: (1) Active, (2) Thirty days delinquent, (3) Sixty days delinquent, (4) Ninety plus days delinquent, (5) foreclosure, (5) Real estate owned (REO) and (6) Paid off.

[43] In this one, dynamic out of sample and static in sample forecasts of NPLs resulted by regressing the actual ratios on the respective forecast values from the preferred univariate or multivariate model of the aggregate and individual banks.

Conclusion

To sum up, most discussions about loans quality begin, reasonably, with an inspection of the firm characteristic of the financial institutions and bank's policies in terms of capitalization, cost efficiency, diversification, and credit growth and so on. Often, as instance, bank vulnerability derives from banker's bad decisions and ability to assess borrowers' creditworthiness, and the business success and capability to satisfy the financial obligations depends on the role and the experience of the management.
However, the focus on the properties of the institutions is incomplete and unclear.
The quality of a bank institution undoubtedly interacts with regulatory and macroeconomic factors.
Banks strength is determined by its investment and funding decisions that in turn are influenced by supervisory and regulatory structures and macroeconomic factors. When macroeconomic forces put sizeable strain on the banking system, the most fragile banks are the ones most likely to fail, but it is the macroeconomic tension, as much as the weakness of individual banks, that causes failures and instability.

Considering the existing literature about the macro drivers of bad debts, this first chapter is focused on its schematization and systematization according to a methodological perspective.
It lays the groundwork for the next by outlining a critical presentation of the extant ways to estimate the nexus between economic activity and asset quality, and individuating some critical lacuna as regard the methodological procedure.
Findings suggest the clearly recognized shortcomings of linear methods and a slightly lower number of applications deviated from those approaches. Although a growing empirical literature is using threshold models, there lack empirical works on the determinant of NPLs that employ the Markov regime-switching framework (an exception is for example the Barbados country-specific study).
Therefore, non-linear methods are presented as a set of suitable alternative models to represent the dynamics of bad debts and specifically, the features of the Markov regime switching approach are considered interesting.
It is suggested as a clear improvement on current methods thanks to its comparative advantage and capability to outperform all linear specifications in capturing the natural asymmetries in NPLs-macro environment nexus.
Indeed, as noted, structural breaks put under scrutiny the assumed linearity of models and adopting such functional forms might not report robust results and affect forecasting.
Social, economic and political changes make it difficult to model macro and financial variables with constant parameter linear models; they make it is unreasonable to expect a single linear model to describe the distinct behaviors. GDP, volatility, interest, and exchange rates are the classical examples of non-linear time series models; similarly, unemployment rate rises more rapidly in recession and declines slowly in expansion but traditional analysis of the NPLs continues to propose a single state relationship.
On the contrary, nonlinear approaches, relaxing the assumption of a single state and involving multiple equations that characterize the time series behaviors in different regimes, account for the possibility that the impact of the business cycle and other macro variables is dissimilar over different phases and capture better the related magnitude.

Accordingly, the chapter is based on the intuition that further studies could be developed to link the macro-economic environment and the response of credit quality.
Taking Grosvenor and Guy (2013) into account, further studies could be implemented in a different setting, with a different sample and data, considering alternative explanatory variables, over a different timescale and examining the determinants of NPLs across different loan categories, rather than the aggregate level of NPLs.
A rethinking and extending dataset could help to confirm the generalization of previous results and to evaluate the suitability of non-linear approach compared to linear models, generating meaningful progress in the context of the thesis.

Moreover, given that forecasting is recommended as a second deserve search path due to few preceding types of studies, extra researches could be undertaken to consider the nonlinear effects in a forecasting model.
Considering the above definition of NPLs as a measure of a country financial stability, it is not difficult to understand the reason why being able to forecast NPLs is necessary and the issue could be challenging and intriguing.
Finally, a time-consuming approach could be based on a methodological comparison exercise between linear and non linear models for predict NPLs ratio and their relation with macro variables. Actually, many studies adopted just a single linear procedure and the paralleling will help to overcome the possibility that results could be influenced by the model specification or sample chosen. At the same time, it will be useful to value what models best replicate business cycle features and to demonstrate whether the predicted exogenous variables are able to affect the amount of NPLs.

Against this background, the next chapter delves into the empirical part of the research.
Our contribution will be merely methodological.
Focusing on the American context and relying on a Markov switching model, we will forecast and unravel non-linearities in the relationship among credit risk with meaningful variables, and we will try to compare the results of the research with past studies in other geographies or timeframes.
We will check whether the impact of macro variables is greater in terms of high NPLs (state 1) rather than low NPLs (state 2). Moreover, given that literature proved both the long calm nature of expansion compared to recessions and the negative relationship between GDP and NPLs, the next chapter will be useful to test whether periods of low NPLs are longer than those with high NPLs (such duration is not captured by linear models). Likely, the likelihood of transitioning from one state to the next.

APPENDIX A: The Hamilton Filter

The classical maximum likelihood estimation method (MLE), detailing the likelihood function by the EM algorithm (Hamilton, 1990), is the parameter estimation tool generally used by the existing literature.

Considering the 2 regimes MS and a first-order autoregressive model $y_t = a_{0st} + a_{1st}y_{t-1} + \epsilon_t$, the aim consists in maximizing for a given value of θ the conditional log-likelohood function:

$$L(\theta) = \sum_{t=1}^{T} \ln f(y_t|\Omega_{t-1};\theta) \tag{1}$$

where $f(y_t|\Omega_{t-1};\theta)$ represents the conditional density function of y_t given the information available in $t-1$. Specifically, as s_t is unobserved, the aim of the estimation procedure is not only the parameters of the autoregressive model in the different regimes, $\theta = (a_{01,}a_{02,}a_{11,}a_{12,}\sigma^2)$, but also the probabilities with which each state occurs at each point in time.

Construction of the conditional log-likelihood function then requires the construction of the conditional density function. Hamilton (1989) developed a non-linear recursive algorithm (the **Hamilton filter)** in which the likelihood function is computed as a by-product of the filter starting with a supposing initial value of $P(s_{t-1} = j|\Omega_{t-1};\theta)$.

In fact, generally, the maximum likelihood estimator $\widehat{\theta_{MLE}}$ is obtained by setting the first derivative of the log likelihood function to zero. However, it may happened that a closed form solution is not obtainable and this orders a nonlinear numerical optimization procedure to maximize the log likelihood function: given initial estimates of the transition probabilities, new estimates are derived iterating the process until obtaining the value of the parameter that maximize the log likelihood function. The regime probabilities are then inferred in a second step.

Conditional density functions can be decomposed as the weighted sum of densities, conditioned on the state, where the weights are state probabilities by applying the total probability theorem:

$$f(y_t|\Omega_{t-1};\theta) = \sum_{i=1}^{2} P(s_t = i|\Omega_{t-1};\theta) f\left(y_t\middle|s_t = i, \Omega_{t-1;}\theta\right) \tag{2}$$

where 2 refer to the different states of the economy.

The term $f(y_t|s_t = i, \Omega_{t-1})$ follows a Gaussian distribution as errors ε_t are assumed to be normally distributed with the same mean and variance of the error.

$$f(y_t|s_t = i, \Omega_{t-1};\theta) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[\frac{-\left(y_t-\alpha_j-\beta_j y_{t-1}\right)^2}{2\sigma^2}\right] \tag{3}$$

The other part, instead, can be once again decomposed using the total probability theorem:

$$P(s_t = i|\Omega_{t-1};\theta) = \sum_{i=1}^{2} P(s_t = i|s_{t-1} = j, \Omega_{\ t-1};\theta)\, P(s_{t-1} = j|\Omega_{t-1};\theta) \tag{4}$$

$$= \sum_{i=1}^{2} p_{ij}\ P(s_{t-1} = j|\Omega_{t-1};\theta) \tag{5}$$

Having $f(y_t|\Omega_{t-1};\theta)$, the next step is to compute $f(y_{t+1}|\Omega_t:\theta)$. This means that we have to update $P(s_t = i|\Omega_{t-1};\theta)$ to reflect the information containing f in y_t as we require $P\left(s_t = i, \Omega_{t;}\theta\right)$. It can be expressed by applying the Bayes's theorem.

$$f\left(y_{t+1}|\Omega_{t;}\theta\right) = \sum_{i=1}^{2} P(s_t = i|\Omega_t;\theta) f\left(y_t \middle| s_t = i, \Omega_{t;}\theta\right) \tag{6}$$

$$P(s_t = i|\Omega_t;\theta) = \sum_{i=1}^{2} P(s_t = i|s_{t-1} = j, \Omega_{\ t};\theta)\, P(s_t = i|\Omega_t;\theta) \tag{7}$$

$$= \sum_{i=1}^{2} p_{ij}\ P(s_t = i|\Omega_t;\theta) \tag{8}$$

and $$P(s_t = i|\Omega_t;\theta) = P\left(s_t = i|y_{t,}\Omega_{t-1};\theta\right) = \frac{P(s_t=i|\Omega_{t-1};\theta) f(y_t|s_t = i, \Omega_{t-1;}\theta)}{\sum_{i=1}^{2} P(s_t=i|\Omega_{t-1}) f(y_t|s_t = i, \Omega_{t-1})} \tag{9}$$

Basically, the Hamilton filter procedure easily calculates iteratively for $t = 1, \dots T$ equations (8)-(9) to finally gives the log likelihood function as a function of just the parameters to estimate and the initial state of the economy in the original period, which can also be expressed as a function of the parameters to estimate.

Therefore, the unknown parameters can be numerically found maximizing that function, and value of parameters that maximizes the loglikelihood function $\widehat{\theta_{MLE}}$ can be obtained using standard numerical techniques.

The filtered probabilities (the probability of being in state j at each period of time, based on the information up to that date) are a readily available result of this method.[44]

The only unsolved element regards the initial value of the transition probabilities needed to start the filter.

Assuming the Markov chain as stationary and ergodic, (Hamilton, 1994) shows that the usual practice to initiate the filter consists of using the unconditional distribution of the first observation, that is:

$$P(s_{t-1} = j|\Omega_{t-1};\theta) = P(s_0 = j|\Omega_0;\theta) = P(s_0 = j) \tag{10}$$

Specifically, in the two regime case, as an example, the eigenvalue of the transition matrix are the solutions of the characteristic polynomial $|P - \lambda I| = 0$.

$$\begin{vmatrix} P_{11} - \lambda & 1 - P_{22} \\ 1 - P_{11} & P_{22} - \lambda \end{vmatrix} = 0 \tag{11}$$

$$(P_{11} - \lambda)(P_{22} - \lambda) - (1 - P_{22})(1 - P_{11}) = 0 \tag{12}$$

$$P_{11}P_{22} - (P_{11} + P_{22})\lambda + \lambda^2 - 1 + P_{11} + P_{22} - P_{11}P_{22} = 0 \tag{13}$$

[44] Additionally, it can be computed the probability of being in state j at each period of time based on the information of the whole sample, the so called "smoothed probabilities" $P(s_t = i|\Omega_T)$. (Kim, 1994) proposed the algorithm.

$$\lambda^2 - (P_{11} + P_{22})\lambda - 1 + P_{11} + P_{22} = 0$$

(14)

$$(\lambda - 1)(\lambda + 1 - P_{11} - P_{22}) = 0$$

(15)

Thus, the eigenvalues for a two-state chain are $\lambda_1 = 1$ and $\lambda_2 = -1 + P_{11} + P_{22}$. The associated eigenvectors and thus the unconditional probabilities that the process will be in regime1 or 2 at any given date are:

$$\begin{cases}(P_{11} - 1)v_1 + (1 - P_{22})v_2 = 0\\(1 - P_{11})v_1 + (P_{22} - 1)v_2 = 0\end{cases}$$

(15)

$$v_{1=} - \frac{(1 - P_{22})}{P_{11} - 1} v_2$$

(17)

$$v_{1=} 1 - P_{22}$$

(18)

$$v_2 = 1 - P_{11}$$

(19)

Since we know $P_{11} + P_{22} = 1$

$$\mathrm{P}(\mathrm{s}_0 = 1) = \frac{1 - p_{22}}{2 - p_{11} - p_{22}}$$

(20)

$$\mathrm{P}(\mathrm{s}_0 = 2) = \frac{1 - p_{11}}{2 - p_{11} - p_{22}}$$

(21)

Hamilton's algorithm is composed of two stages. In the first stage, he estimated the parameters that correspond to the probability density of the state variable. Secondly, filtered and smoothed probabilities are estimated.

A different approach to estimating is the method of Gibbs Sampling introduced by Geman and Geman (1984), as referred to in Casella and George (1992). It is a recursive Bayesian estimation method that treats parameters as random variables and which requires the only knowledge of the full conditional posterior distribution of the parameters also known conditional posterior distribution[45] in the Bayesian analysis. It is particularly attractive because fractionate a high-dimensional estimation problem into lower dimensional ones via full conditional distributions of the parameters.

The algorithm produces random samples from the joint density of a group of random variables by repeatedly sampling to form estimates of the parameters of interest. It starts from some arbitrary set of values of the parameters and then samples alternatively from the density of each element of the parameter vector conditional on the value of the other element sampled in the previous iteration and the data.

Specifically, dividing the parameter vector into j groups, $\theta = (\theta_{1,} \ldots \theta_j)$ given the observed data Ω_T and the full conditional posterior distributions

$$p(\theta_1 | \Omega_T, \theta_2, \ldots, \theta_j)$$

(22)

$$p(\theta_2 | \Omega_T, \theta_1, \theta_{3,} \ldots, \theta_j)$$

(23)

$$p(\theta_j | \Omega_T, \theta_1, \ldots, \theta_{j-1})$$

[45] The conditional posterior distribution is the probability of the parameter given the data $p(\theta|y) \sim p(\theta|y)p(\theta)$. It is proportional to the likelihood function or the probability of the evidence given the parameters and the prior density, that is our initial belief on possible values of θ.

(24)

then, conditional on initial values$\theta_1^0,\ldots,\theta_j^0$, we can obtain a realization of θ_1, denoting θ_1^1 from the full conditional distribution $p(\theta_1|\Omega_T,\theta_2^0,\ldots,\theta_j^0)$.

Randomly we can draw a realization of θ_2, denoting θ_2^1 from $p(\theta_2|\Omega_T,\theta_1^1,\theta_3^0,\ldots,\theta_j^0)$, a realization of $\theta_{3,}$ denoting θ_3^1 from $p(\theta_3|\Omega_T,\theta_1^1,\theta_2^1,\theta_4^0\ldots,\theta_j^0)$ and so on to draw θ_j^1

This procedure can be iterated to obtain $\theta_1^n,\ldots,\theta_j^n$ for n=1,…N.

The element of θ^n form a Markov chain and for a large number of iterations $n \geq N$, this draw can be regarded as taken from the true joint posterior distribution and the estimates are accurate. Geman and Geman (1984) showed that under mild conditions that the Gibbs sequence converges in distribution to the true distribution of θ: $\theta^N \xrightarrow{d} p(\theta|\Omega_T)$. The regularity conditions essentially require that for an arbitrary starting value ì, the prior Gibbs iterations have a chance to visit the full parameter space.

In the context of two regime Markov switching regression, the parameter vector θ can be divided into three groups considering that there are also the state variable and the transition probabilities to be treated as parameters; $\theta_1 = (a_{01,}a_{02,}a_{11,}a_{12,}\sigma^2)$, $\theta_2 = (p_{00,}p_{11})$,$\theta_3 = (s_{1,}s_2)$.

APPENDIX B: Threshold Model Estimation

The estimation of TAR models require determining the threshold order or lag length and the delay parameter. The simplest but not very appropriate method of determining the lag length is to assume that the same number of lags for each of the regimes. It is unlikely that as data have different behavior in different states. An alternative method involves simultaneous selection using an information criterion. The delay parameter can be determined along with the lag orders for each of the regimes or simply can set equal to the most recent past value of the state variable.

Under this assumption, the Hansen (1996)'s method of sequential conditional least squares is the easiest one for estimation to estimate parameters.

Starting from the model specification

$$y_t = \left(\alpha_0 + \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p}\right) I(y_{t-k} < r) + (\beta_0 + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p}) I(y_{t-k} \geq r) + u_t \tag{1}$$

he proposed the alternative representation as

$$y_t = x_t' \alpha \, I(y_{t-k} < r) + x_t' \beta I(y_{t-k} \geq r) + u_t \tag{2}$$

where $\alpha = (\alpha_0, \alpha_1, \dots . \alpha_p)'$, $(\beta_0, \beta_1, \dots . \beta_p)'$ and $x_t = (1, y_{t-1}, \dots, y_{t-p})'$, resulting in

$$y_t = x_t(r)' \theta + u_t \tag{3}$$

with $\theta = (\alpha' \beta')'$ and $x_t(r) = (x_t' \, I(y_{t-k} < r), x_t' I(y_{t-k} \geq r)'$

The parameter of interest became r and θ and since the equation to estimate is a regression, even if nonlinear, the appropriate method is OLS. For a given threshold, the estimate of θ is:

$$\widehat{\theta(r)} = \sum_{t=1}^{n} (x_t'(r) x_t(r))\,'(\sum_{t=1}^{n} x_t(r) y_t) \tag{4}$$

where the residuals of the model are $\widehat{u_t} = y_t x_t(r)' \hat{\theta}$ and residual variance is $\sigma^2(r) = \frac{\sum_{t=1}^{n} u_t^2(r)}{n}$.

The least squares of r are the one that minimizes the residuals variance over the set C of all possible values of the threshold coefficients $\hat{r} = argmin \sigma^2(r)$ with $r \epsilon C$. Finally, the estimates of the autoregressive parameter is found as $\hat{\theta} = \hat{\theta}(\hat{r})$.

REFERENCES

Babouček and Jančar. (2005). A VAR Analysis of the Effects of Macroeconomic Shocks to the Quality of the Aggregate Loan Portfolio of the Czech Banking Sector. *CNB Working Paper Series.*

Berger and DeYoung. (1997). Problem Loans and Cost Efficiency in Commercial Banks. *Journal of Banking and Finance*.

Casella and George. (1992). Explaining the Gibbs Sampler. *The American Statistician, Vol. 46, No. 3*, 167-174.

Garcia and Perron. (1996). An Analysis of the Real Interest Rate Under Regime Shifts. *The Review of Economics and Statistics, vol 78*, 111-125.

Jones. (2005, November). Estimating Markov Transition Matrices Using Proportions Data: An Application to Credit Risk. *An Application to Credit Risk.*

Krolzig. (2003). Construction of Turning Point Chronologies with Markov switching Autoregressive models: the Euro-zone business cycle. *Department of Economics and Nuffield College, Oxford University*.

Tiao and Tsay. (1994). Some Advances in Nonlinear and Adaptive Modeling in Time-series. *Journal of Forecasting 13*, 109-131.

Anas, Billio, Ferrara, Lo Duca. (2007). Business Cycle Analysis with Multivariate Markov Switching Models. *Working Papers Working Papers Ca' Foscari University of Venice, n.32*.

Arpa. (2001). The influence of macroeconomic developments on Austrian banks: implications for banking supervision. *BIS Papers, 1*, 91-116.

Bangia et alt. (2002). Ratings Migration and the Business Cycle, With Applications to Credit Portfolio Stress Testing. *Journal of Banking & Finance, vol 26, n. 2/3*, p. 235-64.

Barisitz. (2013). Nonperforming Loans in Western Europe – A Selective Comparison of Countries and National Definitions. *OENB, Focus on European Economic Integration Q1/13*.

Bauwens, Preminger and Rombouts. (2010). Theory and Inference for a Markov Switching GARCH Model. *Econometrics Journal, 13(2)*, 218-244.

Beck, Jakubik and Piloiu. (2013). Non performing Loans: What Matter in Addition to the economic Cycle? *European Central Bank Working Paper*, n.1515.

Betancourt. (1999). Using Markov Chian to Estimate Losses from a Portfolio of Mortgages. *Review Of Quantitative finance and Accounting* , p. 303-317.

Bexley and Nenninger. (2012). Financial Institutions and the Economy. *Journal of Accounting and Finance, 12(1)*, 221-254.

Bholat. (2016). Non-Performing Loans: Regulatory and Accounting Treatments of Assets. *Bank of England Working Papers n.594*.

Bikker and Metzemakers. (2002). Bank provisioning behavior and procyclicality. *De Nederlandsche Bank Research Series Supervision No. 50*, 1-21.

Billio, Casarin, Ravazzolo, Van Dijk. (2016). Interconnection between eurozone and US booms and busts using a Bayesian panel Markov-Switching Var model. *Journal of Applied Econometrics, 31*, 1352-1370.

Billio, Ferrara, Guégan and Mazzi. (2013). Evaluation of Regime Switching Models for Real-Time Business Cycle Analysis of the Euro Area. *Journal of Forecasting, 32*, 577–586.

Bodman and Crosby. (2000). Phases of the Canadian Business Cycle. *The Canadian Journal of Economics, vol 33, n3*, 618-633.

Bofondi and Ropele. (2011). Macroeconomic determinants of bad loans: evidence from Italian banks. *Bank of Italy, Occasional Papers*.

Bollerslev. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics,31*, 307-327.

Bruno and Otranto. (2004). Dating the Italian Business Cycle: A comparison of procedures . *ISAE Working Paper 41*.

Bruno and Otranto. (2007). models to date the business cycle: The Italian case. *Economic modelling, 25*, 899-911.

Bruno and Otranto. (2008). Models to date the business cycle: The Italian case. *Economic Modelling 25*, 899-911.

Cai. (1994). A Markov Model of Unconditional Variance in ARCH. *Journal of Business and Economic Statistics, 12,* 309-316.

Cao and Tsay. (1992). Nonlinear Time-Series Analysis of Stock Volatilities. *Journal of Applied Econometrics, Vol. 7,* S165-S185.

Castro. (2012). Macroeconomic determinants of the credit risk in the banking system: The case of the GIPSI. *NIPE Working Papers*.

Chaibi and Ftiti. (2015). Credit risk determinants: Evidence from a cross-country study. *Research in International Business and Finance, 33*, p. 1-16.

Chan and Tong. (1986). On Estimationg Thresholds in autoregressive models. *Journal of Times Series Analysis, 7*, 179-194.

Chase, Greenidge, Moore and Worrell. (2005). Quantitative Assessment of a Financial System—Barbados. *IMF Working Paper*.

Chauvet and Piger. (2003). Identifying Business Cycle Turning Points in Real Time. *The Federal Reserve Bank of St. Louis*, 47-62.

Chevallier. (2012). Global imbalances, cross-market linkages, and the financial crisis: A multivariate Markov-switching analysis . *Economic Modelling 29*, 943-973.

Coe. (2002). Financial Crisis and The Great Depression: A Regime Switching Approach. *Journal of Money, Credit and Banking. Vol 34, N.1*, 76-93.

De Lis, Pagés and Saurina. (2000). Credit Growth, Problem Loans And Credit Risk Provisioning" In Spain. *Banco de España — Servicio de Estudios, Documento de Trabajo no 0018*.

De Young. (1999). Birth, Growth, and Life or Death of Newly Chartered Banks. *Federal Reserve Bank of Chicago, Economic Perspectives*, 18-35.

De Young and Hasan. (1998, May). The performance of de novo commercial banks: A profit efficiency approach. *Journal of Banking & Finance*, p. 565-587.

Diebold, Lee and Weinbach. (1994). *Regime Switching with Time Varying Transition Probabilities*. Hargreaves, C. (Ed.), Nonstationary Time Series Analysis and Cointegration.Oxford University Press.

Durland and McCurdy. (1994). Duration-Dependent Transitions in a Markov Model of U.S. GNP Growth. *Journal of Business & Economic Statistics, Vol. 12, No. 3*, 279-288.

EBA. (2016, December). *Risk Assessment of The European Banking System*.

Engel and Hamilton. (1990). Long Swings in the Dollar: Are they in the Data and Do Market Know It? *The American Economic Review, vol 80*, 689-713.

Engle. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of Variance of United Kingdom Inflation. *Econometrica, 50*, 987-1008.

Espinoza and Prasad. (2010). Nonperforming Loans in the GCC Banking System and their Macroeconomic Effects. *IMF Working Paper WP/10/224*.

European Parliament. (2016). *Non-performing loans in the Banking Union: stocktaking and challenges*.

Farhan, Sattar, Chaudhry and Khalil. (2012). Economic Determinants of Non Performing Loans: Perception of Pakistani Bankers. *European Journal of Business and Management*, 87-99.

Fatimoh. (2012). Impact of Corporate Governance on Banks Performance in Nigeria. *Journal of Emerging Trends in Economics and Management Sciences (JETEMS) 3(3)*, 257-260

Ferrara and Guégan. (2005). Detection of the industrial business cycle using SETAR models. *Journal of Business Cycle Measurement and Analysis, Vol. 2, No. 3*, 353-372.

Filardo. (1994). Business-Cycle Phases and Their Transitional Dynamics. *Journal of Business & Economic Statistics, Vol. 12, No. 3*, 299-308.

Filip. (2015). Non-performing loans – dimension of the non-quality of bank lending/loans and their specific connections. *Theoretical and Applied Economics, Volume XXI*, 127-146.

Filosa. (2007). *Stress testing of the stability of the Italian banking system: a VAR approach*. Tratto da http://www.economia.unimore.it.

Fofack. (2005). Non Performing Loams in Sub Saharan Africa: causal analysis and macroeconomic implications. *World Bank Policy Research Working Paper 3769*.

Foglia. (2009, September). Stress Testing Credit Risk: A Survey of Authorities' Approaches. *International Journal of Central Banking*.

Gambera. (2000). Simple Forecasts of Bank Loan Quality in the Business Cycle. *Federal Reserve Bank of Chicago*.

Gasha and Morales. (2004). Identifying Threshold Effects in Credit Risk Stress Testing. *IMF Working Paper*.

Gerlach. (2005, February). Bank Lending and Property Prices in Hong Kong. *Journal of Banking & Finance*, p. 461/481.

Ghosh. (2015, April 15). Banking-industry specific and regional economic determinants of non-performing loans: Evidence from US states. *Journal of Financial Stability*, p. 93-104.

Goldfeld and Quandt. (1973). A Markov Model for Switching Regressions. *Journal of Econometrics, 1(1)*, 3-15.

Gonzales. (2009). The Fundamentals of Procyclicality of the Financial System. *BSP Economic Newsletter, 3*, p. 1-5.

Gray. (1996). Modeling the conditional distribution of interest rates as a regime-switching. *Journal of Financial Economics, 42*, 27-62.

Greenidge and Grosvenor. (2010). Forecasting non performing loans in Barbados. *Journal of Business, Finance & Economics in Emerging Economies, vol5 issue 1*, 79-108.

Grosvenor and Guy. (2013). A Regime Switching Approach to Analyzing Bank Non-Performing A Regime Switching Approach to Analyzing Bank Non-Performing. *CBB Working Paper No. WP/13/7*.

Guidolin. (2012). Markov Switching Models in Empirical Finance. *Working Paper n. 415*.

Guttentag and Herring. (1984, December). Credit Rationing and Financial Disorder. *The Journal of Finance, Vol. 39, No. 5*, p. 1359-1382.

Hamilton. (1994). *Time Series Analysis.* Princeton University Press.

Hamilton. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica 57*, 357-384.

Hamilton and Chaveut. (2005). Dating Business Cycle Turning Points.

Hamilton and Susmel. (1994). Autoregressive Conditional Heteroskedasticity and Changes in Regime. *Journal of Econometrics, 64*, 307-333.

Hansen. (1996). inference in TAR models. *Studies in Non-Linear Dynamics and Econometrics*.

Hansen. (2000). Sample splitting and threshold estimation. *Econometrica, Vol. 68, No. 3*, 575-603.

Hoggarth, Sorensen and Zicchino. (2005, November). Stress tests of UK banks using a VAR approach. *Bank of England Working Paper n. 282*.

Hu, Li and Chiu. (2004, September). Ownership and non performing loans: Evidence from Taiwan's banks. *The Developing Economies*.

Huang, Kuan and Lin. (1998). Identifying the turning points and business cycles and forecasting real GNP growth rate in Taiwan. *Taiwan Economic Review, 26*, 431-457.

IMF: Bloem and Freeman. (2005). The Treatment of Nonperforming Loans IMF Committee on Balance of Payments Statistics. *Eighteenth Meeting of the*, (p. 4-15). Washington, D.C.

Kalirai, Scheicher. (2002). Macroeconomic Stress Testing: Preliminary Evidence for Austria. *Austrian National Bank Financial Stability Report, May, no. 3*.

Kasselaki and Tagkalakis. (2013). Financial soundness indicators and financial crisis episodes. *Working paper 158, Bank of Greece*.

Kaufmann. (2002). Is there an asymmetric effect on monetary policy over time? A Bayesian analysis using Austrian data. *Empirical Economics*, 277–297.

Keeton and Morris. (1987). Why Do Banks‟ Loan Losses Differ? *Economic Review*, p. 3-21.

Keeton, William R. (Second quarter, 1999). Does Faster Loan Growth Lead to Higher Loan Losses? *Federal Reserve Bank of Kansas City, Economic Review*.

Khemraj and Pasha. (2009). The determinants of non-performing loans: an econometric case study of Guyana. *3rd Biennial International Conference on Business, Banking and Finance.* St. Augustine, Trinidad and Tobago.

Kim. (1994). Dynamic linear models with Markov-switching. *Journal of Econometrics, 60*, 1-22.

Kim and Nelson. (1998). Friedman's Plucking Model of Business Fluctuations: Tests and Estimates of Permanent and Transitory Components. *Journal of Money, Credit and Banking, 31*, 317-334.

Kjosevski & Petkovski. (2016). Non-performing loans in the Baltic States: determinants and macroeconomic effects. *Baltic Journal of Economics, 17*, p. 25-44.

Klein. (2013). Non-Performing Loans in CESEE: Determinants and Macroeconomic Performance. *IMF Working Paper*.

Krolzig. (1997). *Markov-Switching Vector Autoregressions. Modelling, Statistical Inference, and Application to Business Cycle Analysis.* Springer.

Krueger. (2002). International Standards for Impairment and Provisions and their Implications for Financial Soundness Indicators (FSIs). *IMF Occasional Paper 212*.

Kuan. (2002). Lecture on the Markov Switching Model. *Institute of Economics, Academia Sinica, Taipei 115*.

Lata. (2014). Non-Performing Loan and Its Impact on Profitability of State Owned Commercial Banks in Bangladesh: An Empirical Study. *Proceedings of 11th Asian Business Research Conference*.

Leventis, Dimitropoulos, and Owusu-Ansah. (2013). Corporate Governance and Accounting Conservatism: Evidence from the Banking Industry. *Corporate Governance: An International Review, 2013, 23 (3)*, 264-286.

Levine and Zervos. (1998, June 3). Stock Markets, Banks, and Economic Growth. *The America Economic Review*, p. 537-558.

Louzis, Vouldis and Metaxas. (2011). Macroeconomic and bank-specific determinants of non-performing loans in Greece: a comparative study of mortgage, business and consumer loan portfolios. *Journal of Banking and Finance*.

Lu and Kuo . (2006). The Default Probability of Bank Loans in Taiwan: An Empirical Investigation by the Markov Chain Model. *Asia Pacific Management Review 11(2)*, 111-122.

Marcellino. (2007). A comparison of time series models for forecasting GDP growth and inflation. *IEP-Università Bocconi, IGIER and CEPR*.

Marcucci and Quagliarello. (2008, June). Credit risk and business cycle over different regimes. *Banca d'Italia- Working Papers*.

Marcucci and Quagliariello. (2008, February). Is bank portfolio riskiness procyclical? Evidence from Italy using a vector autoregression. *Journal of International Financial Markets, Institutions and Money, 18*, p. 46-63.

Mazzi and Scocco. (2003). *http://europa.eu.int/comm/euroindicators*.

Messai and Jouini. (2013). International Journal of Economics and Financial Issues. *Micro and Macro Determinants of Non-performing Loans*, 852-860.

Mohaddes, Raissi and Weber. (2017, March). Can Italy Grow Out of Its NPL Overhang? A Panel Threshold Analysis. *IMF Working Papers*.

Moradi. (2016). Modeling Business Cycle Fluctuations through Markov Switching VAR: An Application to Iran. *Munich Personal RePEc Archive*.

Nguyen and To. (2016). threshold effect in the relationship between foreign direct investment and economic growth: evidence from Asian countries. *Third Asia Pacific Conference on Advanced Research*. Melbourne.

Nickell, Perraudin and Varotto. (2000). Stability of Rating Transitions. *Journal of Banking & Finance, 24*, p. 203-227.

Nkusu. (2011). Nonperforming Loans and Macrofinancial Vulnerabilities in Advanced Economies. *IMF Working Paper*, No 11/161.

Ordóñez, Sala and Silva. (2011). Oil Price Shocks and Labor Market Fluctuations. *The Energy Journal, Vol. 32, n. 3*, 89-118.

Osińska, Kufel, Błażejowski and Kufel. (2016). Modelling and Forecasting Business Cycle in CEE Countries using a Threshold Approach. *Dynamic Econometric Models*, 145-164.

Pain. (2003). The provisioning experience of the major UK banks: a small panel investigation. *Bank of England Working Paper no. 177*.

Parven. (2011, November). Nonperforming loans of commercial banks in Bangladesh. *Munich Personal RePEc Archive*.

Podpiera and Weill. (2008). Bad luck or bad management? Emerging banking market experience. *Journal of Financial Stability. Vol. 4, Issue 2*, 135-148.

Posedel. (2009). Threshold Model of the Exchange Rate Pass-Through Effect: The Case of Croatia. *Eastern European Economics, Vol 47, n.6*, 43-59.

Potter. (1995). A non linear approach to US GNP. *Journal of Applied Econometrics*, 109-125.

Quagliariello. (2006). Banks' Riskiness over the Business Cycle: A Panel Analysis on Italian Intermediaries. *Discussion Papers Banca Italia*.

Quirici. (1996). Aspetti evolutivi nella gestione del rischio di credito alla luce del deterioramento della qualità dei prestiti bancari. *Studi e Note di Economia* , 67-96.

Ranjan and Dhal. (2003). Non-Performing Loans and Terms of Credit of Public Sector Banks in India: An Empirical Assessment. *Reserve Bank of India Occasional Papers*, vol.24, n.3.

Reddy. (2012). Impact of Inflation and GDP on stock market returns in India. *International Journal of Advanced Research in Management and Social Sciences*, 120-136.

Saba, Kouser and Azeem. (2012). Determinants of Non Performing Loans: Case of US Banking Sector. *The Romanian Economic Journal*, 125/136.

Salas and Saurina. (2002, December). Credit Risk in Two Institutional Regimes: Spanish Commercial and Savings Banks. *Journal of Financial Services Research, vol.22, issue 3*, p. 203-224.

Shehzad, de Haan, Scholtens. (2010). The impact of bank ownership concentration on impaired loans and capital adequacy. *Journal of Banking & Finance*, 399–408.

Schirwitz. (2009). A Comprehensive German Business Cycle Chronology. *Empirical Economics, 37*, 287-301.

Shingjergji and Shingjergji. (2013). An Analysis of the Nonperforming Loans in the Albanian Banking System. *International Journal of Business and Commerce, Vol. 2*.

Shu. (2002). The Impact of Macroeconomic Environment on the Asset Quality of Hong Kong's Banking Sector. *Hong Kong Monetary Authority Research Memorandums*.

Škarica. (2014). Determinants of non-performing loans in Central and Eastern European countries. *Financial Theory and practice,* 37-59.

Škarika. (2014). Determinants of non-performing loans in Central and Eastern European countries. *Journal of Financial Theory and Practice, 38(1),* 37-59.

Stern, F. (2004). *Too Big to Fail: The Hazards of Bank Bailouts.* Brookings.

Stiroh. (2004). Do Community Banks Benefit from Diversification? *Journal of Financial Services Research, 25*, 135-160.

Suryanto. (2015). Non Performing Loans on Regional Development Bank in. *Mediterranean Journal of Social Sciences,* Vol 6 No 4.

Teräsvirta and Anderson. (1992). Characterizing Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models. *Journal of Applied Econometrics, Vol. 7,* S119-S136.

Tong. (1983). *Lecture Notes in Statistics. Threshold Models in Non-linear Time Series Analysis.* Springer-Verlag.

Wilson. (1998). Portfolio Credit Risk. *FRBNY Economic Policy Review,* 71-82.

Chapter 2[1]

Forecasting the Sectoral dynamics Macro Variables-Credit Quality: A Markov Switching Analysis

Abstract

As illustrated in the previous chapter, a stable banking system is a necessary base for a prosperous economy. The 2007 financial crisis and subsequent prolonged economic crisis, creating panic in financial markets, massive loan losses and difficulties in maintaining the financial stability, renewed attention in the safety of financial institutions.

The goal of this chapter is to reexamine empirically the successful hypotheses according to which the growth of credit vulnerability is related to macro conditions under the additional assumption that the impact of any future path of the economic activity is subject to one or more regime-switches that characterize asymmetries. To this purpose, we choose the alternative non-linear regime based Markov Switching method, asserting that may occur different regimes of credit quality depending on the level of the Real Gross Domestic Product, Unemployment, house price index and other significant variables.

In tackling this endeavor, the work uses a quarterly dataset lasting thirty years composed of macroeconomic-cyclical and financial variables chosen according to the results that emerged in previous empirical analyses and tested for stationary by both traditional unit root tests and tests with structural breaks. Structural changes, indeed, are quite pervasive in economic time series and ignoring them could be hazardous and could not allow justifying the suitability of our methodology.

The relationship between economic growth and credit quality is likely to be subject to changes due to economic crises, fluctuations in interest and exchange rates, reforms in banking regulation, or changes of economic policy. These changes, affecting the stability of parameters over a period of analysis, need to be considered when studying the direction of the relationship between economic growth and banking system performance. Adopting linear models whose functional forms assume constant relationship might not report robust results and may affect forecasting.

On the contary, beyond the impact of macro variables, the non-linear model will allow us to compare research results with past studies and to forecast credit quality ratio one period ahead, promoting a better insight into the key determinants of credit risk in the focal country. Since loan losses are the major indicator of bank failures, a better prediction could be advantageous for both banks and the whole economy.

Forecasting accuracy and robustness will be compared with the estimation outcomes of a simple linear regression model even to examine later what models, linear versus non-linear, best replicate business cycle features and to finally demonstrate whether the predicted GDP values and the other macro indicators can affect the number of credit vulnerabilities.

Keywords: loan delinquency rates, Markov-Switching Model, forecasts.

---

[1] The main results of this chapter were presented at the 2018 ADEIMF Summer Congress of Cagliari with a paper titled "Forecasting the Macro Determinants of Bank Credit Quality: A Non-Linear Perspective" coauthored with Antonio Forgione and Edoardo Otranto.

1. The dataset and the formulation of testable hypotheses

In recent years, there has been an increased interest in time series modeling and forecasting.
Linear regression models have mainly been used in the earlier literature due to their ability in forecasting as compare to other linear time series models.
However, unstable economic condition[2] and business cycle give asymmetric properties to the macroeconomic variables that cannot be captured by the simple linear models, so that the forecast values based on this approach may not be trustworthy.
The main objective of this study is to describe the nonlinear effects of many macroeconomic variables on credit risk.
There is an extensive amount of non-linear models. Considerations are restricted to the Markov switching (MS) that is essentially a linear process that switches between a number of predetermined regimes according to an unobserved Markov chain.
Using a switching regression setting, we first show that credit risk follows a two-regime Markov process, then we study both in sample and out-of-sample forecast performance of this model by comparing it with a simple linear regression in order to see whether non-linear regime switching model improves predictions.
According to us, the MS framework strengthens the assumption that the statistical properties of the data during stable periods do not remain the same as after a crisis episode or any structural change event.
In addition, as there is a limited amount of research concerning the application of such models on credit vulnerability, we believe that this analysis will serve as a work literature for other researchers who wish to embark on similar studies.
Consequently, we develop two testable hypothesis:

Q1. There is a nonlinear relationship between macroeconomic variables and loan losses.
In the light of the literature reviewed, we analyze and quantify the hypothesis that bank's credit risk is sensitive to predicted external macro environment using a nonlinear data generation process (DGP) to display asymmetric fluctuations or different patterns in the behavior.

Q2. If hypothesis 1 is verified, the non-linear model performs better in terms of forecasting as compare to the linear model.
This question is relevant because it could be that MS delivers poor forecasts once estimated due to imprecision in the estimations process despite the fact that MS is the process that more likely generates the data, and many economic and financial phenomena are non-linear. For the same reason, the linear model could produce better forecasts due to its precision in the estimation procedure and capability to approximate data quite well, even though it is not the DGP.

The empirical research is based on secondary data. It comprises both data on American credit quality and macroeconomic time series on a quarterly basis for a period spanning from March 1987 to March 2017 and consists of 121 observations for each time series.
Variables are drawn mainly from the Federal Reserve Board and the International Monetary Fund, Global Financial Stability Report Tables. Appendix A reports the statistical sources item by item[3].

---

[2] Bull and bear phases in financial markets, depreciation and appreciation of exchange rates, the evolution of unemployment and interest rates (unemployment increases more quickly than it decreases; interest rates are usually lower during recession due to intervention of monetary authorities and a lower demand for financing) are just few examples.

[3] It needs pointing out that the choices of American data as well as the observed period have been dictated by practical reasons. Regarding Europe, as instance, the research of high frequency data on NPLs and over such a long horizon was a rather ambitious demand. Central banks do not collect them except on annual basis, and they systematically monitored them since 2008. Moreover, the IMF publishes just quarterly short series. Special cases were Spain and Portugal provided with a dense statistical dissemination service in an interactive format.

To measure USA loans quality we use the quarterly delinquency rates seasonally adjusted provided by the Federal Reserve Board.
The chosen proxy, as known, monitors the financial health and soundness of the banking system in an economy since high levels restrict the ability of the banking system to provide further credit to support economic growth and provisions are needed to be set aside by banks in case these loans have to be written down.
In the USA, delinquent loans coincide with those past due of over 30 days and still accruing interest as well as those in nonaccrual status, which also covers lease contracts.
Bearing in mind that, we follow Louzis *et al.* (2012) and Ghosh (2017) and we perform our analysis through a sectoral approach. We conduct separate analysis for the primary loans groups of all USA charted commercial banks in order to underline the sensitivity of each loan categories to macro shock. Therefore, in this study, we have three dependent variables, one for each loan type Y_i $(i = 1, \dots, 3)$. Namely, real estate loan losses, consumer loan losses and commercial and industrial loan losses are examined, and the focus of this paper will be on explaining the contributing factors to different categories[4], revealing possible similarities and differences.
Delinquency rates are measured as a percentage of end of period loans, and for any loan category are calculated as the ratio of the dollar amount of a bank's delinquent loans in that category to the dollar amount of total loans outstanding in that category.
Figure 2-1 presents their development over the sample period[5] giving a useful picture of the complete loan performance.

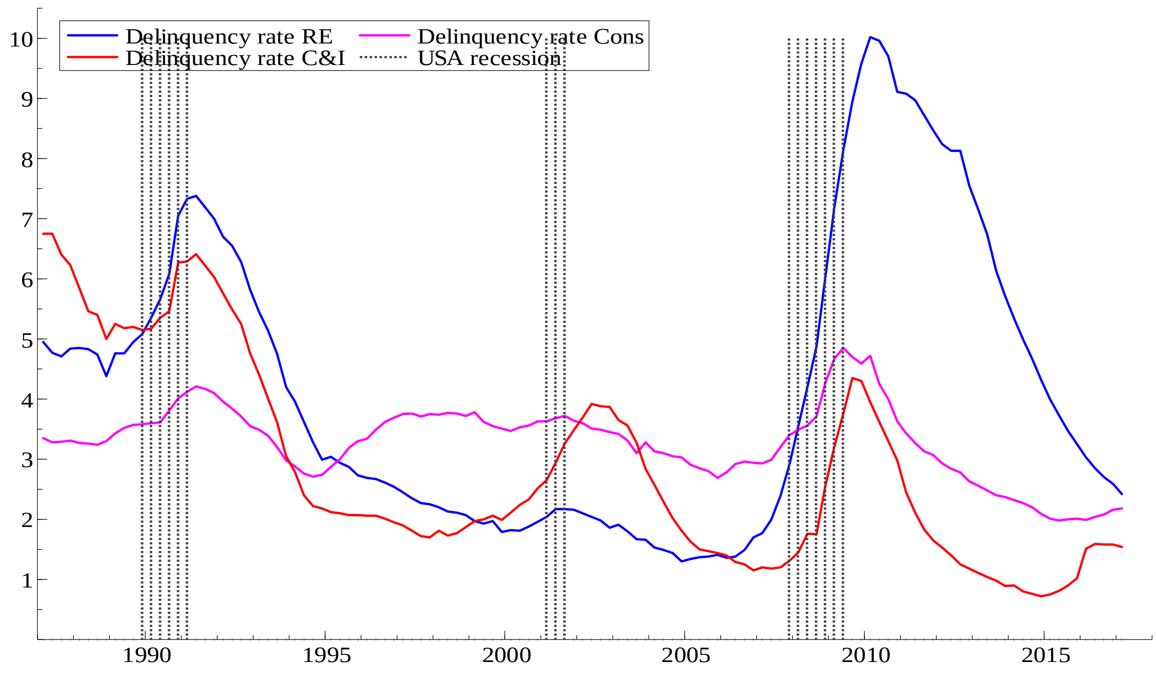


Figure 2-1 The evolution of loan quality in the USA. Source: Federal Reserve.

The graphical comparison shows that all delinquency ratios share a common qualitative evolution: during the first year of the 1990s, they augmented substantially reflecting the related mild recession. In the succeeding

[4] Real estate loans include loans secured by one to four family residential properties, including home equity lines of credit; construction and land development loans, loans secured by multifamily residences, loans secured by non-farm and nonresidential real estate. Consumer loans refer to secured and non-secured financing given to clients for family, personal or household scopes, or for consumable items. Differently, C&I loans are provided to business or corporation (not to an individual) either to finance working capital or physical assets expenditures.

[5] It is relatively large to include three economic recessions: 1989q4 to 1991q1, 2000q1 to 2000q3, and 2007q4 to 2009q2.

years, credit quality trended downward and remained relatively low until the 2001 economic recession that, however, was not as severe as the 2007 recent financial crisis. The latter had the major significant impact on loan quality. Loans experienced the highest credit loss rate. They weakened considerably and hovered globally in ranges between 2.5 and 7%, with the peak of real estate loans between 2 and 10%.

Today, the overall performance of the banking industry is much better with loan portfolios improved considerably since 2009.

In spite of such similar dynamics, each loan performance is different and the influence exerted by the macroeconomic variables on delinquency loan ratios changes depending on the specific segment to which the credit is transferred. For instance, while real estate appears as the primary source of problem loans in the 1990s, commercial and industrial lending was the driver of credit decline in 2002. Continuing, in 2007 and beyond the magnitude of commercial and industrial bad loans was considerably smaller and there was a more concentration of mortgage loans.

Evidently, divergences across types of loans mirror qualitative difference in the nature of the three economic slowdowns itself: the first one was caused or likely linked to the Saving and Loans crisis[6], the second was sustained by the burst of the Internet bubble, and the last one was exacerbated by the subprime mortgage crisis[7].

Obviously, despite ratio hide important variations across types of loans and they are disturbed by macroaggregates with different lags, we will adopt the same procedure for studying each of the four models.

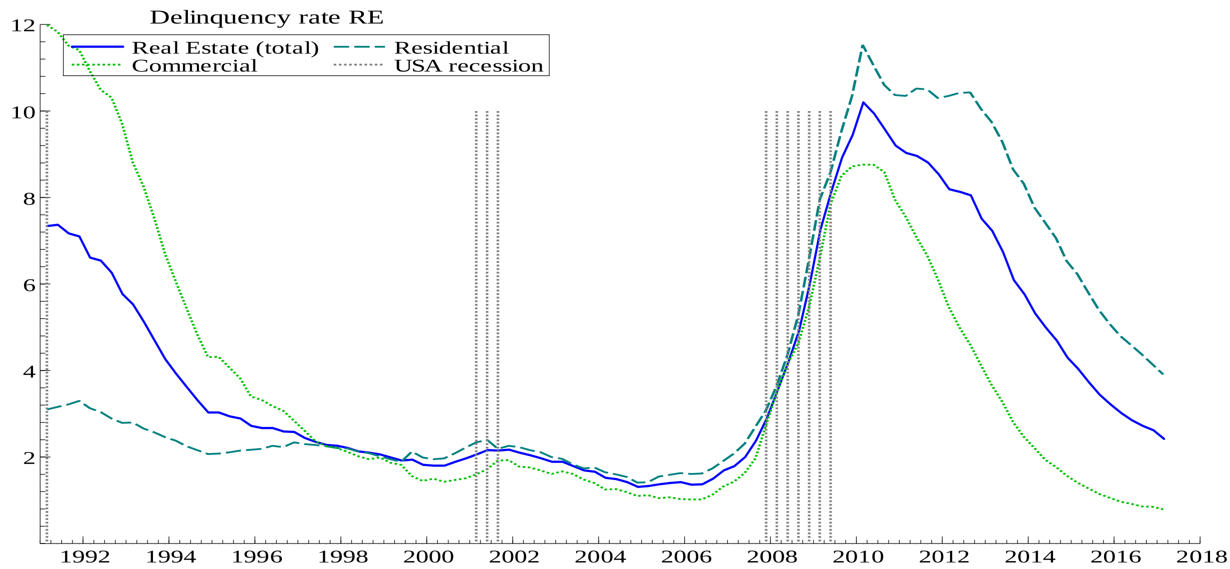


Figure 2-2 Delinquency rates on real estate loans - insured commercial banks in the United States from 1991 to 2016. Source: Federal Reserve.

[6] It was the greatest scandal and slump of financial institutions in US history. Over one thousands of thrift institutions (specialized in offering mainly housing and consumer loans) failed with a total cost around one hundred fifty billion dollars, requiring a bailout plan from the federal government. The maturity mismatching between borrowing and lending (they used short term saving to finance long fixed rate lending) and rising interest rates were the main causes of the crisis. Indeed, when inflation and interest rates at which borrowed rose between the late 1970s and early 1980s, S&L could not adjust the lending interest rate to reflect those paying on borrowed funds, becoming insolvent.

[7] "Irrational exuberance" (Alan Greenspan, 1996) in the high tech and house market caused these two latest recessions. As name suggests, the internet bubble or information technology bubble was based on speculative investments in internet companies. The investors excitement about the progress of internet led them to invest massive amount of money in any tech company with the belief of a new era and the hope of high returns, without caring wheatear or not the company was profitable or eventually could became profitable or not. From 1997 to 2000, prices went up about five times with a spillover from the technology sector towards sectors like finance, trading and services. However, many dot-coms were highly overvalued, they had often little or no profit and as a result, they crashed from 2000 to 2002, leaving investors with significant losses exacerbated by the terrorist attacks of September 2001.

A strong excitement and the overconfidence on asset value beyond its real economic value were among the reasons of the United States housing bubbles and related global crisis as well. Around the 2000, many people, exploiting low interest rate and lending standards, bought houses they could not afford thinking that house prices could only rises.

Figure 2-2 and 2-3 focus on each loan category subpartitions (available separate data for residential and commercial real estate, consumer credit cards and other). The shaded areas correspond to periods of recessions in the USA dated by the NBER.
The representations aim to provide a further breakdown of residential and commercial loans. However, the data availability is reduced to a subsample of our reference period.
It is not surprising that in figure 2-2, residential real estate loans underwent a rapid rise during the financial crisis driven by the mortgage credit boom and the subsequent falling in real estate prices. In 2010, the delinquency rate on residential real estate loans in the United States stood at 10 percent. Prior, their performance had relatively low credit risk. All subcategory clearly lags the business cycle.

Consumer loans, instead, apart the surge in correspondence to 2007-2009 appear as the least correlated with business cycle phases. Regarding the contribution of each size class, credit cards delinquency was and continue to be the main source of losses.

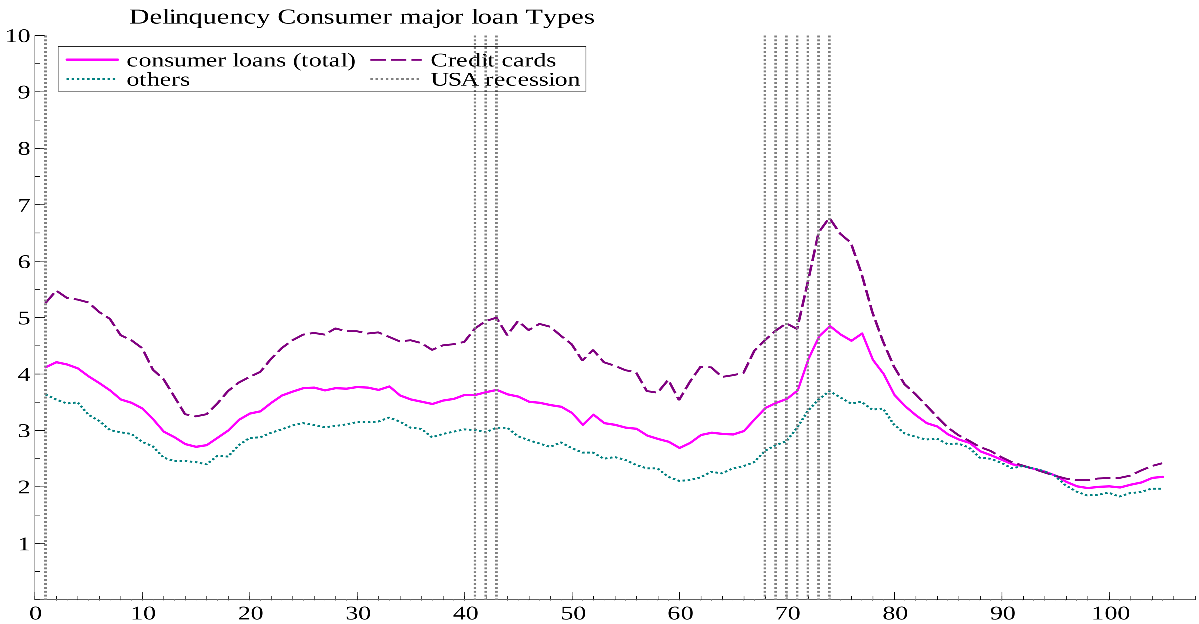


Figure 2-3 Decomposition of delinquency consumer. Source: Federal Reserve.

In what follows, we provide a brief overview of the macro dataset used in this paper. Section 3 outlines the tests for unit roots summarizing several tests and their specifications. Next, we specify in detail the set of variables, we introduce the methodologies and the consequent empirical results.
Finally, Section 4 concludes the paper with the forecasting evaluation analysis and conclusions sum up the entire works.

2. A look at the data: macroeconomic determinants

Various researches during the two past decades confirmed that macroeconomic conditions matters for credit quality and the several determinants had been an area of investigations.
In answering the research questions, our study found it appropriate to select eight recurring factors based on the abovementioned literature. In table 2-1, we have a summary of the variables used, their definitions and the expected signs demonstrating the connection between the two variable as provided by existing literature. The negative one implies that higher value of a chosen macro indicator tends to be associated with lower bad debts. The contrary holds for the positive one. The table also points some references to the existing studies using the same variables.

To portray the general state of the economy, we refer to the growth rate of Real GDP and unemployment rate. These variables are connected intimately to the capacity to satisfy debt obligations on the immediate hypothesis that a growing economy is associated with reduced financial distress. High levels of GDP growth, in fact, mirror higher inflows of income for both householders and firms.
The extensive set of exogenous variables also includes two indicators of price stability (the consumer price index (CPI) and the broad money supply M2[8]), two measures of changes in financial and real wealth (the house price index and the stock market index). It also comprises a measure of the burden of debt (the interest rate) and an indicator of competitiveness (the real effective exchange rate (REER)). We finally include the lagged value of the dependent variable itself to investigate the effect of previous delinquency rate on the current amount[9].
For homogeneity with the dependent variable, all variables are presented in quarterly frequencies and they are measured as percentage growth (percentage change since the same period in the previous quarter)[10].

Table 2-1 Variables definition and sign of the relation between delinquency rate and macro determinants. Source: author elaboration

Variables	Definition	Relation with delinquency loans	References
Real GDP	It is the core indicator of a country's standard of living and current health. It is the measure of the size and growth of an economy over time.	(-)	Salas and Saurina (2002), Ranjan and Dhal (2003), Fofack (2005), Quagliarello (2006), Espinoza and Prasad (2010), Bofondi and Ropele (2011), Louzis et alt. (2011), Castro (2012), etc.
Unemployment	It is a measure of the extent the economy operates at full capacity, exploiting the entire labor force[11]. Therefore, the unemployment rate is the percent able-bodied health adult able to work who are not doing it.	(+)	Gambera (2000), Nkusu (2011), Bofondi and Ropele (2011), Louzis at alt. (2012), Castro (2012), Skarika (2014), Gosh (2015), Flip (2015).[12]
Inflation (CPI)	It is the most widely used measure of inflation or price instability. The CPI represents the average changes over time in prices paid by urban households for consumption goods and services[13].	(-/+)	Arpa (2001), Shu (2002) Fofack (2005), Khemraj and Pasha (2009), Nkusu (2011), Klein (2013), skarika (2014), Gosh (2015), Flip (2015).

[8] As later, we do not observe a strong correlation between inflation and broad money supply in our analysis reason why we include both of them.
[9] Researchers found different results regarding the effect of the lagged NPLs on the current NPLs. Louzis et alt. (2012) found a negative and significant effect of the lagged NPLs, claiming that NPLs ratio is likely to decrease when it has increased in the previous quarter, due to the write-offs.
[10] Interest rate is in percentage per annum.
[11] The Bureau of Labor Statistic classifies people as unemployed if they meet the following criteria: they had no employment during the reference week; they were available for work at that time; and they actively looked to find employment sometime during the past four weeks. People who were temporarily laid off from a job and expecting recall are counted as unemployed.
[12] There are also study that found an insignificant relationship such as Quagliarello (2007).
[13] User fees (such as water and sewer service), sales and excise taxes paid by the consumer are also included. Income taxes and investment items (like stocks, bonds, and life insurance) are not included.

M2	It is an indicator of money supply also referred to as monetary aggregates. It measures the amount of money circulating in the economy and consists of narrow money (M1)[14] plus short-term deposits in the bank and 24 hours money market funds.	(-/+)	Kalirai and Schicher (2002), Fofack (2005), Babouček and Jančar (2005).
House price index	It is an accurate indicator of house price trend at a county level. It is used as a proxy to monitor individual investment behavior.	(-)	Arpa (2001), Hoggarth et alt (2005), Castro (2012), Gosh (2015), Hoggarth et alt (2005), Bofondi and Ropele (2011).
Stock market index	It is a measurement of the performance and activity of various financial markets within the economy. S&P 500[15] is taken as a proxy.	(-)	Bofondi and Ropele (20119, Castro (2012), Beck (2013).
Interest rate	It is the cost of borrowing money or better, the particular amount that a borrower is required to pay to a lender to receive a sum of money as financing. We use the T-bills 3 months short term as proxy series, commonly used by banks for pricing.	(+)	Shu (2002), Ranjan and Dhal (2003), Fofack (2005), Hoggarth et alt 82005), Quagliarello (2006), Bofondi and Ropele (2011), Castro (2012).
REER	Also known also as Trade-Weighted Exchange Rate Index, it is a measure the average prices of home manufactured goods relative to those of a country's major trading partners. It describes the strength of a currency relative to a basket of other currencies.	(+/-)	Fofack (2005), Hoggarth et alt. (2005), Quagliarello (2006), Khemraj and Pasha (2009, +), Castro (2012), Beck (2013), Klein (2013).

3. Preliminary econometric analysis

3.1. Unit root tests

3.1.1. Traditional tests: a brief presentation

Our econometric analysis starts with pre-testing all-time series for stationary performing the commonly known tests prominently in literature, namely the Augmented Dickey-Fuller (ADF), Phillips-Perron (PP) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests.

Indeed, before the execution of the switching regression analysis, the time series properties of the variables need to be examined. Testing for unit root is an imperative condition because time series analysis and forecasting assume stationary but it is widely known that most macro and financial variables exhibit mean and variance varying over time. Asset prices, exchange rates and all macroeconomic aggregates like real GDP are leading examples of variables that experience episodes in which their behavior change quite importantly compared to that exhibited previously.

[14] It is defined as a country money stock that consists of coins and notes in circulations plus other money equivalent that are readily available and convertible into money.

[15] It is the best gauge of the large cap U.S. equities market. The index includes 500 leading companies in leading industries of the U.S. economy, capturing 75% coverage of U.S. equities.

The theme of unit roots became prevalent in modern time series analysis given that incorporating nonstationary in a regression model gives serious econometrics consequences and misleading inferences (Glynn, 2007), such as spurious causality results[16] and influences on the properties of the regression estimators. Most important, the limiting distributions of unit root tests are non-standard and non-normal; they can be expressed as functional of Brownian motion and critical values have to be calculated using simulation methods like Monte Carlo.
Despite "nonsense regression"[17], the relevance for checking for the unit root hypothesis has at least two additional econometric advantages. First, it can help to detect the source of non-stationary and, in consequence, the different strategies to reverse it depending on the features of the process (Escudero, 2000).
Second, testing can also be important for forecasting. Indeed, the characteristics of the process determine whether the effects of shocks in Y_t h-periods ahead are transitory or permanent and imply very different predictions. Therefore, it will help us to decide what kind of process we will have to use in order to make accurate predictions (Diebold and Kilian, 2000).

It was Nelson and Plosser (1982)'s influential article that gave to the topic great momentum among economists. Applying Dickey-Fuller types tests to 14 annual U.S. of the most important aggregate economic variables including measures of output, employment, prices, wages, stock prices, and interest rates, authors failed to reject the unit roots hypothesis in all but one (unemployment) of the series and overturned the effective general view that macroeconomic data series were stationary around a deterministic trend. The only exception of the unemployment rate was recognized *a priori* to be stationary from the autocorrelation function representation inconsistent with the behavior of a random walk (this last one decays slowly with increasing lags).
However, authors stressed how the non-rejection of the null hypothesis does not necessarily imply that the null is true because none of the tests presented, formal and informal, can differentiate between unit roots and a stationary alternative with an AR root arbitrarily close to unity (power matter).
Moreover, authors found an important implication: the nature of non-stationary in macroeconomic time series "arises from the accumulation over time of stationary and invertible first differences" (page 160).
To say it differently, just when series are integrated, random shocks will have a permanent effect on the economy and fluctuations are highly persistent. Clearly, they referred to the discrepancy between trend and difference stationary process, also known as deterministic non-stationary and stochastic non-stationary to state that the difference-stationary process more adequately represents most macroeconomic series.
Indeed, trend-stationary series, characterized by stationary movements around a deterministic trend, can be transformed into a stationary process by extracting a time trend. They exhibit mean-reverting behavior meaning that any change has only a temporary effect that vanishes to zero as time passes so that the series reverts to its steady trend after the shock.
In contrast, if the series is difference stationary or has a unit root, and is therefore characterized by a random walk, shocks have an enduring effect. As a result, the series does not return to its former path and its level shifts permanently.
Authors concluded that in this case of non-stationary, economic fluctuations are related to movements in the secular component (that has to be modeled as a stochastic process and that results from long terms effects of real factors, such as changes in tastes and technology) rather than to movements in the cyclical component.

[16] Spurious regression is the most important implication of unit root. Granger and Newbold (1974) discovered it. It occurs when a pair of two independent random walks X_t and Y_t are regressed over the other and OLS results show an apparently relation that it is incorrect. Let be $Y_t = Y_{t-1} + \epsilon_{1t}$ and $X_t = X_{t-1} + \epsilon_{2t}$ with $E[\epsilon_{1t}\epsilon_{2t}] = 0$ for all t. Suppose we estimate the model $Y_t = \mu + \beta X_t + u_t$, since the two series are independent we should expect an estimated coefficient very close to zero and a small R^2. On the contrary, Granger & Newbold (1974) observed that the frequency of rejection of the true hypothesis $\beta = 0$ is much greater than the nominal 5% significance level, and $R - square$ is unusually high. Their outcomes, better, showed a good regression that actually is meaningless. Later, Philips (1986) formally showed that in the Granger-Newbold regressions all the statistical tools used to prove results, such as law of large number and central limit theorem are no longer valid and it inevitably leads usual t-ratio significance test to be bias towards the rejection of no relationship.
[17] Udny Yule, 1926.

“Since cyclical fluctuations are assumed to dissipate over time, any long-run or permanent movement (non-stationary) is necessarily attributed to the secular component” (pages 139-140).
In other words, authors gave evidence that cycle is present in the stationary first difference and unit root time series contain stochastic trend rather than deterministic.

In order to identify the stationary of the variables under study in this research, we choose to begin the analysis with the ADF as empirical studies indicate it as the initial and groundbreaking work on testing for a unit root in time series and it is the most popular of many competing unit roots tests.
Based on the underlying concepts of the original Dickey-Fuller test, the ADF is an expanded version to include lagged differences of the observed variable in order to capture the dependence of the error term within the model. One property of the basic Dickey-Fuller test, in fact, is that the disturbance term is not white noise and to account for this and to avoid an oversized test[18], the ADF test’s regression includes k lags of the first differences of y_t in the model:

$$\Delta\gamma_t = u + \beta t + \alpha y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-1} + \epsilon_t \tag{1}$$

where γ_t is the series being tested, t is the time trend variable, $\alpha = \theta - 1$; Δ denotes the first differences and k the number of lags added to the model to ensure that residuals are uncorrelated and whose optimal length is determined using the frequency of the data or the information criteria.
The basic object of the test is to examine the null of a random walk against the alternative of stationary by the estimation of the parameter θ. If the estimated θ is statistically equal to one in equation 1 or, alternatively, when $\alpha = 0$ in equation 1, then γ_t is non-stationary. In fact, if $\alpha = 0$, then $\theta = 1$, that is we have a unit root, meaning the time series under consideration is non-stationary.
Formally:

$$H_0 = \alpha = 0 \quad \theta = 1 \quad \sim I(1)$$

$$H_1 = \alpha < 0 \quad \theta < 1 \quad \sim I(0)$$

where $I(0)$ and $I(1)$ stay for a non-integrated stationary process and a series with a unit root that needs to be differentiated one time to be stationary.
The test is a normal one side (it discerns on the left) “t” test on the coefficient of the lagged dependent variable; hypotheses are assessed using the t-ratio $t_\alpha = \hat{\alpha}/_{SE(\widehat{\alpha})}$[19] and the critical values are the same as those given for the DF test[20]. If the ADF statistical value is smaller in absolute terms than the critical values, and the value of the (p-probability) is bigger than five percent, then the null hypothesis is accepted of a unit and data is concluded to be no stationary.
On the other hand, if the value of the DF statistical is bigger in absolute terms than the critical value and the value of the (p-probability) is less than 5% then the null hypothesis of a unit root is rejected and it is concluded that y_t is a stationary process.
Next, (Phillips and Perron, 1988) and KPSS (Kwiatkowski, Phillips, Schmidt and Shin, 1992) tests are performed as alternative and complementary unit root tests as well. This variety of approaches is justified by

[18] It consists in highly reject a correct null hypothesis.
[19] $\hat{\alpha}$ is the estimate of α and $SE(\widehat{\alpha})$ is the standard error.
[20] The basic DF test does not have a conventional t distribution and the critical values are the originally calculated by Dickey-Fuller derived from simulation experiments and tabulated by (MacKinnon, 2010).

the aforementioned tests size and power properties: the tendency to over-reject the null incorrectly when it is true and under-reject the null when it is false.
Specifically, a common limitation of ADF regards its sensitivity to the different lag length chosen and to the functional manner it is conducted, with severe size distortion. It may happen that the true distribution of the test statistic is different from the one we are using. For example, if the true model is a random walk but we are testing a random walk with drift, conclusions may be wrong since the true significant level is different from the estimated one, increasing the probability of making a type I error. Additionally, because the power of the test is inversely related to the size, ADF has notoriously low power; it does not reject the null of unit root more often than it is justified, concluding that the variable has a unit root even when none exists (Byrne and Perman).
In the light of this, the comparison between alternative unit root tests with useful variations as well as the implementation of the test from the opposite direction of the null of stationary against a unit root help to get around the problem of low power. Following the words of Kwiatkowski *et alt.* (1992), "it would be useful to perform tests of the null hypothesis of stationary as well as tests of the null hypothesis of a unit root because most economic time series are not very informative about whether or not there is a unit root". That is to say, they support that standard unit root tests are not very powerful against relevant alternatives.
PP and ADF tests, for instance, share the same limitation of power but differ in the treatment of autocorrelation and heteroskedasticity in the errors.
Under the same null hypothesis, the former corrects the issues non-parametrically by modifying the Dickey-Fuller test statistics rather than adding an autoregressive part such that the Philips Perron test statistics can be viewed as Dickey-Fuller statistics that have been made robust to serial correlation by using the Newey–West heteroskedasticity and autocorrelation consistent covariance matrix estimator (HAC).
It is based on the OLS estimation of the following regression equation:

$$y_t = u + \beta\left(t - {}^{T}\!/_{2}\right) + \theta y_{t-1} + v_t \tag{2}$$

where y_t is the particular time series under investigation, $t - {}^{T}\!/_{2}$ is the time trend, T is the sample size and $v_t$the error term.
Clearly, despite the disadvantage of working well only in large samples because based on an asymptotic theory, the advantage of not specifying a lag length for the test regression and the robustness to any form of heteroscedasticity in the error term makes the test more powerful than the ADF.
Besides, the PP test tends to accentuate the results relative to the ADF tests. In fact, the asymptotic distribution of the PP t statistic is the same as the ADF t and therefore the critical values are identical but the estimated statistical values are usually larger in absolute value; thus, for given critical values, one is more likely to reject the null.

However, testing for unit root is not as straightforward and direct as it can seem.
Several authors observed that conventional testing procedures have an inconvenience related to the occurrence of anomalous events such as a structural break or change in a data series that influence the tests results.
Plots of the data, as well as our knowledge of financial crisis events and policy changes, make highly plausible that significant structural breaks are present in time series.
The 1997 Asian financial crises and the 2007/2008 Global Financial Crisis are two common points of changes but also other dates are subjected to abundant shocks and regime shifts given that data span over a long period of time.
Because of the observed sudden changes, applying conventional unit root test alone is insufficient and problematic.
In general, ignoring the presence of structural breaks in economic and financial time series leads to a misspecification of the data generating process. We not only obtain an incorrect behavior and estimation of parameters of the variables but also unreliable forecasts.

In the last decade, lots of interest has been paid to the issue and a number of tests have been developed to take into account this feature, revisiting and contradicting Nelson and Plosser (1982) unit root results.
Along with the conventional tests, we will also use these different testing procedures to investigate whether or not results are reversed or robust when a structural break is included in the model.

3.1.2. Structural breaks definition and related unit root tests

The term structural break is not exactly defined. Although we have repeatedly used it, there is not a concrete definition in the literature.
By instinct, a structural break is a deviation in the parameters of an entity. In econometric theory, structural change is a "statement" about any or all parameters of the process that produces a time series (Hansen, 2001): it is the point in the sampled period at which there is a considerable unexpected change in the normal path of a time series.
Perron (1989) talked about a sudden change or external disturbances in the mean level of the series at a given time period; similarly Maddala and Kim (2003) and Gujarati (2007) argued about "changes of regression parameters" and unexpected shift in a time series that can lead to huge forecasting errors and unreliability of the model in general. All authors aimed to indicate that the value of the parameters of the model[21] do not remain the same through the entire period and create difficulties in determining whether a stochastic process is stationary or not.
Since the economy is quite flexible and dynamic to various influential factors, change points can arise from a number of possible causes starting from institutional, social and legislative changes up to economic, technical and climatic shocks.
Therefore, international and national events like globalization, social movements, abrupt policy changes in the direction of a State, wars, changing world economic conditions, oil crises, technology advances, natural disasters, and various other factors could be responsible for any parameter shifts observed in the time series data, altering the way the economy functions both positively and negatively.
"Many economic and financial time series undergo sudden, large breaks reflecting institutional changes, regime switches or breakdowns in market mechanism" (page 508); "breaks or jumps in the parameters could arise from factors such as major changes in market sentiments, bursts or creation of speculative bubbles, change in monetary and debt management (page 496) (Pesaran and Timmermann, 2004). According to the authors, ignoring breaks can be costly and determine loss in the forecasting accuracy.
A voluminous and growing literature has supported the applications and estimation of structural breaks.
It consists of two main kind contributions that according to Hansen (2001) have "dramatically altered the face of applied time series econometrics" (page 118).
The first class includes a considerable number of tests whose aim is detecting the timing and number of such changes. Not surprising, each of them reaches a different conclusion; results depend on the specification of the test. Initially, this class of test was restricted to cases of a single break in a regression equation model (Chow test[22]). Later, Bai Perron (2003) incorporated multiple shifts in a univariate regression[23].

---

[21] The reference is not purely limited to the mean specification but it extends into the series second moment as well.

[22] It is a test to determine whether a structural break exists. The procedure splits the sample into two sub-periods and estimates three regressions: one over the whole period and the other for the sub-samples separately. The F-test is used to see whether there are significant differences in the estimated equations. An important discrepancy, in fact, indicates the presence of a structural break in the data. The test runs in three stage: the first one consists in running a regression with all the data, before and after the structural break and in collecting the residual sum of squares, RSS_C. The second stage consists in running two separate regressions on the data before and after the structural break and collecting the RSS in both cases, giving RSS_1 and RSS_2. Later, using these three values, one has to calculate the test statistic from the formula: $F = \frac{RSS_c-(RSS_1+RSS_2)}{RSS_1+RSS_2} \times \frac{n-2k}{k}$, where k are the number of coefficients that are estimated for each regression. Finally, if the value of the test statistic is greater than the critical value from the distribution $F(k, n-2k)$, the null hypothesis that the parameters are stable over time is rejected.

[23] Bai Perron (2003) is a sequential test that generalize the Chow test for multiple unknown structural breaks. It is not a unit root test, but a commonly methodology approach with the goal of determining multiple break dates and number endogenously. It considers a

Differently, the second class of tests addresses the issue of structural breaks devoted to distinguishing between the hypothesis of random walk and that of broken trends. In our research, we will focus on this tests category, as the aim is not only break detection but also accounting for unit roots in the presence of shifts.
Perron (1989) developed the earliest one in a seminal paper where the effects of structural changes in the implementation and interpretation of unit root tests are underlined. Following the important findings of Nelson and Plosser, Perron argued that in presence of break the ADF produces misleading inference and inconsistent results because it is biased towards accepting the null hypothesis of unit root even when the series is stationary around the deterministic break component; the non-stationary might reflect neglected structural changes. Perron simply modified the standard ADF procedure to include dummy variables. The aim was to account for one ex-ante known exogenous break (determined by the tester, pre-tested) and to make the analysis as similar as possible to the previous one.
Using the same dataset of Nelson and Plosser and quarterly postwar real GNP series for the US economy, Perron (1989) rejected the unit null for eleven series that previously was found to be non-stationary and he argued that in the presence of an exogenous shock, most macroeconomic time series persistence arises only from large and infrequent shocks. After small and frequent shocks, the economy returns to a deterministic trend: "…fluctuations are indeed transitory. Only two events (shocks) have had a permanent effect on the various macroeconomic variables: the Great Crash of 1929 and the oil price shock of 1973" (page 1361). The quarterly GNP was also found to be stationary[24].
Nevertheless, since Perron supposed an ex-ante identification of the break date referring at economic information that depends on visual inspection of the data, it cannot be applied where such breaks are unknown, reason why Zivot and Andrews (1992) and the contemporary literature considered it inappropriate[25].
Criticisms sprouted from the arbitrary selection of break-dates, without any diagnostic test. In Perron's, the Great Crash and the oil-price shock, as instance, were chosen as obvious candidate among all the historical records; they were modelled as exogenous shocks in the sense that they are not realizations of the data generating processes of the various series but, they are assumed to occur at a priori known date.
Moreover, objections claimed the related inability to identify the exact occurrence objectively, with the likelihood to select a break date at the suboptimal point and stressed how in practice none researcher would choose the break date without pre-examining the data. "In practice, researchers use a combination of visual

---

multiple linear regression with l breaks: $y_t = x_t'\beta + z_t'\delta_j + u_t$ where y_t is the observed dependent variable, x_t is a vector of exogenous variables not subject to breaks, z_t is a vector of exogenous variables subject to breaks, β and δ_j $(j = 1, \ldots, l + 1)$ are the corresponding vectors of coefficients and u_t the disturbance term. The break points are explicitly treated as unknown. When β is not subject to shifts, the model is called a partial structural change model, if they are also allowed to shift or zeros; it is a pure structural change model in which all the coefficients vary.
For locating the breaks, the unknown regression coefficients together with the break points can be estimated using the Ordinary Least Square (OLS) method. Formally, for each l partition the least squares estimate of β and δ_j are obtained by minimizing the sum of squares residuals: $\sum_{i=1}^{l+1}\sum_{t=T_{l-1}+1}^{T_i}(y_t - x_t' - z_t'\,\delta_i)$^2.
Given the resulting estimates $\hat{\beta}(T_j)$, $\hat{\delta}(T_j)$ and substituting it into the objective function, the residual sum of squares over all partitions it is derived, $S_T(T_1, \ldots, T_m)$. The estimated break date are the global minimizer $(\hat{T}_1, \ldots, \hat{T}_m) = argminS_T(T_1, \ldots, T_m)$.
In addition, the regression parameter estimates are the associated least squares estimates at the estimated partition: $\hat{\beta} = \hat{\beta}(T_j)$; $\hat{\delta} = \hat{\delta}(T_j)$.

[24] Independently and using the Nelson-Plosser data, Rappoport and Reichlin (1989) reached the same conclusions. They attributed the changes in the trend rates of growth (the long-term properties of output) to rare events that occur infrequently such as substantial technological innovations, the Great Depression and the Second War World with the subsequent governmental economic policies, the oil shock of the early 1970.

[25] See also Christiano (1992). He used the bootstrap method to search for possible breaks points in the US GNP series and argued that the selection of the date of break independently of any prior information about the data is implausible and hardly credible. Author argued that that choice of the break point based on an ex post examination or knowledge of the data can lead to fallacious rejection of the unit root hypothesis and the p-values overstate the likelihood of the trend break alternative. Indeed, the critical values obtained in this way are much larger in absolute terms than those obtained when the break date is exogenously fixed. Therefore, the approach may have substantial size distortion as it invalidates the asymptotic distribution of the conventional testing and produce a not entirely negligible bias on the F-test.

examination of data plots, consultation with colleagues, and formal techniques to select a break date which is then tested for statistical significance" (Christiano, 1992).
In response to this criticism, a number of scholars proposed new methods.
Zivot and Andrews (1992) (ZA) developed a methodology for endogenizing the time break in the analysis of unit root, which it is to say to treat the break data as the outcome of the estimation procedure rather than fixed. This test is based on Perron specification of the form of structural break but slightly differs in the treatment of break under the null hypothesis and in the determination of the breakpoint.
In view of this, the test procedure considers a null as a unit root process that excludes exogenous structural change (clearly, in this test, a break under the null is not considered) against the alternative that the process is stationary with one break in the level, in the trend or both.
The test uses the full sample and is based upon a sequential estimation of the date of the hypothetical break by running a regression for every potential break change (every point is considered as a potential breakpoint), from the date after the starting until the last observation.
Specifically, the null hypothesis of unit root with a drift that excludes any structural break in ZA is $\alpha = 0$, while the alternative hypothesis $\alpha < 0$ implies a trend-stationary process with a one-time break occurring at an unknown point in time.
So, as pointed out by Lee and Strazicich (2004) "despite their popularity, these tests are invalid if structural breaks occur under the null; as rejection of the null would not necessarily imply rejection of a unit root per se, but would instead imply rejection of a unit root without break" (pages 1-2).
By endogenously determining the time of structural breaks, ZA demonstrated that bias in the usual unit root tests can be reduced but they found different break dates[26] and reported less evidence against the unit root hypothesis than Perron did, supporting Nelson and Plosser's original conjecture[27].
Nevertheless, the ZA (1992) test incorporates only one structural break in the data, the most significant structural break in each variable, even if more than one break is present. That is its obvious greatest inability and to overcome this issue, Clemente, Montañés, and Reyes (1998) (CMR) continue in this direction and extend the minimum ZA unit root test to include two structural breaks.
Several scholars, in fact, indicated the inadequacy of considering only one endogenous break, giving evidence of economic time series with more than one break. Notable studies in this regard include, among others, Ben-David, Lumsdaine, and Papell (2003). He warned that "just as failure to allow one break can cause non-rejection of the unit root null by the Augmented Dickey-Fuller test, failure to allow for two breaks if they exist, can cause non-rejection of the unit root null by the tests which only incorporate one break" (page 304). On the same issue, Lumsdaine and Papell (1997), Lee and Strazicich (2004) cautioned that considering only one break when in fact two are present can result in loss of power of the test and leads to loss of information.
Therefore, in order to submit our data to more rigorous analysis, we also extended the analysis for the case of more than one structural break and used the Clemente, Montañés, and Reyes (1998)'s testing method.
This unit-root approach is an extension of the Perron and Vogelsang (1992)'s approach[28] for non-trending data, to allow for two structural breaks in the mean.
Specifically, this test, similar to ZA, does not require an *a priori* knowledge of the structural break dates. Contemplating the two breaks distinctions or the two models of additive outliers (AO) and innovational outliers (IO), it considers a null hypothesis of unit root with structural breaks against the alternative hypothesis that

[26] Clearly, such difference emerged due to the selection of the time of the break as the outcome of an estimation procedure, rather than pre-determined exogenously.
[27] The difficulty in rejecting the unit null rises as Zivot Andrews's critical values differ with respect to Perron's in absolute value (they are more negative). And specifically, authors reverted the unit root null conclusion at five percent significance level for only three out of 13 variables (nominal and real GNP and IP) using the Nelson and Plosser data.
[28] Perron and Vogelsang (1992) proposed a class of test statistics, which allows two singular forms of structural break: the Additive Outlier (AO) model, where changes occur instantaneously, and the Innovational Outlier (IO) model, where structural breaks occur gradually over time.

series are stationary with breaks that is rejected if the calculated *t* statistic is greater in absolute values than the critical value[29].
In this thesis, we will use both AO and IO even if the IO model seems to be more suited for our variables as they all seem to exhibit gradual shifts rather than sudden structural changes.
For further detail about the test procedures, see Appendix B.

3.2. Empirical results

As mentioned, to ascertain the integration order of the variables, we begin our analysis by testing delinquency rates with conventional tests (ADF, PP, and KPSS).
Results are reported in Table 2-2 below where all values are the corresponding p-values.
For simplicity, we choose to shorten delinquency rates in delq, while the terms RE, Cons, and C&I indicate each loan category.

Series in growth rate			
	ADF	**PP**	**KPSS**
Delq RE	0.1639 0.1507 (c) 0.4119 (c&t)	0.2411 0.4270 (c) 0.7500 (c&t)	>0.10 (c) <0.01 (c&t)*,**;***
Delq Cons	0.4148 0.1133 (c) 0.1143 (c&t)	0.3860 0.3969 (c) 0.4555 (c&t)	0.029 (c)* 0.038 (c&t)**,***
Delq C&I	0.0694 0.0602 (c) 0.0164 (c&t)	0.0389 0.1695 (c) 0.3516 (c&t)	<0.01 (c) *,**,*** <0.01 (c&t)*,**,***

Table 2-2 P-values of several tests for unit roots. Source: author calculation.
Note: (c), (c&t) indicates that unit root tests were conducted with a constant and a constant plus trend respectively.
*, **, *** denote that we reject the null of stationary at the 1, 5, 10% confidence levels respectively. The selection of the appropriate number of lagged coefficients introduced into the model is based on the Bayesian information criterion (BIC) proposed by Gideon E. Schwarz (1978) to determine the optimal lag order.

As it is evident, all tests fail to reject the null hypothesis of a unit root in each time series at almost all significance levels, implying that all the delinquency rate variables considered in this study are non-stationary in level.
The overall ADF results show more supportive evidence of random walk regardless of whether including a time trend or without a time trend in the series. These findings suggest that any macroeconomic shocks to the credit quality variables have a permanent effect; it will have effects on the variable into the very long run and may affect accuracy in forecasting.
The use of the PP test makes little difference in our conclusions. Results of PP test (and KPSS) also display strong confirmation of unit root for the all variables under study for the sample period[30]. It is still true that the null of a unit root cannot be rejected.
Nevertheless, as we all know, a weakness of the ADF and PP unit root test is their potential bias towards flawed non-rejection of the non-stationarity hypothesis. As many time said, they may fail to reject the unit root hypothesis if the series has a structural break such that for the series that are found to be $I(1)$ there may be a

[29] Perron and Vogelsang (1992) provide the critical values, as they do not follow the standard 'Dickey-Fuller' distribution (Baum 2001).
[30] KPSS test indicates that the null hypothesis of stationary can be rejected mostly for the cases with constant and trend. This conflicting result can be attributed to power and size distortion issues discussed.

possibility that they are in fact stationary around the structural break but, they are classified erroneously as *I*(1).
Clearly, the confusion of structural breaks as evidence of non-stationarity leads to misspecifications and misinterpretations. We thus include further an endogenous structural break in ADF-type unit root tests to guarantees robustness in our conclusion.
When the ADF-type tests are expanded along the lines of ZA, it can be concluded that the unit root null is rejected more frequently.
The next table presents the value obtained by ZA's procedure for our data. Lag length, again, is obtained by Schwarz's Bayesian information criterion.

Table 2-3 Zivot-Andrews one break test. Source: author calculation.

Series in growth rate (original values)			
	Time break	Lag	t-statistics
delq RE	2007q2	2	-4.972 (c)**,***
	1997q3	2	-3,398(t)
	1997q3	2	-5,060(c&t)***
delq Cons	2012q1	2	-4,008(c)
	2009q1	2	-3.947 (t)
	2008q3	2	-4.823 (c&t)***
delq C&I	1992q4	2	-5.297 (c)**,***
	1995q4	2	-5.104 (t)*,**,***
	2000q1	2	-5.465 (c&t)**,***

Note: If the computed value of the Zivot-Andrews test is less than the critical value at a chosen significance level, reject the null of a unit root with drift and without structural break.

Results indicate that in the case of random walk with drift (except delinquency rate consumer) and random walk with drift and trend we can reject the null of unit root for all variable at original level, clearly contradicting outcomes from tests without structural breaks.
Simultaneously, the test locates endogenously the point of most probable/significant structural break (this is reported for every time series examined in table 2-3 too).
In like manner, we find strong rejections of unit root with two endogenous breaks. Indeed, despite breaks, unit roots cannot be rejected with Montañés, unless in the cases of IO estimation method, which provides further evidence. In that case, all series have a minimized t-statistics (-5.645,-5.548 and -6.711 respectively) less than the t-critical value, meaning that we can reject the null hypothesis and conclude that the series is stationary when considering two changes in the mean of the series.
We reject the null hypothesis of unit root at the 5-percent level for delq cons and delq C&I. We reject the null hypothesis of a unit root for delq RE at the 1-percent level. Results demonstrate the effect of allowing two breaks instead of one but there is no potential bias from omitting the second break variable.
Therefore, great evidence is found in supporting trend-stationary processes for all delinquency rates.
However, to validate our conclusions even more, we refer to Baum (2005). Indeed, according to the authors, the results derived from ADF and PP tests are questionable and the models excluding structural breaks are misspecified if the estimates of the CMR unit root tests testify significant additive or innovational outliers in the time series, as in our cases.
If these tests in the presence of one structural break show no evidence of a structural break, the ADF and PP tests can be accepted.
The next table shows the value obtained by implementing Clemente, Montañés, and Reyes (1998)'s.

Table 2-4 Clemente-Montañés-Reyes unit-root test with double mean shifts. Source: author calculation

				Clemente-Montañés-Reyes (CMR)
Series in growth rate	Additive Outlier Model		Innovation Outlier Model	
	t-stat	optimal break	t-stat	optimal break
delq RE	-5.004	1995q1*, 2010q3*	**-5.548****	1990q3*, 2006q4*
delq Cons	-4.173	2009q3*, 2011q4*	**-5.645****	2008q1*, 2009q4*
delq C&I	-4.455	1994q3*, 2012q4*	**-6.711****	1990q3*, 2009q2*

Table 2 Critical value is -5.490 at 5% significance level, -5,96 at 1% and -5.24 at 10%.
***,** in the t-statistics values indicates rejection of the null hypothesis at 1 and 5% levels.
* in the optimal break values indicates that break is statistically significant at 5% level.

Evidently, the structural break dates are different concerning variables and in terms of ZA and CMR tests. Some break dates are closely connected with global events as we presumed: 1997 is the year of developing country debt crisis[31], the 2007 and beyond coincide with the financial crisis and its responses; the second quarter of 1990 is the period of USA economic recession.

However, not every break is associated with US recessions and the dates of the breaks differ across loans classes suggesting that events exert a heterogeneous effect.

Given such breakpoints and the plausible existence of others, changes in delinquency rates can be modeled by different approaches.

We choose the regime-switching model rather than insert breakpoints as exogenous variables in a linear model generally adopted in literature.

We think that linear models on first differences of the delinquency series and without changes in unknown breaks could be based on the wrong hypothesis.

Next sections are devoted to the empirical investigation. Typically, two finite number of regimes are defined, one representing periods of high vulnerability, the other period of less stress.

The so-called Markov switching models define a probability transition matrix, which governs the shifts between regimes. They can capture for example business cycle effects.

3.2.1. Summary descriptive statistics and correlation matrix

This section presents the descriptive statistics of dependent and independent variables used in this study. The following tables depict each variable used in this study.

Table 2-5, giving an overall about the data used in the model, reports mean, standard deviation and minimum and maximum values for the dependent variables.

[31] The Asian financial crisis, also called the “Asian Contagion”, was a period of currency devaluations and market declines that spread though the “tiger economies” of SE Asian. Starting with a severe Thai bath depreciation followed by the other currencies in the region with a domino effect (the Indonesian rupiah and the Korean won), it was caused by the unprecedented economic growth of Asian countries (i.e. Korea, Thailand etc.…) experienced before 1997. The turmoil in the stock markets and the collapse of Asian economies were also felt in USA, Europe and Russia. Specifically, USA were among the chief investors in the region with investments in financial instruments, bank loans and US subsidiaries.

Despite of the important differences among individual countries, exports and foreign investments were the common generators of the Asian miracle.

In fact, a monetary policy of fixed exchange rate tied with the USA dollar, eliminating the exchange rate risk, led Southeast Asia to enjoy an increased inflow of foreign capital and hence debt that helped investments in commercial and residential property. When USA began to rise interest rate to control inflation, its more attractiveness relative to Asia for investment raised the value of dollars and caused a decline in Asian exports and a deterioration of the economic status. Real estate prices started to fall due to the overvaluation of assets, foreign lenders withdrew credit and local currencies depreciated seriously.

Variable	Mean	Median	S.D.	Min	Max
Delq RE	4.18	3.47	2.45	1.3	10
Delq Cons	3.29	3.35	0.629	1.98	4.85
Delq C&I	2.85	2.11	1.69	0.72	6.75

Table 2-5 Key descriptive statistics for the dependent variables involved in the regression model.

Data confirm our graphical intuition.

Delinquency real estate shows the highest values reaching a maximum of 10% and a mean of 4.18%. Consumer delinquency ratio ranges from a minimum value of 1.98 to a maximum of 4.85 percent having a mean value of 3.29%. In addition, although the average rate of the business part is relatively low, the banking sector in the USA is still characterized by a high level of vulnerability.

The macro statistics in table 2-6 reveal that the stock market index has a high standard deviation with a large difference between the maximum and minimum value, suggesting that the market is highly volatile. The unemployment rate is 6.00%.

Variable	Mean	Median	S.D.	Min	Max
GDP	2.55	2.7	2.41	-8.2	7.8
CPI	2.65	2.7	1.34	-1.62	6.22
Unemployment	6	5.6	1.5	3.67	10.4
REER	-0.002	0.26	2.28	-4.88	8.34
M2	1.32	1.32	0.784	-0.36	4.52
House price index	0.861	0.99	1.17	-3.13	3.83
Tbills3 months	3.2	3.4	2.52	0.01	8.67
Standard Poors500	2.11	2.36	6.71	-25.5	15.6

Table 2-6 Macroeconomic data statistics. Source: author calculation.

Finally, table 2-7 gives correlations amongst all variables.

We notice how real estate delinquency rates are highly correlated with the unemployment amount meaning that unemployment increment positively weights on.

Correlation coefficients between the different explanatory variables are low except for the relationship between the inflation rate and interest rate (0.7). It is theoretically expected and justified because lowering interest rates allow people to borrow more money with the result that they have to spend, causing the economy to grow and inflation to increase. On the contrary, higher interest rate reduces consumer spending and the incentive to borrow, increasing the motivation to save. With less usable income, the economy gets slower and inflation decreases. Theoretically justified should be a correlation between inflation and broad money supply but we do not observe it, the reason why we include both variables in the analysis.

Correlation	DelqRE	DelqCons	DelqC&I	GDP	CPI	Unempl	REER	M2	Houseprice	Tbills	S&P500
DelqRE	1										
DelqCons	0.3	1									
DelqC&I	0.3	0.6	1								
GDP	-0.2	-0.1	0.02	1							
CPI	0.0	0.2	0.4	0.0	1						
Unempl	0.9	0.3	0.2	-0.2	-0.2	1					
REER	-0.1	-0.1	-0.2	-0.1	-0.1	-0.1	1				
M2	-0.1	-0.1	-0.3	-0.3	-0.1	-0.2	0.2	1			
Houseprice	-0.6	-0.4	-0.1	0.3	0.0	-0.5	0.1	0.1	1		
Tbills	-0.3	0.3	0.5	0.2	0.7	-0.5	0.0	-0.2	0.2	1	
S&P500	0.0	0.0	0.0	0.3	-0.1	0.1	-0.2	-0.3	0.0	0.1	1

Table 2-7 The Pearson correlation matrix. Source: author calculation.

4. Methodology

4.1. Regression analysis and results

Against this data analysis backdrop, as a starting point of our empirical investigation, we first test the significance of our regressors through a simple multiple linear regression model.
With forecasting ambition, we start by regressing delinquency rates on the entire set of variables lagged a quarter to select a significant set of regressors for each dependent variable, using the restricted observations from 1987 q1 to 2015 q1 for the in sample analysis, and all remaining information for the out-of-sample analysis.
Eliminating the non-significant variables, we then obtain parsimonious models with the consequence that models differ in term of the number and kind of regressors considered[32]. Precisely, GDP, M2, House price index and lagged delinquency rate are always present whereas the remaining variables are present only in some models.
According to the data analysis of the least squares estimation, results are presented in table 2-8 below, with the main diagnostic tests needed to decide whether models fulfilled the assumptions of a multiple linear regression model, test for heteroscedasticity, test for multicollinearity, and an autocorrelation test.
Not all residuals series pass both tests. An exception is the series of delinquency consumer that, on the contrary, does not exhibit autocorrelation. Hence, we calculate robust standard errors.
The highest coefficients of determination between 0,96-0,99 % reveal that the variation in delinquency ratio is broadly explained by the independent variables; they suggest that all variables are useful in explaining loan losses.
Overall, single loan estimation supports the influence of the macroeconomic environment but differs for the sensitivity to peculiar variables.
Regression analyses confirm Louzis *et al.* (2011) and Ghosh (2017) hypothesis according to which macro variables have a dissimilar repercussion on different loan types. However, besides confirming their results, which ascribed the vulnerability of the Greek sector mainly to GDP and unemployment, we found that the USA credit quality depends significantly on both financial and real wealth factors as well.
Indeed, all amounts of delinquency rates depend on not only the GDP but also on the house price index and the amount of delinquency of the previous quarter.
The quantitative difference among classes is proved with real estate delinquency loans more susceptible towards macro environment changes (it is evident a greater impact of all variables), followed by business and consumer loans.

[32] In the case of commercial and industrial loans, it has been necessary to exclude gradually inflation and interest rate variables.

Table 2-8 Reduced Ols estimation results, omitting some explanatory variables. SE in parentheses. The symbol * means that the t-statistics has a p-value less than 0,01,** between 0,01 and 0,05, * more than 0,05.**

Explanatory variable	***delqRE***	***Delq Cons***	***Delq C&I***
GDP_1	-0.041 (0.011)***	-0.014 (0.004)***	-0.038 (0.008)***
Unemployment_1		-0.039 (0.010)***	-0.083 (0.013)***
M2_1	0.062 (0.032)*	0.031 (0.014)**	0.060 (0.026)**
House priceindex_1	-0.173 (0.024)***	-0.045 (0.010)***	-0.058 (0.018)***
Tbills_1	0.039 (0.010)***	0.017 (0.006)***	
S&P_1	-0.006 (0.004)*		
Delinquency variable_1	0.950 (0.012)***	0.964 (0.021)***	0.997 (0.011)***
Intercept	0.250 (0.103)**	0.320 (0.096)***	0.531 (0.111)***
σ^2	0.057	0.010	0.034
R^2 **Adjusted**	0.991	0.967	0.988
Breusch-Godfrey	5.20E-10	0.140	9.99E-04
Breusch-Pagan	8.85E-06	1.83E-08	6.99E-08
VIF	1.454 1.307 1.564 1.220 1.196 1.634	1.365 2.388 1.408 1.516 2.125 1.364	1.362 1.405 1.379 1.435 1.155

Generally, in all estimations real GDP effect is omnipresent. The one lag is highly significant, at 1 percent level of significance supporting previous studies, that is that if the economy rate decelerates this will influence the delinquency ratios in the next quarter.

Moreover, all type of loans, except real estate exhibit large sensitivity to change in the unemployment rate.

In contrast to the majority of most previous empirical works, it has been recorded a negative sign. For consumer loans, the negative sign may mean that when people experience short-term unexpected financial adversity such as unemployment, they continue to pay back the credit cards, auto loans and other types of personal debts, becoming delinquent only when they suffer sizeable losses in income through long duration increases in unemployment. In other words, unemployment may alter individuals' capability to repay as unemployment duration increases, or better, it may happens that repayments do not decline as much as income.

The linkage could also imply that banks finance high income or capable workers less likely to get unemployed, the reason why increased unemployment does not damage loan repayment and propensity to default.

For the commercial and industrial model, the negative sign could be related to the short-term effect of the unemployment on firm profitability and cash flow. In detail, a rise in the unemployment rate should reduce, in the short term, the labor costs. Indeed, workers' bargaining power bend in favor of firms because workers are less reluctant to reject lower wage than the prevailing rate; if they do, someone else will do not it. As a result, wage rates fall very slow allowing firms to repay debts easier.

After controlling GDP and the general state of the economy, the wide fluctuation of house price with the recent history of price deceleration and related lack of confidence in the financial system result as the major significant channel for explaining changes in delinquencies in the United States.

Intuitively, price declines imply a contraction in the value of the assurances given and an increase in the risks for the banks financing the loans; banks experience a worsening in the value of their assets and balance sheets. In contrast, the increase in house prices improves the quality of loans.

Indeed, housing is the most important sector in the USA economy and the largest component of household’s wealth (13, 3% of total GDP in 2017 q1). Consequently, any shocks in the house markets have a spillover effect on the rest of the economy and household’s probability of defaulting loans. A fall in house prices may make housing a less attractive investment and rise the individuals in negative equity[33]. In such scenario, a household may find not easy to sell his house and extinguishes the loan and if he faces unfortunate circumstances like fewer jobs, evidently, default cost is better than maintain the mortgage; it is less onerous.

In another word, in case of negative equity borrowers have an interest in selling houses in exchange for the remaining debt.

A feature of the US real estate market is that mortgages are non-recourse that is to say that banks can manage the house in case of insolvency but they cannot rely on other debtor’s assets. Therefore, the borrower is the holder of an American put option that allows him to sell the house at any moment.

Moreover, the other loan categories endure the negative fluctuations in house prices as well, given that housing became less collateralizable as an asset. Individuals, indeed, perceiving a reducing wealth (wealth effect) and tense constraints, revise consumer spending and saving. It is instructive, what happened in 2008: an overall fall in GDP followed the sharp downturn in house price.

Changes in property prices also influence a firm’s ability to borrow and finance large-scale business activities. In the case of firms, the ability to borrow depends essentially on the market value of their net worth. Indeed, in certain situations, commercial property assets can be used as collateral and any economic disturbance changes its values such that whenever it goes down, related loans are judged more likely to default [34].

“The net worth of borrowers changes not only in response to variations in cash flow but also (and often, more dramatically) to changes in the valuation of the real and financial assets that they hold” (page. 1377-Bernanke *et alt.* 2015).

Proceeding with the variables, the rate of inflation, *per se*, is found to have a less effect and has not been considered in all models. The same for REER.

It was slightly surprising the insignificance of the REER coefficient.

Exchange rate, indeed, is expected to cause an increase of bad debts in countries with a high degree of openness to international trade, and the U.S. is the second leading exporter of goods and services in the world and the number one importer.

An exchange rate appreciation, for instance, worsens the performance of export-oriented firms, contributing to a deterioration of bank portfolios. It produces a rise in the international cost of exports, which reduces both export demand and the domestic cost of imports, leading to a substitution away from domestic production. Thus, it is plausible that its delinquency ratio reacts greatly to exchange rate volatility.

[33] The case where the value of the collateral of a mortgage loans decreases, making difficult debt renegotiation or better, is the situation where the market value of a house is below the outstanding loan secured on it.

[34] This refer to the key idea of the financial accelerator effect of Bernanke and Gertler (1989). It consists on the effect that shocks in economic acitvity have on the net worth of economic agents, resulting in their borrowing capacity and ability to repay loans. Namely, the asymmetric informations between lenders and borrowers necessitates that a borrower pledges a solid collateral as a security againts the financing of a loan at the time the loan is granted; lenders require collateral assets to disclose borrower ability to pay. All arise because borrowers have informational advantages on their reliability and the values of their investments, and to avoid consequences originated from moral hazard and adverse selection.

However, the reason behind the existence of credit vulnerabilities associated with exchange rates goes beyond trade and is associated with foreign currency borrowing. In this latter case, the appreciation of the local currency compensates the trade channel as it improves the debt servicing costs in local currency terms for borrowers.

Aware of these relationships, the fact that the US has not a large share of debt denominated in foreign currency may explain the reason why exchange rate had not affected NPLs significantly.

Moving forward, the growth of delinquency rates depends on interest rate and money supply. Both of them display an expected positive sign, obviously with a quarter lag. The lagged behavior of the money supply is explained by the fact that it is an indicator of monetary policy meaning the set of tools through which Central Banks create money, determine its size and growth. More money in circulation means that more income is held and available by individuals and firms, but also investment, increased sales and consumption. In other words, inflation that motivates the not reimbursement of loans from a perspective of high-interest rates[35].

Regarding, the relationship delinquency-interest rate, it is immediate.

High-interest rates cause increases in the costs of loans charged on the borrowers, the price of borrowing, which in turns encourages debtors to save instead of spending more money and causes them not to succeed to meet their obligations. The higher the interest rates, the lower money in an economic agent's pocket.

The same reasoning applies for the case of deflation with the difference that expansionary policy actions and cutting of interest rate as Central Bank's main tool result in a reducing cost of mortgage payments and, hence, delinquency loans.

Once more, depreciation of stock price S&P 500 explains only real estate loans fluctuations that respond faintly. It works similarly to change in the housing market or property ownership with the difference that stocks represent a liquid wealth. In a good market, insolvency diminishes because the stock index is an indicator of the value collateral and, in addition, people became more and more brave, capable, and enthusiastic to meet the loan repayment more easily, they see their portfolio appreciates and confidence rises.

The effect of financial wealth is weaker and less significant (0.094%), confirming the greater extent of American households housing assets. Consequently, in the USA the size of such wealth is more important over stock market assets.

Finally, yet importantly, a common feature is the dependence on the previous amount of delinquency loans. Such persistence over time is consistent with the broad literature. Louzis et alt. (2011) explained the negative sign found in Greece with the practice of write off, as we underlined. Evidently, in the USA it does not occur implying a high recursive value of delinquency loans. They are not immediately devaluated and remain on balance sheets at least up to the successive quarter.

4.2. Markov Switching approach: switching regression models and identification

Results of linear model raise the question whether the sensitivity of delinquent loans to changes in macroeconomic variables remains constant among the various phases of the economy and whether all variables continue to have the same significant effect.

Moreover, the problems of autocorrelation and heteroscedasticity faced up in the linear models make us question whether a variance-switching model can catch the absence of homoscedasticity.

[35] Relationship among inflation and money supply is the notable concept shown by Fisher (1911) in the "Purchasing Power of Money" who textually wrote "the normal effects of an increase in the quantity of money is an exactly proportional increase in the general level of prices" (page 821). Such statement in monetary economics is known as quantity theory of money and in its simplest form is expressed by:

$$M \times v = P \times Y$$

where M is the money supply, v is the velocity of circulation in the economy, P is the average price level and Y the volume of transaction of good and services. The theory assumes v and Y constant in the short term so that changes in the volume of money supply determine a change in the same direction of the general price level. When the money supply increases it should be inflation and deflation vice versa.

Thus, to control inflation, monetary authorities respond by acting a restrictive policy addressed to absorb the money supply.

In this section, we extend regression analysis to include non-linearities in macroeconomic relationships arising from a discrete change in regime.
In fact, ignoring the effect of any change means imperfect modeling that suffers from data fit. Hence, for parameter estimation, we implement a Markov switching regression (MS).
At first, we consider three couples of structural models, the first one has the constant subject to regime change, the second one has constant and regressors subject to regime change; in the last one, we assume regime specific error variances.
Therefore, using the likelihood ratio test and the Akaike information criterion (AIC)**,** Schwarz criterion and Hannan–Quinn information criterion (HQC) we choose the best model for all loans category among the candidates and we compare the in sample fitting to test which model provides a better explanation.
Again, we use all the observations until 2015 q1 included for the in sample analysis and all remaining observations for the out of sample analysis.
The general regime-switching model we consider is a linear regression where the dependent variable y_t is generated by two distinct regimes:

$$y_t = \delta_{s_t} + \alpha_{s_t} y_{t-1} + x'_{t-1}\beta_{s_t} + \epsilon_t \qquad \text{with} \qquad \varepsilon_t \sim N(0, \sigma^2) \tag{3}$$

y_{t-1} the autoregressive term and x'_{t-1} a $n \times 1$ vector of exogenous explanatory variables measured at time $t-1$, δ_{s_t}, α_{s_t}, β_{s_t} are unknown coefficients which can change according to the state s_t and ϵ_t is a zero mean normal disturbance with variance constant or depending on the state.
Indeed, to deal with regime specific variance heterogeneity, the model has also been extended under the assumption $\varepsilon_t \sim N(0, {\sigma^2}_{s_t})$ so to have switching intercepts, autoregressive parameter, variances and regressors.
Analogously to the linear regression model, the estimated betas are a measure of the sensitivity of credit quality to changes in the GDP, unemployment, inflation, money supply, exchange and interest rate, the house of price, and stock index. Again, these explanatory variables are lagged and not contemporaneous due to our forecasting interests. We assume the lag time to be one quarter.
Finally, the model is a two-regime MS model and the latent state variabless_t could take the value of one or two depending on the state or regime of the delinquency rate. In other words, the effect of each of the explanatory economic variables depends on whether delinquency rate amount is high or low.
The model is estimated using a maximum likelihood procedure.

4.2.1. The likelihood ratio test and information criteria

An essential challenging part of the applied statistical analysis and inference consists of selecting the appropriate model or the one that is best[36] supported by the data among the candidates set.
The likelihood ratio test that ranks competing models that differ in the number of parameters opposing the suboptimal is a broadly established statistically procedure typically used to face with this problem; in fact, the choice of one model over the other can make a substantial difference, and a subjective or somewhat arbitrary approach would have left much to be desired.
Obviously, it is impossible to have a model that represents exactly our information[37] but scrutinizing their precision, completeness, and accuracy is reasonable and valuable.

---

[36] Formally, the "best" model is the one with least information lost relative to other models in the set or, the one minimizing the Kullback-Leibler information. Let f denotes the full reality or truth and g an approximation model, K-L information $I(f, g)$ is the information lost when model g is used to approximate$f$; it is the distance between full reality and model (Burnham and Anderson , 2004).
[37] Keep in mind the famous mathematician Mark Kac who described models as caricatures. He observed, "models are, for the most part, caricatures of reality, but if they are good, then like caricatures, they portray though perhaps in distorted manner, some of the features of the real world".

Box (1979) originally expressed this sentiment in his famous statement "models, of course, are never true, but fortunately, it is only necessary that they be useful. For this, it is usually needful only that they are not grossly wrong" (page 2) meaning that models cannot reflect full reality but approximate it.
The likelihood ratio test is an easy procedure to compare two alternative nested models[38]. On an intuitive level, the test gives an impartial criterion to measure the goodness of fit of two models, a full or unrestricted one that has all the parameters and a reduced model that specifies some sort of restriction on the parameters. It is based on the comparison of the likelihood scores.
Let $\hat{\theta}_R$ be the maximum likelihood estimator of θ obtained imposing restrictions, $\hat{\theta}_U$ the unconstrained maximum likelihood estimator, l_R and l_U the likelihood functions evaluated at these two estimates.
The likelihood ratio test is the twice difference between the log-likelihood functions. It follows a χ^2 distribution with the number of degree of freedom equal to the number of restrictions imposed or equivalently, the number of additional parameters included in the complex model:

$$LR = -2(l_R - l_U)$$

or

$$LR = 2(l_U - l_R).$$

Such difference must be positive as additional parameter always result in a higher score and $l_U \geq l_R$.
The null hypothesis states the goodness of the smaller model so that the impact of the explanatory variables omitted in the regression is not significant.
So, the null is accepted when the test statistic is smaller than the critical value of the chi-squared. On the contrary, if the statistics exceed the critical value, the hypothesis is rejected and the restrictions are likely to be incorrect; the full model is an improvement over the small one meaning that model fit outbalances the cost of added model complexity.

Below in table 2-9, there are the outputs of our pairwise comparisons.
In order to establish a unique notation for each model throughout the work, we specify the regime dependent terms with the abbreviations MS-I (models with only the switching constant), MS-C (models with switching coefficients) and MS-CV (models with switching coefficients and variance switch).

Table 2-9 Likelihood ratio test results. Source: author calculation

Statistic1	Statistics2	MS-I	MS-C	MS-CV	d.o.f.	critical value
LR2=51.958	LR2=61.201	35.363	61.342	65.963	6	12.592
LR3=14.072	LR3=6.903	128.320	131.356	131.771	6	12.592
LR4=13.363	LR4=26.331	50.905	57.586	64.070	5	11.071

Over all models, test specification shows the unrestricted model as the most appropriate to capture the features of the analyzed data.
This applies especially when the MS-C is compared against the restricted MS-I; whereas comparing the case of MS-CV with the MS-C, the former is always favored except for the consumer series, where the test prefers the limited model.
However, we can generally conclude that the fit is significantly worse under the null; restrictions implied by the shifting intercept can be rejected while the additional parameters are meaningful.

[38] Models are said to be nested if one of them is a special case of the other.

Nevertheless, the likelihood ratio test suffers some limitations; it is restricted to nested models and allow us to consider and compare just two types of them.
For further evidence of our conclusions, and to make inference from several models simultaneously, along with hypothesis testing and as its supplement, we consider the alternative set of methods based on complexity-penalized likelihood criteria as well; they can be used to compare two models which have different specifications and which are not necessarily nested.
Moreover, as emphasized by Psaradakis and Soagnolo (2003) citing Granger, King, and White (1995), these methods are more appropriate because, unlike testing, they do not favor unfairly the model chosen under the null.
Thus, for each model fitted, we calculate AIC**,** BIC and HQC criteria:

$$AIC = -2\frac{Likelihood}{T} + \frac{2n}{T}$$

$$BIC = -2\frac{Likelihood}{T} + \frac{n}{T}\ln(T)$$

$$HQC = -2\frac{Likelihhod}{T} + \frac{2n}{T}\ln(\ln(T))$$

where n is the numbers of parameter estimated and T is the sample size. Evidently, all criteria consist of a measure of model fit that is based on the log-likelihood of a model and a penalty term for model complexity given by the number of estimated parameters.
Interestingly, standard likelihood ratio tests and goodness of fit criteria are agreed despite discrepancy related to AIC criteria outcome for consumer loan and BIC's commercial.
Given that a lower score reflects a better fit, our results show at first that MS models always outperform linear models confirming the presence of regimes in the delinquency rate series; or differently, MS models fit delinquency series better but we cannot draw a univocal conclusion regarding the most suitable MS model specification.
Specifically, when we compare the AIC, BIC and HQC values with different MS regressions, we find that all criteria prefer the MS-CV in the case of real estate series. MS-C and MS-I are chosen respectively by AIC and BIC, HQC criteria in the case of consumer data; MS-CV and MS-C are chosen respectively by AIC, HQC and BIC criteria for commercial loans[39]. These last outcomes are consistent with much of the related literature which revealed BIC to tend to underestimation and parsimony and AIC not prone to penalize model complexity so that there is no single correct solution[40].
Overall, even if no criterion is superior to others and useful information for model selection can be obtained from using them together, according to Burnham and Anderson (2004) and Aho *et alt.* (2014), criteria are appropriate for different tasks. The AIC procedure derived by Akaike is asymptotic efficient. It maximizes predictive accuracy selecting the dimension which leads to the smallest average mean squared error and it is the selection strategy that should be used when the prediction is the purpose of the study. Schwarz 'approach, on the contrary, although not asymptotically efficient, is consistent: when the sample size increases, the correct model is selected from any group of models.
Moreover, the same Burnham and Anderson (1998) provide arguments contrary to the BIC claiming that while AIC remains a viable method, the usage of BIC is theoretically inappropriate. It performs well just in large sample size and it "results in a selected model that is under fit meaning, biased parameter estimates, overestimates of precision, and achieved confidence interval coverage below that achieved by AIC". Likewise,

[39] All results are available upon request from the author.
[40] Zou and Yang (2004) cite several authors who proposed the term "model uncertainty" to emphasize the difficulty to recognize the plausible structure of the model.

Psaradakis and Spagnolo (2003) argued against the BIC in a paper on the determination of the number of regimes in MS autoregressive models based on complexity penalized likelihood methods. Authors stressed the capability of the AIC to the estimation of the state dimension more successfully with respect to BIC and the HQC, more prone to underestimate the true regime.

Table 2-10 Estimated models. Source: author calculation.

Time Series	**AIC**	**BIC**	**HQC**
Delq RE	MS-CV	MS-CV	MS-CV
Delq Cons	**MS-C**	**MS-I**	**MS-I**
Delq C&I	MS-CV	**MS-C**	MS-CV

On this sense, since literature is not clear on the best method of model selection and we are interested in a model that provides the best out of sample forecast results we follow (Psaradakis and Soagnolo, 2003) and mainly concentrate on AIC[41] indications, selecting the MS-CV for delq RE and delq C&I and the MS-C for delq Cons.

4.2.2. MS parameter estimation and empirical results

Identified a possible model for the data, the next step consists of estimating the parameters of the selected models.

All of them allow for different specification in the conditional mean of each regime. The models for real and business loans show that the two states generally differ in variance as well, so we have regime heteroscedasticity.

Table 2-11 provides the estimated parameters of all MS models. It gives a detailed overview of how changes in macroeconomic variables influence the growth in credit vulnerability in the two regimes.

Models outputs display clear differences mainly in the significances and sizes of the estimated coefficients depending on the regime. The coefficients estimates are quite similar in signs to the results of the prior baseline regressions.

The value of the intercept of each regime indicates the first regime as the high credit vulnerability state (or high NPLs regime) for all loan classes, while the low NPLs period is the second one.

Besides, looking at the transition probabilities outputs, it is evident that the probabilities that the process moves from regime one to regime two and vice versa are tiny in the case we consider MS-CV, meaning that when the process is in one state, the probability of remaining there is very large. Mixed results are obtained in the case of consumer parameter estimation with only parameter switching (MS-C): consumer loans tend to vary position more frequently.

Expected regime durations confirm such tendency. Moreover, periods of low NPLs (Regime 2) are globally longer than periods with high NPLs (Regime 1).

In the case of delinquency RE, for instance, the chance of remaining in high regime given that delinquency loans were in the same regime in the previous period is 97%. This regime has slightly less persistence as the probability to stay in the low one is 98%; in fact, the expected duration of the high regime is 30 quarters, whereas the expected duration of the low regime 41 quarters. In other words, credit vulnerability in high loan status is expected to last over than 7 years and a low credit quality to last over ten years. There is a 3.4% probability of switching from the high regime to the low regime that is greater to the opposite one (2.4%) indicating a greater opportunity for credit betterment.

[41] Notice that not necessarily, the selected model provides best forecasting results because the model can yield a good fit in the sample used for estimation but it need to translate it good forecasting performances, and vice-versa. It was Geurts and Ibrahim (1975, pag 186) who remarked: “the model that fits best is not necessarily the one that forecasts best”.

Focusing on the effects of each variable, not surprising all factors influence delinquencies in different ways, some in diminution, and other in growth.
The most significant macro impact is that of GDP, money supply and the house of price index.
The negative association with the GDP is generally corroborated and evident for all class loans. However, changes in GDP are more accentuate for real estate loans. They influence the commercial and industrial loans in high regime only, while they are even insignificant for the consumer class. In such cases, while an increase in real sector allows consumers to repay their loans successfully, its deterioration cannot be considered a fundamental reason of credit vulnerability as graphically we saw at the beginning; business loans, on the contrary, appear to be sensitive only to a deterioration of the business cycle.
Regarding the magnitude of the impact of GDP, contrary to the expectation, the rate at which delinquency rates decline in periods of economic expansion is higher than the rate at which delinquency rates grow during times of recession.
The negative relationship between delinquency aggregates and unemployment found in the linear model is still generally evident.
Specifically, mixed results are observed for consumer loans. In regime one, unemployment is positive before turning to be negative under low phase. The positive outcome is consistent with the overwhelming literature besides being obvious. In an economy that goes bad, loss of jobs or an increased unemployment reduces the purchasing power of households making it hard to pay back the loans acquired.
On the contrary, in periods of the low regime, higher unemployment leads to lower delinquent loans. The outcomes may reflect the fact that lower labor participation does not allow individuals to enjoy more disposable income. Therefore, they do not invest and increase their debt burden with the consequence that after a lag of a quarter their inability to pay does not grow too.
Moreover, the missed push of demand does not conduct to grant credit also too bad creditworthiness borrower or the fall in the demand of debt as a consequence of the fall in the level of income means a reduction in the cost of credit repayments such that individuals are less likely to miss a payment and to stay past-due.
Similarly, the negative association of the low state can be explained by the fact that the amount of loan-financed is well guaranteed. For instance, consumer loans may have supplied to eligible individuals able to manage the ongoing repayment even if fallen in unemployment thanks to a security like welfare payments, unemployment benefits, public pension plans or income earning assets. The simple ownership of a boat, a car, or some properties are decisive guarantees. Additionally, the association may be attributed to the fact that already employed borrowers, with a poor probability of default, hold loans.
Proceeding with the business delinquency amount, the nexus is negative and insignificant for the high regime, whereas it returns to be significant and negative for the lower one. Again, as previously stated, this result could be related to the short-term effect of unemployment on firm cash flow, which should increase consequently of the reduction in labor cost.
Specifically, firms' ability to hire workers at lower wages due to the greater pool of jobless, allow them to produce more and a lower unit price. Consequently, in a situation of less job layoff, the firm's output does not reduce, and it does not increase the firm's tendency to default on loan repayment.
Furthermore, regarding the linkage between delq and money supply, differences regard the significance of the variable and its magnitude across regimes; they do not include sign that registers the expected positive ones.
Specifically, while real estate and business loans show a significant response to restrictive monetary policies, consumer loans are sensitive just to the expansionary measures. Insignificances may be linked to the protracted mechanism between monetary policies and market price interest rate adjustments. Stated differently, they can be a matter of lags. As pointed out by Friedman (1961), monetary actions affect economic conditions only after one or more lag.
In line with linear model, it is also generally supported the negative sign of house prices but, again, the link differs across loans category regimes in terms of significance. The most significant effect of falling house prices is on delq RE and delq Cons. It is conceptually consistent and well-founded that, in the case of real

estate loans, a fall in the growth rate of house prices is significant just for the boom regime. Historically, indeed, housing markets have been the causes of the banking sector crisis in the 1990s and 2007. Dependency on foreign capital and losing credit standard facilitated, respectively, the expansion of mortgage credit particularly in risky area with the consequence that when house market experienced difficulties, the banking and financial sector became squeezed.

Similarly, it is reasonable the more pronounced effect of house price falling on consumer and real estate loans than on firms as well. In fact, while household investments decisions, ability to obtain funds and to repay them are exclusively (or primarily) based on the buffer offered by housing wealth and proxied by changes in house values; firms, on the contrary, can count on alternative sources such as equipment, purchase order, inventory, intangible assets, cash flows and so on.

As in the linear model, there is a positive sign associated to the coefficient of the interest rate.

However, only real estate loans and consumer losses of regime two are directly dependent on short-term variations of interest rates whereas in regime one the estimated coefficients are not statistically significant. Finally, contrary to the linear case, there is an insignificant asymmetric response of real estate loans to changes of the stock market.

The outcomes of the lagged dependent variables are consistent with the directional impacts. Previous quarters delinquency amounts are highly significant and positively related in both regimes confirming the dynamism and persistence of the losses. Slight differences concern the magnitude of the impact under the regimes and with the linear model results.

Figure 2-4 shows the smoothed probabilities of delinquency Real Estate for regime 1.

The completely smoothed probabilities for the other loan categories indicating the regime classification quality during the entire period scrutinizing are reported in figures in Appendix C.

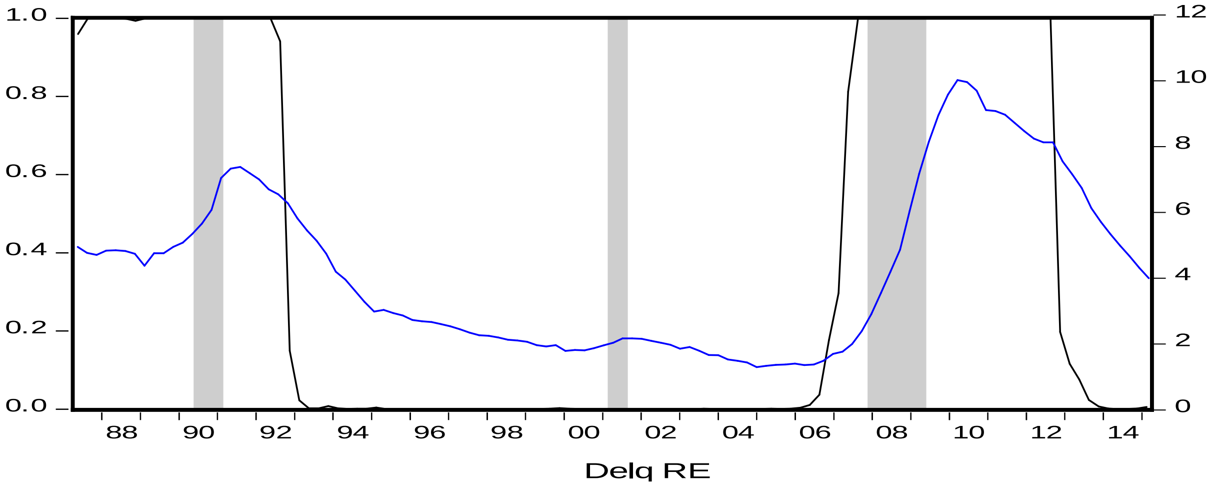


Figure 2-4 Delq RE MS-CV Smoothed. Source: author calculation

The comparison of the graphs of the different series stresses how the high regime of delq RE and delq CI is identified in correspondence of the first and the third recessions indicated by NBER, obviously with different duration. Delq CI presents another high state in correspondence of the brief intermediate recession of 2001. Contrary, the regimes of Delq Cons seem not related to the recession dates (excluding the last one of 2007) and there are several values of the smoothed probabilities around 0.5, not saying if the corresponding date falls in a recession or growth period. This series presents fewer activities evident even from a simple linear assessment.

In other terms, series Delq RE and Delq CI show a clear nonlinear pattern, whereas there is not a strong empirical support for the existence of two clear regimes in the case of Delq Cons.

Table 2-11 Markov Switching Model –Parameter estimates with all switching coefficients and variance, and parameter estimates with coefficients switching for consumer loans. SE in parentheses. The symbol * means that the t-statistics has a p-value less than 0,01**

Explanatory variable	*delq RE*	*Delq Cons*	*delq C&I*
GDP_1	-0.059[1] (0.019)***	-0.003[1] (0.007)	-0.036[1] (0.019)**
	-0.009[2] (0.006)	-0.008[2] (0.004)**	-0.005[2] (0.007)
Unemployment_1		0.057[1] (0.021)***	-0.034[1] (0.054)
		-0.039[2] (0.008)***	-0.057[2] (0.010)***
M2_1	0.022[1] (0.067)	0.085[1] (0.030)***	0.031[1] (0.061)
	0.035[2] (0.015)**	0.016[2] (0.010)	0.031[2] (0.017)**
House price index_1	-0.174[1] (0.055)***	-0.089[1] (0.017)***	-0.071[1] (0.044)
	-0.022[2] (0.018)	-0.037[2] (0.008)***	-0.029[2] (0.011)***
Tbills_1	0.035[1] (0.030)	-0.005[1] (0.013)	
	0.016[2] (0.007)**	0.022[2] (0.004)***	
S&P_1	-0.007[1] (0.006)		
	-0.000[2] (0.002)		
Delinquency variable_1	0.934[1] (0.031)***	0.772[1] (0.050)***	0.927[1] (0.039)***
	0.910[2] (0.009)***	0.933[2] (0.016)***	0.916[2] (0.015)
Constant	0.547[1] (0.350)	0.567[1] (0.217)***	0.688[1] (0.342)**
	0.096[2] (0.072)	0.383[2] (0.071)***	0.407[2] (0.071)***
σ	0.138[1] (0.032)***	Common 0.065 (0.010)***	0.040[1] (0.011)
	0.063[2] (0.013)***		0.008[2] (0.002)
Log Likelihood	65.693	135.356	64.070
Expected duration of regimes	29.512[1]	2.936[1]	14.952[1]
	40.946[2]	13.034[2]	25.714[2]
Mean	4.277 [1]	3.460[1]	3.013[1]
	3.959[2]	3.259[2]	2.681[2]

$\boldsymbol{p_{11}}$	0.966	0.659	0.933
$\boldsymbol{p_{12}}$	0.034	0.341	0.067
$\boldsymbol{p_{21}}$	0.024	0.077	0.039
$\boldsymbol{p_{22}}$	0.976	0.923	0.961

5. Forecast comparison and results

Against the estimation backdrop, we want to investigate whether the MS models provide more accurate forecasts than the linear ones.

In fact, our important objective of considering such nonlinear model with respect to the linear counterpart is to adequately describe the dynamic behavior of the observable series under consideration but also to produce adequate prediction values that are much better than those produced by simple linear models.

We address our key research questions through a comparison of the forecast evaluation measures, checking how models perform in terms both in sample and out-of-sample[42] forecasts.

For the out of sample scheme, we first use the estimated equation to produce eight steps ahead forecasts and to compare them to the actuals; afterward, we follow the recursive method. We estimate the model in sample using data from 1987 q1 to 2015 q1 to forecast 2015 q2 and in the next steps, we re-estimate model parameters from 1987 q1 to 2015 q2 to forecast 2015 q3, and etc.

Specifically, we focus on the Mean Squared Error (MSE) and Mean Absolute Error (MAE) that can be considered as the very early measures.

They are scale dependent statistics meaning that they depend on the scale of the dependent variable, and multiplying values of the explained variable by any scalar multiplies the measure by that scalar as well.

In mathematical notation, they are based on the forecasts error.

They are used to compare forecasts for the same series[43] and sample across different time series models and the better the forecasting ability of the model the smaller MSE and MAE error statistics; a good model for forecasting produces the minimum forecast errors as compared to other competing models.

Let be Y_t the observation at time t, $\hat{Y}_t$ the forecasted value of Y_t and n the forecast sample size. Then, the forecast error is defined as $e_t = Y_t - \hat{Y}_t$ and

$$MSE = \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

The former is the average of the squares of the errors of each item in the data; the latter is the mean of all absolute errors.

Throughout our work, as we do not compare the forecast accuracy of different variables, we mainly concentrate upon such measures for a preliminary descriptive analysis of the forecast errors.

Proceeding, to have a most exhaustive of model performance, we also have results of the scale-free or invariant measures[44] based on percentage errors and on the Theil U statistic as well as we take into account the three

[42] In sample forecast means predictions generated for the same data set used to develop the model; the out of sample forecasts are those made for a period outside the data set used to estimate the model's parameters.

[43] They should not be used for comparing variables across data sets that have different scales (Hyndman and Koehler, 2006).

[44] Their purpose is to facilitate comparison between different variables. As the percentage error is given by $p_i = \frac{100\, e_i}{y_i}$,

indicators derived from the decomposition of the mean squared forecast error: the bias, variance, and covariance proportions. They show, respectively, how the mean and variance of the forecast deviate from those of the actual series; the covariance is a measure of the unsystematic forecasting errors. If the forecast is “good”, the bias and variance proportions should be small so that there is no indication of the systematic error, the mean and variance of the dependent variable are well tracked by the forecasts, and most of the bias should be concentrated on the covariance proportions that should be ideally equal to one.

Along with measuring forecasting accuracy *per se,* our evaluation of predictive accuracy comprises the comparison of various forecasting models through a formal test.
In fact, we can find that the forecasting results of Markov switching models look intuitively better than the other models for all loan category estimation, both in sample and out of sample. The higher forecast accuracy of the MS model has generally been supported by the results of both MSE and MAE.
However, it is interesting to know whether there is significant difference in the forecast from the two models. Some small differences between linear and Markov model forecast accuracies in the consumer and commercial case make it difficult to distinguish and decide whether the difference is due to chance and “the victory in sample is merely good luck, or truly indicative of a difference in population” Diebold (2013).
Therefore, it is necessary and interesting to employ a formal quantitative criterion to select the optimum method.
The Diebold and Mariano (1995) test statistic (DM) offers the statistical tool to know whether there is significant difference in forecasts from the two models. It will allow to evaluate the forecast accuracy and to differentiate the forecasting performance of models.
Therefore, to assess the difference of the forecasting performance more rigorously, we too compute the DM test.
According to Diebold (2013), while the statistics were originally intended to evaluate forecasts that aren´t based on econometric models, the so-called model-free forecasts, successively literature took a very different course and it has been common to use the DM-statistic to compare the forecasting ability of econometric models.
The test compares model loss functions to verify the null hypothesis of equal predictive ability against the alternative that there is difference between the forecast accuracy from the two models so that one is better than the other. The test, in other words, verifies the null hypothesis that the forecasting performance of the two models is equally good.

Let be $(\hat{y}_{1t}, \hat{y}_{2t})$ the h-steps ahead[45] competing forecasts of the time series y_t, with $t = 1, \dots, T$.
The corresponding forecast errors from the $i-th$ competiting models are e_{it} $(i = 1, \dots, m)$ where m is the number of models. The loss associated with forecast i is assumed to be function of the forecast error and it is denoted by $g(e_{it})$. It is a loss function that maps forecasts on the positive real number R_+ and may be equal to the square (squared error loss) or the absolute value (absolute error loss) of e_{it}:

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_i}\right|.$$

It is the mean of the absolute percentage errors of forecasts divided by the actual value.
Finally, the Theil inequality coefficient always lies between zero and one, where zero indicates a perfect fit. It is just the root mean squared error normalized by the dispersion of actual and forecasted series:

$$U = \frac{\sqrt{\frac{\sum_{i=1}^{n}(y_t - \hat{y}_t)^2}{n}}}{\sqrt{\frac{\sum_{i=1}^{n} y_t^2}{n}}\sqrt{\frac{\sum_{i=1}^{n} \hat{y}_t^2}{n}}}$$

[45] In this work, h is set equal to one.

$$g_1(e_{it}) = g(y_t, y_{it}) = \sum_{i=1}^{T}(e_{it})^2$$

$$g_2(e_{it}) = g(y_t, y_{it}) = \sum_{i=1}^{T}|e_{it}|$$

In hypothesis testing terms, the DM null and the alternative are given as:

$$H_0 = E[g(e_{it})] = E[g(e_{jt})]$$

$$H_0 = E[g(e_{it})] \neq E[g(e_{jt})].$$

Equivalently, the null can also be expressed in terms of loss differential series associated with each of the two forecasts, $d_t = g(e_{it}) - g(e_{jt})$.
More precisely, under the assumption of loss differential covariance stationary[46], the expectation operator in the above equation can be replaced and the key hypothesis of equal predictive accuracy corresponds to testing whether the loss differential has zero expectation for all t:

$$H_0 = E(d_t) = 0$$

According to Diebold and Mariano (1995), the test is asymptotic and standard results can be used to deduce the asymptotic distribution of the sample mean loss differential.
Denoting the sample mean loss differential, $\bar{d}$, with:

$$\bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t = \frac{1}{T}\sum_{t=1}^{T}[g(e_{it}) - g(e_{jt})]$$

Then the distribution of the difference between the sample mean and the population mean of the loss differential when multiplied by the factor T is described by:

$$\sqrt{T}(\bar{d} - \mu)\vec{d}N(0, 2\pi f_d(0))$$

where $f_d(0) = \frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty}\gamma_d(\tau)$ is the spectral density of loss differential at frequency 0 and $\gamma_d(\tau) = E[(d_t - \mu)(d_{t-\tau} - \mu)]$ is the autocovariance of loss differential at the time difference τ.
Following Diebold and Mariano's argumentations, the serially correlation of forecast errors that emerges from the formula for $f_d(0)$ requires the calculation of robust standard errors of the loss differential. Given that, in large sample, the sample mean loss differential is approximately normally distributed with mean μ and variance $\frac{2\pi f_d(0)}{T}$, the large-sample statistics for testing the null hypothesis of equal forecast accuracy is:

$$DM = \frac{\bar{d}}{\sqrt{\frac{2\pi \hat{f}_d(0)}{T}}} \xrightarrow[d]{} N(0,1)$$

[46] The assumption is summarized by the following properties: 1) $E(d_t) = \mu \;\; \forall t$; 2) $cov(d_t, d_{t-\tau}) = \gamma(\tau)$; 3) $0 < var(d_t) < \infty$. In other words, the series has contant mean, constant variance and constant autocovariance structure.

where $2\pi\hat{f}_d(0)$ is a consistent estimator of the asymptotic variance.
Summarizing, under the DM assumptions, the test statistics is asymptotically normal distributed. This leads to use the normal critical values for model comparison: the null hypothesis of no difference is rejected if the computed DM statistic falls outside the range of $-z\alpha/_2$ to $z\alpha/_2$, that is $|DM| > z\alpha/_2$, where $z\alpha/_2$ is the upper z value from the standard normal table corresponding to half of the desired α level of the test.
However, as noticed by the same authors, the normal distribution can be a very poor and an unsatisfactory approximation of small sample distribution. In such cases, the test tends to be somewhat oversized, rejecting the null too often.
Harvey, Leybourne and Newbold (1997) suggested a DM statistics modification to improve finite samples properties, based on Student t critical values rather than those from a standard normal; the statistic they propose can be written as:

$$HLN = \sqrt{\frac{T + 1 - 2h + h(h-1)}{T}}DM$$

Consequently, since our forecast evaluation sample is of a moderate size, we apply the modified Diebold-Mariano test suggested in Harvey *et al*.

Our results from the in-sample forecasting exercise seem to be satisfactory and robust across the two statistical criteria. They greatly lean toward the switching regression, indicating its outperformance over the benchmark; they are homogenous across loans divisions, barely mixed in nature.
Likewise, the simple out of sample results show the regime-switching model to be globally better for the first two loan classes. The linear regression model is mainly preferred for business loans.
In the recursive case, the non-linear model does not forecast better than linear ones for a business loan when the criterion is MSE.
Proceeding and comparing the out of sample forecasts of different models within the same type, we also verify that the model with minimum AIC and BIC values does not necessary provides best forecasts. Meaningful the case of consumer delinquency rate: the two models that give the minimum information criteria (the estimated constant and parameters switch) do not produce minimum forecast errors as compared to third model (all parameters and variance switch) that is highly significant both in sample and out of sample.
The Diebold Mariano test recognizes the Markov switching accuracy at highly significant levels; in the in sample case, it is only significant in AE difference for consumer loans; for the other cases, differences are statistically significant according to both the squared errors and absolute errors.
We verify the results looking at the p-values based on the t-Student distribution, and since those of the DM test based on the absolute and squared error loss and (DM-AE) (DM-SE) are lower than 0.05, the null hypothesis can be rejected.
The same predictive power is supported out of sample where the DM favors the opposite linear regression model just in the case of business loans. Mixed in nature results concern the case of the recursive forecasting exercise where the superiority of the MS is significant only for the delq RE case whereas the DM partly yields insignificant results in the consumer and business cases.
Hence, the MS globally proves our starting hypothesis, we show that both in sample and out of sample, the forecast performance is superior to the linear regression and we also find that the regime switching models are plausible alternative models that can be used under the nonlinearity and structural change conditions.
Surely, the strong detectable evidence of nonlinearity in the data is a first valid explanation of the relatively large general performance of nonlinear forecasting models. It is strong enough to make a difference in forecasting.
In addition, it is plausible that MS is the correct type of model to capture the presence of non-linearity.

In fact, there is a consistent part of literature that has noticed the inability of Markov switching models to outperform linear model in macroeconomic out of sample forecast in spite attractive in sample properties[47] and it justifies in some measure our outcome as well. However, a minimum degree of switching non-linearity in out and in sample relationships enhances forecast accuracy (Arora, Little and McSharry, 2013).

In this sense, the out of sample features in the data can explain the negative results for commercial and industrial loans: the absence of variation in the data in the period of forecasting can explain the inferior performance of MS. It may be that non-linear observations became infrequent so that there is no gain from using MS.

The same reason holds for the recursive forecasting where equality between models can be ascribed to the fact that in the small number of out of sample forecasts, the non-linearity does not show up and it does not arise the edge of the non-linear ones.

[47] See Dacco and Satchell (1999) for a theoretical explanation. They ascribe the poorly performance to a missclassification of future regimes or saying it differently, to the difficulty of forecasting the regime that the series will be in. Hansen (2010) also makes this point arguing that estimated models tend to have a better fit in sample than out of sample.

Table 2-12 Forecast errors using static scheme.
MS-CV for delq tot, RE, and C&I; MS-C for consumer.
Source: author calculation

Forecast evaluation statistics: in sample fit ((1987q1-2015q1))[48]			
	Delq RE	**Delq Cons**	**Delq C& I**
Simple linear regression static forecast			
MSE	0.0540	**0.0095****	0.0326
MAE	0.1770	0.0686	0.1263
Markov Switching static forecast			
MSE	**0.0347*****	0.0107	**0.0294*****
MAE	**0.1247*****	**0.0675***	**0.1158*****

Table 2-13 Simple out-of-sample forecast

Forecast evaluation statistics: out of sample fit (1987q2-2017q1)			
	Delq RE	**Delq Cons**	**Delq C& I**
Simple linear regression static forecast			
MSE	0.0206	0.0026	**0.0250*****
MAE	0.1265	0.0420	**0.1190*****
Markov Switching static forecast			
MSE	**0.0016*****	**0.0025*****	0.0376
MAE	**0.0339*****	**0.0415*****	0.1471

Table 2-14 Recursive out-of-sample forecast

Forecast evaluation statistics: recursive out of sample fit (2015q2 first obs, 2017q1 last obs)			
	Delq RE	**Delq Cons**	**Delq C& I**
Simple linear regression static forecast			
MSE	0.0201	0.0027	**0.0250**
MAE	0.1273	0.0430	0.1212
Markov Switching static forecast			
MSE	**0.0013*****	**0.0026**	0.0292
MAE	**0.0290*****	**0.0426**	**0.1178**

[48] Asterisks stay for the significance of the MSE and MAE difference between models estimated by the DM statistics. Again, ***, **, * indicates statistical significant at the 1%, 5%, and 10% level, respectively.

Conclusion

Financial stability is a necessary condition for a sustained economy. NPLs or delinquency loans are among the prior indicators of financial soundness reflecting asset quality; they are very crucial for banks' liquidity and solvency: the direct consequence is bank failure.
Therefore, the importance of large amounts of problem loans in any developed or developing country cannot be denied, acquiring growing attention in recent times.
A broad area of studies has noticed and commented on the impact of macroeconomic factors on loan performance but despite the abundance of studies, none has ever focused on forecasting the nonlinearities inherent in the relationship.
This study is motivated by the intention to bridge the gap, capturing the natural asymmetries and changes along different phases of the economic cycle.
It presents a macro model that reexamines the link between bad debts and macro factors for the USA using the Markov switching approach to understanding whether results are homogeneous across regimes and different financing classes, stressing on the disaggregate specification.
The study provides evidence about the statistical validity of the non-linear approach, documenting the uniformity across loans and supporting the greater and accurate predictive ability of MS as well.
Indeed, after presenting the non linearity issue, it compares the forecasting performance of MS with the linear regression.
The model has been proved good in modeling and in forecasting since it is able to manage the asymmetries features in the data.
Specifically, findings are largely consistent with literature regarding the sign of the aforementioned variables except for the surprising results of the unemployment variable that contradict past research. Moreover, it has been underlined that delinquency rates display different features especially regarding the magnitude of the impact during economic expansion and recession, a period of unemployment and no, high and low-interest rate phases and so on, that cannot be captured by the simple linear model.
Comparisons of the forecasts from a nonlinear functional form with the benchmark linear model show the advantage of considering nonlinearities. We find evidence that MS globally dominates the linear one specifically for the delinquency real estate amount.
Therefore, it can be concluded that hypothesis one and two holds and the study is unique not only in making a non-linear comparative analysis of the macro determinants that affect delinquency rates of different types but also in assuming a forecasting perspective.

Consequently, results can be a guide for investors to have a better asset allocation and portfolio risk management.
Likewise, results can be beneficial to American financial institutions.
Indeed, the examination could have implications in terms of regulation and policy. Namely, as macro factors are important predictors of an increase of delinquency loans, authorities could use results to tailor proper and responsive policies to deal with different types of loans and to diversify strategies during the different economic trends in accordance with the basics of macroeconomics. Identifying in which periods loans became more troubled, it would be useful for the betterment of the financial health, in order to strengthen the banking sector against failures and prevent a rise of credit risk.
Besides, the management of the banks can refer to the performance of macroeconomic conditions to predict the performance of their bad loans when they manage the credit risk of their loans in order to avoid the possibility of increasing defaults.

In the light of the above findings, the work can be extended in many ways.

Obviously, the limits of this document represent opportunities for future research. As we have noted, the work is based on American data referring to commercial banks, and the decision to conduct analysis on exclusively this data was caused by practical reasons of data availability. In fact, it has been difficult to find data differentiated by categories of banks over a long time horizon, and for European countries, data on NPL or bad debts are not available for each category.

Therefore, bearing in mind that data availability will be a major challenge for achieving different results in terms of comparability, future works could come to substantial conclusions on the non-linear impact of macroeconomic variables on bad debts by using a larger country data set and a broader time framework (consider that we have used quarterly data, but MS may be more comprehensive with more observations and longer research periods).

Further developments include a differentiated analysis depending on the different types of banks, on the sub-components of each specific loan category and an enlarged analysis in order to consider the role played by regulation and supervision instead of relying only on the macro environment in the strict sense.

To put it differently, banks operate in one of the most regulated industries at both the state and federal level; therefore, a consideration of the macro regulatory and supervisory practices and their potential coincidence with changes in regimes could provide additional insights into the influence of the macro environment on banks loan quality and extend the paper’s work.

In fact, regulatory devises like Basel I, II and III Capital Accords, working as instruments of capital, liquidity requirements and market discipline, systematically impact the risk taking behavior of banks and hence their amount of bad debts: they should reduce market failures by monitoring banks and improving the quality of bank lending.

An important implication is that changes in the level of capital requirements influence not only bank lending, but also the whole economy (the famous pro-cyclical effects of Basel). As instance, stringent capital requirements shrink credit growth to households and firms, raises spreads on home mortgages and on corporate bonds, damps down housing market activity, with both a drop in house price and a rise mortgages in arrears as well.

Therefore, it is not surprising that regulatory and deregulatory strategies create the economic environment and they not just suffer from its whims.

Keeping this in mind, future researches and a new manuscript could focus this topic extending the comparative assessment.

APPENDIX A: Sources of Macro-economic Variables

Macroeconomic variables	Source
Real GDP	Bureau of Economic Analysis
Unemployment	International Monetary Fund
Inflation (CPI)	International Monetary Fund
Money supply (M2)	International Monetary Fund
House price index	FHFA United States (Federal Housing Financing Agency)
S&P 500	Standard &Poors's
T-bills 3 months	Federal Reserve Board
Real Effective Exchange rate (REER)	International Monetary Fund

APPENDIX B: Perron, ZA and CMR tests

Perron's testing procedure allows for a break under both the null and alternative hypothesis. The null considers a unit root with an exogenous break at time $1 < T_B < T$, and the alternative of a stationary process around a deterministic time trend function at time T_B, where T_B is the point of break and T is the total number of observations.

In particular, the author proposed three models according to the type of breaks: 1) change in the intercept/level of the series (crash model), 2) change in the slope of the rate of growth (changing growth model), 3) change in the intercept and the slope simultaneously (combined model). Practically, each of these models is an extension of the standard Dickey-Fuller procedure by adding dummy variables to represent different intercepts and slopes. These models provide different results but suppose a unit root with break under the null and a broken stationary process under the alternative and are tested via these following regressions:

Null Hypothesis:

1) $y_t = \mu + dD(TB)_t + y_{t-1} + \varepsilon_t$

2) $y_t = \mu_1 + y_{t-1} + (\mu_2 - \mu_1)DU_t + \varepsilon_t$

3) $y_t = \mu_1 + y_{t-1} + dD(TB)_t + (\mu_2 - \mu_1)DU_t + \varepsilon_t$

Alternative Hypothesis:

1) $y_t = \mu_1 + \beta t + (\mu_2 - \mu_1)DU_t + \varepsilon_t$

2) $y_t = \mu + \beta_1 t + (\beta_2 - \beta_1)DT_t^* + \varepsilon_t$

3) $y_t = \mu + \beta_1 t + (\mu_2 - \mu_1)DU_t + (\beta_2 - \beta_1)DT_t + \varepsilon_t$

The first model allows for an exogenous change in the levels of the series under the null hypothesis ($D(TB)_t = 1$ if $t = T_B + 1$, zero otherwise) and a one-time change in the intercept of the trend function under the alternative hypothesis ($\mu_2 < \mu_1$). Model 2 specifies an exogenous change in the rate of growth under the null hypothesis and in the alternative hypothesis a change in the slope of the trend function without any sudden change in the level at the time of break. Model 3 allows the combined effects to take place simultaneously: a sudden change in the level and in the growth path.

Following Perron's, ZA proceeds with three models to test for a unit root. While the first model allows a one-time change in the level of the series and the second permits a one-time change in the slope of the trend function, the last model combines one-time changes in the level and the slope of the trend function of the series[49]:

Null Hypothesis H_0:

$y_t = \mu + y_{t-1} + \epsilon_t$

Alternative Hypothesis:

[49] The modified models do not include the crash dummy DTB.

1) $y_t = \mu_1 + \beta t + (\mu_2 - \mu_1)DU_t + \varepsilon_t$

2) $y_t = \mu + \beta_1 t + (\beta_2 - \beta_1)DT_t^* + \varepsilon_t$

3) $y_t = \mu + \beta_1 t + (\mu_2 - \mu_1)DU_t + (\beta_2 - \beta_1)DT_t + \varepsilon_t$

Hence, nesting the models under the null and alternative hypotheses, the model involves the estimation of the following regression equations corresponding to the above three models:

1) $y_t = \mu + \theta DU_t(\hat{\lambda}) + \beta t + \alpha y_{t-1} + \sum_{j=1}^{k} c_j \Delta y_{t-j} + \varepsilon_t$

model with intercept (1)

2) $y_t = \mu + \beta t + \gamma DT_t^*(\hat{\lambda}) + \alpha y_{t-1} + \sum_{j=1}^{k} c_j \Delta y_{t-j} + \varepsilon_t$

model with trend (2)

3) $y_t = \mu + \theta DU_t(\lambda) + \beta t + \gamma DT_t^*(\hat{\lambda}) + \alpha y_{t-1} + \sum_{j=1}^{k} c_j \Delta y_{t-j} + \varepsilon_t$

model with both intercept and trend (3)

In the above equations, DU_t is an indicator variable capturing the mean shift occurring at each possible break date (it takes the value zero for all year prior to the break and one for all dates thereafter: $DU_t(\lambda) = 1$ if $t > \lambda T$, 0 otherwise); DU_t is the corresponding trend shift variable ($DT_t^*(\lambda) = t - \lambda T$ if $t > \lambda T$, 0 otherwise) and $\hat{\lambda}$ the estimated value of the break function.

Amongst all possible break-points TB, the procedure selects as its choice of break-date $\lambda = {}^{TB}/_{t}$ the date where the t-statistic from the ADF test of unit root is at a minimum (most negative); that is to say, λ is chosen in such a manner that one-side t-statistic $\hat{\alpha} = 1$ is minimized. Consequently, the break date will be chosen where the evidence is least favorable for the unit root null or differently: it will choose "the breakpoint that gives the most weight to the trend-stationary alternative" (Zivot and Andrews, 1992), page 254.

If λ_{inf}^{i} denotes the minimizing value for the model i, then the criterion to determine the point of structural break endogenously corresponds to:

$$t_{\widehat{\alpha i}}[\lambda_{inf}^{i}] = {}_{\lambda\Delta}^{inf} t_{\widehat{\alpha i}}(\lambda) \tag{4}$$

where $\boldsymbol{\Delta}$ is a closed subset of (0,1).

Clemente- Montañés-Reyes (CMR) is similar to Zivot Andrews.

A finite order autoregressive model estimates the unit root hypothesis in the IO model, which assumes a change in the intercept term supposed to affect the level of the series gradually through a transition period:

$$y_t = \mu + \alpha y_{t-1} + \delta_1 DTB_{1t} + \delta_2 DTB_{2t} + d_1 DU_{1t} + d_2 DU_{2t} + \sum_{j=1}^{k} c_j \Delta y_{t-j} + \varepsilon_t \tag{5}$$

where $DU_{it}(= 1$ for $i = 1{,}2$ if $t > TB_i$, and zero otherwise) is an intercept-shift parameter and $DTB_{it}(= 1$ for $i = 1{,}2$ if $t > TB_i + 1$, and zero otherwise) is a pulse variable or crash dummy.

This test provides evidence against the unit root null hypothesis if an estimate of α is significantly less than unity.

Moreover, the unit root hypothesis is tested by a two-steps procedure if shifts are better described as additive outliers.

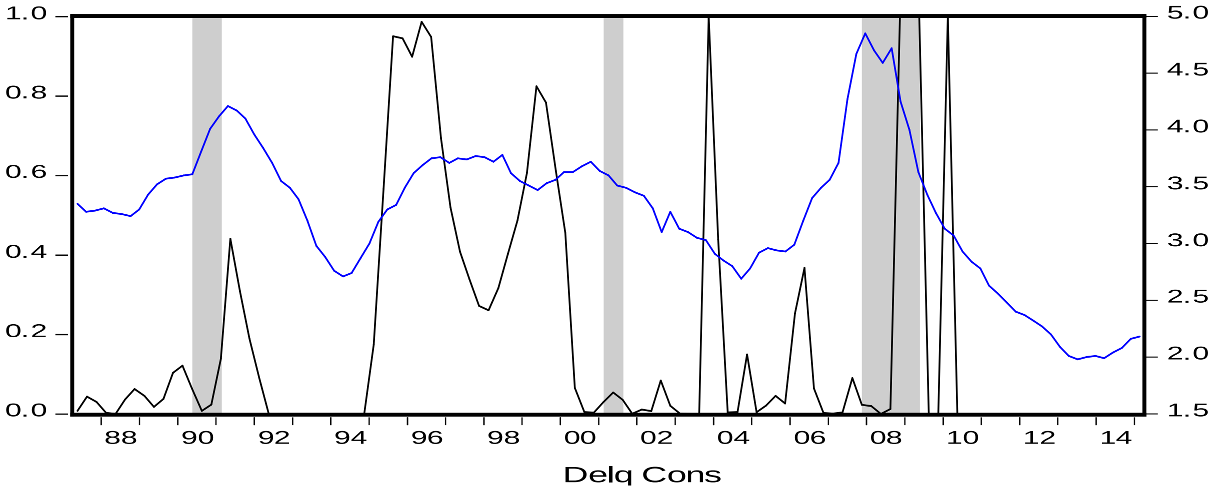
First, it is removed the deterministic part of the series using the estimates of the regression:
$y_t = \mu + d_1 DU_{1t} + d_2 DU_{2t} + \tilde{y}_t$ where $\tilde{y}_t$ denotes the residuals; DU_{1t} is equal to 1 for periods t after the first break and 0 otherwise, DU_{2t} is equal to 1 for periods t after the second break and 0 otherwise, TB_1 and , TB_1 are the breakpoints located by a grid search.
Then, residuals (that are the dependent variable in the equation to be estimated) are regressed on their lagged values and lagged differences in order to test whether the remaining noise is characterized by unit root:

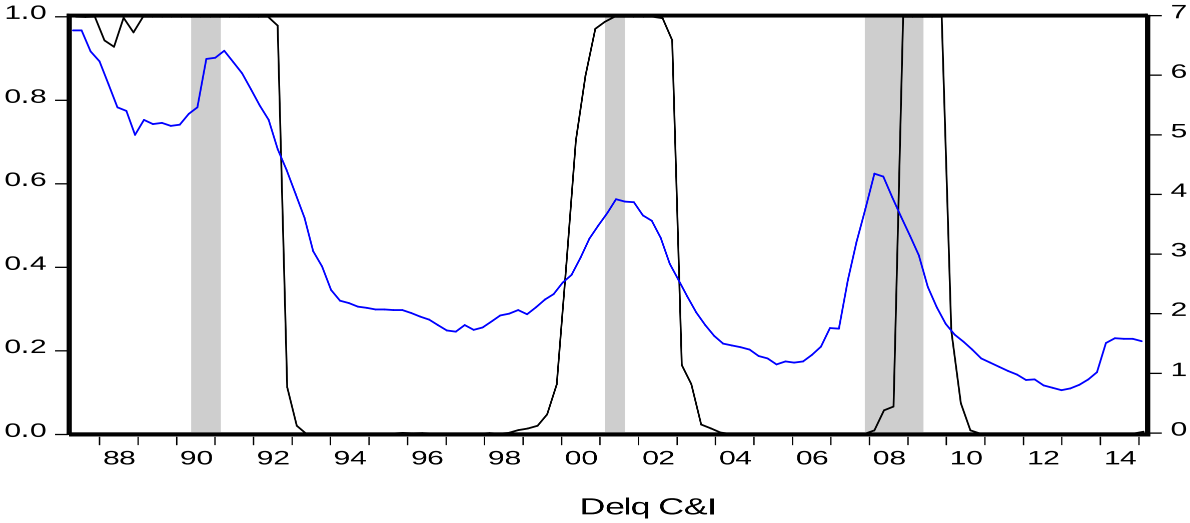
$$\tilde{y}_t = \sum_{i=0}^{k} w_{1i} + DTB_{1t-i} + \sum_{i=0}^{k} w_{2i} + DTB_{2t-i} + \alpha \tilde{y}_{t-1} + \sum_{i=1}^{k} c_i \Delta \check{y}_{t-i} + \varepsilon_t \qquad (6)$$

The regression is estimated over possible pairs of periods, searching for the minimal t-ratio for the hypothesis α = 1; that is, the strongest rejection of the unit root null hypothesis. The value of this minimal t-ratio is compared with critical values provided by Perron and Vogelsang, as they do not follow the standard Dickey-Fuller distribution.

APPENDIX C

Smoothed probabilities of regime 1 derived from the MS model for Delq Cons and Delq C&I





Note. Completely results are available from the author upon request.

REFERENCES

Babouček and Jančar. (2005). A VAR Analysis of the Effects of Macroeconomic Shocks to the Quality of the Aggregate Loan Portfolio of the Czech Banking Sector. *CNB Working Paper Series.*

Aho, Derryberry and Peterson. (2014). Model selection for ecologists: the worldviews of AIC and BIC. *Ecology, 95(3),* 631–636.

Armstrong. (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners.* Springer Sciences.

Armstrong, Green and Graefe. (2010). Forecasting Principles. *International Encyclopedia on Statistical Science.*

Arora, Little and McSharry. (2013). Nonlinear and nonparametric modeling approaches for probabilistic forecasting of the US gross national product. *Studies in Nonlinear Dynamics and Econometrics 17, no. 4.*

Bai Perron. (2003). Computation and Analysis of Multiple Structural Change Models. *Journal of Applied Econometrics,18*, 1-22.

Baum. (2005). Stata: The language of choice for time-series analysis? *The Stata Journal 5, Number 1*, 46-63.

Ben-David, Lumsdaine and Papell. (2003). Unit Root, Postwar Slowdowns and Long-Run Growth: Evidence from Two Structural Breaks. *Empirical Economics, 28 (2),* 303-319.

Bernanke and Gertler. (1989). Agency Costs, Net Worth, and Business Fluctuations. *American Economic Review*, 14-31.

Bernanke, Gertler and Gilchrist. (2015). *The Financial Accelerator in a Quantitative Business Cycle Framework.* Handbook of Macroeconomics, Volume 1, Edited by J..B. Taylor and M. Woodford.

Box. (1979). Some problems of Statistics and everyday life. *Journal of the American Statistical Association, 74,* 1- 4.

Burnham and Anderson. (2004). Understanding AIC and BIC in Model Selection. *Sociological Methods & Research,33,* 261-304.

Burnham and Anderson. (1998). *Model Selection and Multimodel Inference.* Springer.

Byrne and Perman. (s.d.). Unit Roots and Structural Breaks: A Survey of the Literature. Glasgow: Department of Economics, University of Glasgow.

Christiano. (1992). Searching for a break in GPN. *Journal of Business and Economic Statistics, 10*, 237-249.

Clemente, Montañés and Reyes. (1998). Testing for a unit root in variables with a double change in the mean. *Economics Letters 59*, 175-182.

Dacco and Satchell. (1999). Why do regime-switching models forecast so badly? *Journal of Forecasting, vol 18, Issue 11*, 1-16.

Dickey and Fuller. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association, 74,* 427-431.

Diebold. (2013). A Personal Perspective on the Use and Abuse of Diebold-Mariano Tests. The University of Pennsylvania.

Diebold and Kilian. (2000). Unit Root Tests are Useful for Selecting Forecasting Models. *Journal of Business and Economic Statistics, 18*, 265-273.

Diebold and Lopez. (1995). *Forecast Evaluation and Combination.* Federal Reserve Bank of New York, Research Paper.

Diebold and Mariano. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics, 13*, 253-265.

Escudero. (2000). A Primer on Unit-Roots and Cointegration. *Trabajo Docente Nº 3, Departamento de Economía, UNLP.*

Fisher. (1911). The Purchasing Power of Money. *Publications of the American Statistical Association, Vol. 12, No. 96,* 818-829.

Geurts and Ibrahim. (1975). Comparing the Box-Jenkins Approach with the Exponentially Smoothed Forecasting Model Application to Hawaii Tourists. *Journal of Marketing Research, Vol. XII,* 182-188.

Ghosh. (2017). Sector-specific analysis of non performing loans in the US banking system and their macroeconomic impact. *Journal of Economics and Business, 93, 29-45*

Glynn, P. V. (2007). Unit root tests and structural breaks: a survey with applications. *Journal of Quantitative Methods for Economics and Business Administrations, 3(1),* 63-79.

Granger and Newbold. (1974). Spurious Regression in Econometrics. *Journal of Econometrics 2,* 111-120.

Granger, King and White. (1995). Comments on testing economic theories and the use of model selection criteria. *Journal of Econometrics*, 173-187.

Gujarati. (2007). *Basic Econometrics.* New Delhi: Tata McGraw-Hill.

Hamilton. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica 57*, 357-384.

Hansen. (2001). The new econometrics of structural change: Dating breaks in US labor productivity. *Journal of Economic Perspectives, Vol. 15*, 117-128.

Hansen. (2010). A Winnerís Curse for Econometric Models: On the Joint Distribution of In-Sample Fit and Out-of-Sample Fit and its Implications for Model Selection. *working paper*.

Harvey, Leybourne and Newbold. (1997). Testing the Equality of Prediction Mean Squared Errors. *International Journal of Forecasting, 13*, 281-91.

Hyndman and Koehler. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting 22 (4)*, 679-688.

Kwiatkowski, Phillips, Schmidt and Shin. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. How sure are we that economic time series have a unit root? *Journal of Econometrics, 54*, 159-178.

Lee and Strazicich. (2004). Minimum LM Unit Root Test with Two Structural Breaks. *Working paper, Department of Economics, Appalachian State University*.

Louzis et alt. (2011). Macroeconomic and bank-specific determinants of non performing loans in Greece: a comparative study of mortgage, business and consumer loan portfolios. *Bank of Greece working Paper*.

Lumsdaine and Papell. (1997). Multiple Trend Breaks and the Unit Root Hypothesis. *Review of Economics and Statistics, 79 (2)*, 212-218.

MacKinnon. (2010). Critical Values for Cointegration Tests. *Queen's Economics Department Working Paper No. 1227*.

Maddala and Kim. (2003). *Unit Roots, Cointegration, and Structural Change.* Cambridge, Cambridge University Press.

Nelson and Plosser. (1982). Trends and random walks in macroeconomic time series. *Journal of Monetary Economics 10*, 139-162.

Ng and Perron. (1995). Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of American Statistical Association, 90*, 268-281.

Perron. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica, 57*, 1361-1401.

Perron. (1989). The Great Crash, the Oil Price Shock,and the Unit Root Hypothesis. *Econometrica, 57*, 1361-1401.

Perron and Vogelsang. (1992). Nonstationarity and Level Shifts with an Application to Purchasing Power Parity. *Journal of Business and Economic Statistics, 10*, 301-320.

Pesaran and Timmermann. (2004). How costly is it to ignore breaks when forecasting the direction of a time series? *International Jornal of Forecasting, Vol 20, Issue 3*, 411-425.

Philips. (1986). Understanding Spurious Regression in Econometrics. *Journal of Econometrics, Vol.33, Issue 3*, 311-340.

Psaradakis and Spagnolo. (2003). On the determination of the number of regimes in Markov-Switching Autoregressive models. *Journal of Time Series Analysis Vol 24*, 237-252.

Rappoport and Reichlin. (1989). Segmented Trends and Non-Stationary Time Series. *The Economic Journal Vol. 99, No. 395*, 168-177.

Yang. (2005). Can the Strengths of AIC and BIC Be Shared? *Biometrika 92, no. 4*, 937–950.

Zivot and Andrews. (1992). Further evidence on the great crash, the oil price shock, and unit root hypothesis. *Journal of Business and Economic Statistics, 10*, 251-270.

Zou and Yang. (2004). Combining time series models for forecasting. *International Journal of Forecasting 20*, 69-84.

Chapter 3

Links between credit quality and macro variables: Evidence from USA with a Dynamic Conditional Correlation GARCH

Abstract

Credit quality depends on macroeconomic indicators. Despite this relationship has been considered in many empirical studies over the past decades, in this chapter, we review the phenomenon more systematically by performing a useful and popular methodology that addresses the time-varying structure of the relationship using the dynamic conditional correlation model. Quarterly data for the US from April 1985 to December 2017 are adopted.

Our main findings are as follows.

(1) The correlation "delinquency rate-GDP" is indeed time-varying and negative over the entire period; (2) the variability remains when considering lagged values with an observed lower increment in absolute terms during periods of financial crisis; (3) no evidence supports the declaring dynamic relationships with other macro variables, except in lagged terms; (4) regime switching dynamics into the conditional correlation do not fit empirical data.

Keywords: delinquency rates, macroeconomic factors, DCC, MS-DCC, USA.

Introduction

As seen in the previous chapters, over the past decades, the demand for a better understanding of the causes of financial instability has put a pressure on researchers to investigate the relationship between credit quality and macro variables.
The variability of time spans and country frameworks used in researches produced a growing academic debate and the theme became of one of the top spoken topics especially after the 2008 financial crisis.
Many practitioners inside and outside of the USA reasserted intensively the association between macroeconomic variables and loan portfolio quality and they always found clear evidence of the interaction among the variables: the existence of correlation and an almost unique, quite rarely different, sign of the correlation direction.
However, to the best of our knowledge, no work revealed and evaluated the nature of time-varying correlation among such time series adopting advanced methodologies in financial econometrics and questioning the assumption that correlations are constant with respect to time, a declaring assumption in nearly all the literature in this area. See Socol and Iuga (2010), Iuga and Lazea (2012), Murumba (2013), Messai and Jouini (2013), Filip (2015), Anjom and Karim (2016), Adeola and Ikpesu (2017), Filip (2017), etc.
Specifically, few of them face the co-movement of problem debts with key macroeconomic variables. In the case where works investigate the statistical dependence, they adopt the classical correlation, while the real causality may be nonlinear and dynamic due to changeable economic or financial conditions or major events.
The previous chapters had introduced the issue of non-linearities in the time series relations by proving, through a Markov switching approach that allows the series to follow different behaviors at a different point in time, the existence of different delinquency rates regimes that is of different relationships between credit quality and economic environment.
This chapter wants to move forward in the analysis of non-linearity focusing on modeling co-movements and correlations with the motivation of giving a quantitative representation of how much stable and lasting linear relationship between real and financial sector is. Results testify that correlations might increase during the period of financial and economic turmoil and reverse from positive to negative. Consequently, the simple correlation coefficient cannot be able to give a sufficient interpretation of the true nature of the relationship.
Bearing this in mind, this work adds to the broader literature on credit quality in several aspects.
Firstly, it will contribute to the analysis of the long-term relationships among variables by examining the time series properties, measuring the number of periods in which they are correlated and providing an in-depth time-varying examination.
Secondly, the obtained empirical results might be relevant for bank management, central banks, policymakers and regulators working to stabilize financial markets. Understanding what typically corresponds to magnified correlations could help to control the amount of risk, minimize bank expenses, implement a mechanism that restrains and limit potential contagious effects, making financial systems more resistant to shocks.

The chapter is organized as follows. Section 1 summarizes concisely the international works on bad loans concentrated on correlations as a tool to explain their macro dependence.
Section 2 contains the data description; section 3 reviews several MGARCH alternatives together with the typical technical features; section 4 provides empirical results and examines the dynamic correlation between delinquency rates and gross domestic product, the main driver of defaulted loans. It also looks at the time-varying correlations with unemployment, inflation, interest rate, money supply, house price and stock market index rates. Compared to previous literature, we perform two models, the Dynamic Conditional Correlation Model (DCC) and the Markov Switching DCC (MS-DCC), that among econometric approaches seem to be the most fruitful to describe whether correlation changes over the sample period. Preliminary tests are done in advance to support the use of the chosen univariate models.
The last sections conclude with final comments.

1. Review of earlier researches.

Dealing with the credit risk analysis in terms of bad loans, the literature of the sector could not refrain from referring to the study of the correlations existing between the variables to demonstrate the presence or absence of a relationship between factors, indicate the precise degree of relationship and measure the direction signs.
Both the Pearson correlation coefficient and the Spearman rank correlation[1] have been widely adopted as exclusive survey techniques or as complementary analyses of the most sophisticated statistical techniques.
Just to cite few scholars, Socol and Iuga (2010) develop an empirical study based exclusively on the Pearson correlation to demonstrate the inverse annual connection between interest rate and non-performing loans in Romania; Iuga and Lazea (2012) expand the precedent work to consider the influence of the unemployment rate as an independent variable over the period 2008-2011. Murumba (2013) elaborates a fifteen years research study on real GDP and bad loans in the Nigerian banking industry. He determines a positive significance of GDP by converting the R^2coefficient of determination into a t score such that a value less than the critical value in a two-tailed test is adopted as a rule to accept the null of no significant relationship.
The authors that use correlation as a complementary analysis, Fofack (2005), Filip (2015) and Anjom and Karim (2016), Sandada and Kanhukamwe (2016), Adeola and Ikpesu (2017), etc. detect the direction and intensity of the causal association between the amount of bad debts and the other variables through the correlation matrix.
Fofack (2005), for instance, shows a low intensity of the coefficient of correlation between inflation and credit quality, a negative association with the GDP variable across all state and private banks considered in the African countries, a non-homogenous linkage throughout the sample between exchange rate and non-performing loans.
Filip (2015) finds a strong positive correlation between bad loans and their amount during the previous period and GDP, while the linkage with inflation and unemployment result negative and less important. Similarly, Anjom and Karim (2016) support the studies about the positive relations with GDP.
Sandada and Kanhukamwe (2016) gather the degree of correlation among credit risk and macro, industry and bank-specific factors in Zimbabwe through both the non-parametric Spearman correlation rho and regression analysis, observing a statistically positive significant association between credit and macroeconomic performance.
Proceeding, Mondal (2016) questions the causes of the downfall of loans carrying out Pearson correlation, Granger causality test, and regression analysis.
Durafe and Singh (2016) base on the Pearson coefficient the analysis of the pro-cyclical behavior of the bad loans of the Indian public sector banks group, confirming their negative and significant correlation with GDP.
Christodoulou-Volos and Hadjixenophontos (2017) employ correlation to pre-test the presence of high correlation between dependent and independent variables, verify the hypothesis that change in GDP impact non-performing loans with a certain time lag of a quarter, and later create different models for each regressor considered as a predictor of bad loans of Cypriot commercial banks.

[1] Generally, while the Pearson explain the linear relationships of normally distributed variables, the Spearman benchmarks the monotonic linkage between ordinal variables.

2. Data Description and preliminary tests.

Against this theoretical setting, in this section, we introduce the empirical part of the work, presenting the dataset and the testable hypotheses. As in chapter 2, we analyze credit risk on the basis of one of the major proxies of bank asset quality, loan delinquency rate. Quarterly data about USA series are continued to be used and preferred for the easy availability, completeness and wide range of credit quality data with respect to other countries. However, we consider the aggregate credit quality series and we adjust the sample period a little by including all observations from 1985 q2 to 2017 q4, for a total of 131 observations for each series. The usage of date up to the end of 2017 allows the work to result as a rigorous and exhaustive analysis.
The quarterly macrographs are reported in figure 3-1.

Macroeconomic variables

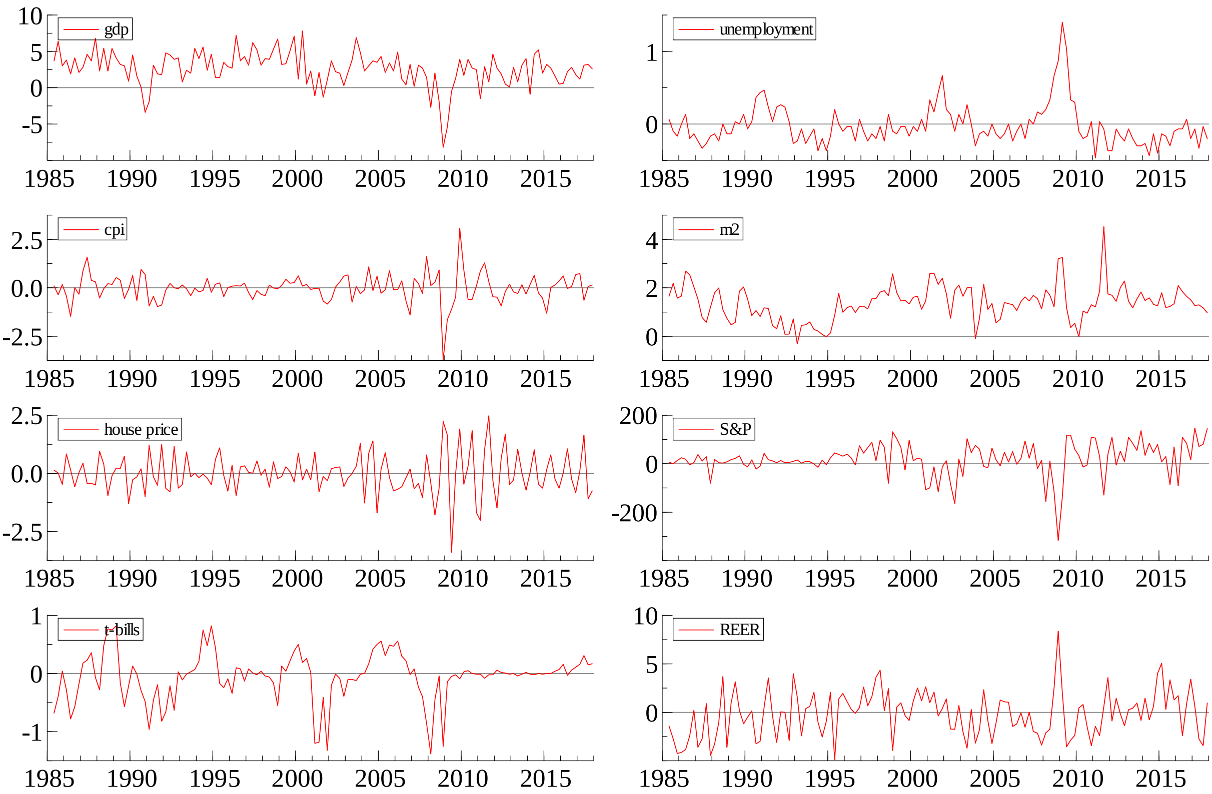


Figure 3-1 Macro dynamics, April 1985- December 2017

The visual inspection of the series suggests a stationary dynamic in the process evolution confirmed by the formal unit root tests of Dickey-Fuller and Philips-Perron that reject the hypothesis of non-stationary at the 95% confidence level. Thus, such data in the sample are found to be I(0).

Focusing on such regressors, in line with the broad object highlighted above, we formulate two research questions for investigation purposes:

- **Q1: Are contemporaneous and lagged correlations levels constant over time?**
- **Q2: Does the dynamic conditional correlation between delinquency rates and macro variables contain asymmetries?**

Both hypotheses are built on the existing theories and past empirical studies that have been conducted on the determinants of bank's delinquency rates.
The first hypothesis postulates the time-varying nature of the correlation between delinquency rates and macro variables, with the aim of measuring the trends in correlations to understand how the changes in default rates over the previous 30 years have contributed to financial instability.
Question 2 investigates state dependent, time-varying transition probabilities within regimes. Specifically, it addresses the potential presence of a regime switch between two unconditional correlation levels through the performance of an MS-DCC, or better, it considers whether correlation dynamics are subject to changes in regime and characterized by speedy and abrupt changes driven by a Markov chain.

3. Methodology

Considering the nature of the study, a quantitative approach is adopted to gain the research objects and a richer knowledge about the problem.
Correlation analysis is the basic tool to quantify interdependence and the degree of synchronization between economic variables but, unfortunately, the simple static method, being a pointwise local analysis, fails to capture any dynamics in the co-movement, showing little about the historical links.
As each economic and financial variable may tend to exhibit a time-varying behavior, or simply depend on varying external conditions and point changes, such as shocks and crisis, in the same way, the co-movement between variables may not be constant but change dramatically over time and the assumption of constant correlation not conceivable and credible.
Motivated by the ambition to analyze the evolution of time-varying correlation between bad loans in terms of total delinquency rate (henceforth Delq) and macro factors, a suitable methodology consists of modeling the time-varying characteristics of the correlation matrix.
The DCC-GARCH (Engle, 2002) is the method that provides a means for adequately assessing continuous changes in correlations between variables. Indeed, accounting for the dynamic evolution of the relationship between variables at each point in time, it is an appropriate and reasonable tool to use in such kind of researches for avoiding the drawbacks of the simple correlation analysis.
Moreover, it will allow studying whether and how correlations vary through time, to investigate the correlation behavior during specific periods of interest such those of financial crisis and to obtain all possible pairwise conditional correlation coefficients, highlighting the true nature.
The acronym DCC-GARCH stands for Dynamic Conditional Correlation Generalized Autoregressive Conditional Heteroscedasticity.
It is a generalization of univariate ARCH/GARCH, the Multivariate GARCH models (MGARCH, Multivariate Generalized Autoregressive Conditional Heteroscedasticity), that is a family of nonlinear models generally used in portfolio selection to model co-movements of financial assets' returns and risks.
Specifically, this technique depicts the current volatility of one-time series as influenced not only by its own past innovation but also by past innovations to volatilities of other time series.
Indeed, in reality, most relevant concerns and applications are multivariate[2], and although the univariate ARCH/GARCH frameworks developed by Engle (1982) and Bollerslev (1986) are very efficacious in describing univariate time series, they are incapable of taking into account the possibility of interaction between more time series and the volatility spillover effect. Univariate expansions represent the appropriate solution to capture these characteristics and for modeling and forecasting the variable of interest so much so that they became the most popular approach in financial econometrics.

[2] For example, in analyzing indexes, portfolios, markets, volatility, portfolio optimization problem, hedging, option pricing, analysis of contagion, volatility spillovers, risk assessment, etc.

Citing Silvennoinen and Teräsvirta (2005) “modeling individual time series separately is thus an insufficient method as it leaves out information about co-movements and interactions between the instruments of interest” (Silvennoinen and Teräsvirta, 2005).
Therefore, limitations of univariate models motivated the research of multivariate volatility models. In addition, the analysis of financial interdependence has gained advantages from the development and the extensions of multivariate volatility models.
Formally, MGARCH is vector volatility model whose basic framework requires the estimation of n different mean and corresponding variance and covariance equations. For a $N \times 1$ dimensional process $\{y_t\}$ and a vector of innovations $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{1N})$, y_t can be written as

$$y_t = \mu_t + \epsilon_t, \tag{1}$$

where $\mu_t = E[y_t|F_{t-1}]$ is the conditional expectation of y_t provided the information set F_{t-1} generated by the observed series up to and including $t-1$, and $\varepsilon_t = H_t^{1/2} z_t$ where z_t is a $N \times 1$ sequence of iid variables with mean zero and identity covariance matrix I_N. $H_t^{1/2}$is a $N \times N$ positive definite covariance matrix whereby

$$H_t = V(\epsilon_t|F_{t-1}) = V(y_t|F_{t-1}) = \begin{pmatrix} \sigma_{11,t} & \cdots & \sigma_{1N,t} \\ \vdots & \ddots & \vdots \\ \sigma_{N1,t} & \cdots & \sigma_{NN,t} \end{pmatrix} \tag{2}$$

the diagonal element $\sigma_{ii,t}$ is the variance of the i^{th} element and $\sigma_{ij,t}$ is the covariance between ε_{it} and ε_{jt}.
However, as the general multivariate GARCH (MGARCH) is considered unfeasible for most problems, different specifications have been developed. In particular, it involves a larger number of unknown coefficients and it is very difficult to guarantee the positive definiteness of H_t.
The alternative models aim at achieving the best compromise between parsimony and flexibility to describe a wide variety of dynamics in the data and covariances.
Bauwens, Laurent and Rombouts (2006) classify MGARCH models into three categories that space from slightly easy configurations to ones that are more complex.
Models mainly differ in the decision of how modeling and specifying the multivariate time series, modeling the conditional variance-covariance matrix directly and modeling the correlation between the time series indirectly.
In general, we have:

a) direct generalizations of the univariate GARCH model: VEH (Bollerslev, Engle and Wooldridge, 1988), BEKK (Engle and Kroner, 1995), Risk-Metrics and factor models;
b) linear combinations of univariate GARCH models: generalized orthogonal models (Van Der Weide, 2002) and latent factor models (Engle, Ng and Rothschild, 1990);
c) non-linear combinations of univariate GARCH models: Dynamic Conditional Correlation (DCC) models (Engle, 2002).

While VECH and BEKK were successful attempts towards the first direction of multivariate volatility models, DCC is the most recent direction of researches developed in order to address the structural and empirical weaknesses of the previous models; they identify accurately and parsimoniously the dynamic pattern of the correlation. Historically, the large body of literature has fruitfully used MGARCH to investigate market spillover and contagion effects.

3.1. The DCC model: Engle and Tse-Tsui specifications

As stated above, the diagonal VECH and the BEKK model are extensions of the univariate GARCH to determine the dynamics of the covariances, and not those of correlations, that have to be computed from the estimated covariance matrix; it cannot be constructed as a simple generalization of univariate GARCH.
For this purpose, the DCC model is the most widespread approach used in the literature due to its simplicity. The versions of Engle (2002), Tse, and Tsui (2002) are prominent among those that have been proposed.
They are a modification and generalization of Bollerslev (1990)'s baseline constant conditional correlation (CCC) model developed to overcome the shortcomings of the assumption of constant correlation not realistic and too strong in practice[3].
For instance, Engle (2002), at first, considered the daily correlation between Down Jones Industrial Average and NASDAQ; then he examined stocks, bonds, and the currencies correlations between the Deutschmark, the pound and the lira around important historical events like the lunch of Euro in 1999. Tse and Tsui (2002) introduced an MGARCH model with time-varying correlations to prove the inadequacy of the BEKK model adopting three data sets, namely exchange rate data, national stock market price data, and sectoral stock price data.
To be more specific, a CCC model is based on the decomposition of the conditional covariance matrix into conditional standard deviations and correlations. Each conditional variance is modeled by separate univariate GARCH(1,1)[4] and the conditional covariance are proportional to the product of conditional standard deviations.
To have an idea of the model setup, let be $h_{ij,t}$ the covariance between two series i and j to be modeled and $h_{ii,t}$ the conditional variance modeled by some univariate GARCH, under the assumption of constant correlation ρ_{ij} between $i$and j, we have

$$h_{ii,t} = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{ii,t-1} \ \ \forall i \neq j \tag{3}$$

$$\rho_{ij} = \frac{h_{ij,t}}{\sqrt{h_{ii,t} h_{jj,t}}} \rightarrow \ h_{ij,t} = \rho_{ij}\sqrt{h_{ii,t} h_{jj,t}} \tag{4}$$

In matrix terms, the model can be written as:

$$H_t = D_t R D_t \tag{5}$$

where $D_t = diag\{\sqrt{h_{ii,t}}\} = diag(h_{1,t}^{1/2}, h_{2,t}^{1/2}, \dots, h_{n,t}^{1/2})$ is a $n \times n$ diagonal matrix of standard deviations from any univariate GARCH models with $\sqrt{h_{ii,t}}$ being the i^{th} position on the diagonal; R denotes a positive definite $n \times n$ matrix that contains the unconditional correlations ρ_{ij}.
In the simple bivariate case, the CCC model is defined as follow where the temporal variations of H_t are clearly determined by only the conditional variances:

$$H_t = \begin{bmatrix} \sqrt{h_{11,t}} & 0 \\ 0 & \sqrt{h_{22,t}} \end{bmatrix} \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix} \begin{bmatrix} \sqrt{h_{11,t}} & 0 \\ 0 & \sqrt{h_{22,t}} \end{bmatrix} \tag{6}$$

---

[3] Just for example, in high volatility periods, higher correlation is usually observed.

[4] Univariate GARCH is the standard specification but one could use any process for conditional variance that satisfies non-negativity constraints and stationarity conditions.

Engle (2002) and Tse and Tsui (2002) generalized the CCC model to make the conditional correlation matrix time-varying too and so that the conditional covariance matrix previously specified in (6) becomes a time-varying counterpart:

$$H_t = D_t R_t D_t \tag{7}$$

Engle's formulation is defined by the following set of equations where an auxiliary variable that follows a GARCH equation is used to compute the correlation dynamics:

$$h_{ii,t} = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{ii,t-1} \tag{8}$$

$$D_t = diag(h_{ii,t}^{\frac{1}{2}} \dots h_{nn,t}^{1/2}) \tag{9}$$

$$R_t = Q^* Q_t Q^* = diag(Q_t)^{-1/2} Q_t diag(Q_t)^{-1/2} \tag{10}$$

$$Q_t = (1-\alpha-\beta)\bar{Q} + \alpha \hat{\varepsilon}_{t-1}\hat{\varepsilon}'_{t-1} + \beta Q_{t-1} \tag{11}$$

where D_t is the diagonal matrix of time varying standard deviations from univariate GARCH; R_t is the time varying correlation matrix; Q_t a $m \times m$ time varying covariance matrix of standardized residuals resulting from univariate GARCH; α and β are two non negative scalar parameters satisfying $0 \leq \alpha + \beta < 1$, $\beta = 0$ if $\alpha = 0$.

$\bar{Q} = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}'_t$ [5] is a $n \times n$ unconditional variance matrix of the vector of residuals $\hat{\varepsilon}_t = (\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_n)$ standardized by their conditional standard deviations $\hat{\varepsilon}_{it} = {}^{\varepsilon_{it}}\big/_{\sqrt{h_{ii,t}}}$ and Q^*a diagonal matrix composed by the square root of the diagonal elements of Q_t.

Consequently, the matrix of conditional correlations is obtained by the elements of Q_t:

$$\rho_{ijt} = \frac{q_{ijt}}{\sqrt{q_{iit} q_{jjt}}} \text{ where } i, j = 1 \dots n;\ i \neq j \tag{12}$$

We will focus on bivariate cases where the correlation coefficient is simply expressed by the equation:

$$\rho_{12t} = \frac{(1-\alpha-\beta)\bar{q}_{12} + \alpha\epsilon_{1,t-1}\alpha\epsilon_{2,t-1} + \beta q_{12,t-1}}{\sqrt{(1-\alpha-\beta)\bar{q}_{11} + \alpha_{1,t-1}^2 + \beta q_{11,t-1}}\sqrt{(1-\alpha-\beta)\bar{q}_{22} + \alpha_{2,t-1}^2 + \beta q_{22,t-1}}} \tag{13}$$

Generally, in summary, the estimation procedure is done in a three-step procedure.
At first, mean equations are estimated for each time series. Secondly, the residuals from the first step are used to estimate univariate GARCH to obtain the standard deviations used to build D_t and to define the standardized

[5] In practical application, the technique of putting $\bar{Q}$ equal to the sample covariance matrix of standardized residuals is called correlation targeting.

residuals vector. In the next step, the standardized residuals are applied to obtain the parameters of the unconditional variance and to estimate the time varying conditional volatility correlation matrix.
Regarding the role of the scalars α and β, we have to stress that these parameters govern the conditional correlation dynamics. Specifically, when $\alpha + \beta = 0$, the model collapses to one with constant conditional correlation matrix. Vice-versa, the estimated correlation becomes increasingly dynamics as soon as the sum of the two parameters approaches one. In particular, α measures the extent to which the previous shocks of innovations (lagged standardized hocks) affects the dynamics of correlation and β captures the decay in the dynamics or persistence in correlations, and as α increases, an increasing role in the estimation of the correlations is played by the lagged squares errors.

Tse and Tsui (2002) essentially formulated the same model but specified the correlation process directly similarly to an ARMA process and as a weighted sum of past correlations. Contrary to the model of Engle, his specification has the disadvantage of requiring an arbitrary number m of observations to compute the sample correlations:

$$R_t = (1 - \theta_1 - \theta_2)R + \theta_1 \Psi_{t-1} + \theta_2 R_{t-1} \tag{14}$$

where θ_1 and θ_2 are positive scalars satisfying $\theta_1 + \theta_2 < 1$, $R_t = [\rho_{ij}]$ is a time varying positive definite correlation matrix with unit diagonal elements and Ψ_{t-1} is the correlation matrix of the K innovations terms:

$$\psi_{ijt} = \frac{\sum_{m=1}^{K} \hat{\varepsilon}_{i,t-m} \hat{\varepsilon}_{j,t-m}}{\left(\sum_{m=1}^{K} \hat{\varepsilon}_{i,t-m}^2\right)\left(\sum_{m=1}^{K} \hat{\varepsilon}_{j,t-m}^2\right)} \tag{15}$$

where $\hat{\varepsilon}_{i,t} = \frac{\varepsilon_{it}}{\sqrt{h_{iit}}}$.

3.2. Further developments in DCC literature

The standard DCC model can be extended using richer parameterizations capable of allowing asymmetry in conditional variances and covariances. Indeed, a major drawback and an unnecessary restriction is the assumption that the dynamics of conditional correlations have the same behavior as they all depend on the scalars α and β; that it is to say, the patterns of correlation between groups of even non homogenous variables are identical. On the contrary, in practice, correlation between and within groups of different categories or geographical areas does not follow the same pattern; it is more likely to assume series-specific evolution and smoothing parameters or asymmetries in correlation dynamic structure.
"When N is large, the restriction of common dynamics gets tighter, but for large N the problem of maintaining tractability also gets harder" (Bauwens, Laurent and Rombouts, 2006).
Many scholars have devoted works to overcome these limitations. Among the principal refinements of the original DCC model formulated to overcome this limitation, the Flexible DCC of Billio, Caporin and Gobbo, (2006), the Smooth Transition Conditional Correlation (STCC-GARCH) model of Silvennoinen and Teräsvirta (2005), the Asymmetric Generalized (AG-DCC) GARCH model of Capiello, Engle, Sheppard (2006) and the Regime Switching Dynamic Correlation (RSDC) GARCH of Pelletier (2006).
While the FDCC model introduces a diagonal block structure and constrains the dynamics of correlations to be equal only among the groups of variables within each block and not for the whole correlation matrix, the AG-DCC introduces asymmetry in the correlations dynamics to capture heterogeneity in the data.
Silvennoinen and Teräsvirta (2005)'s STCC model assumes the conditional correlations to change smoothly over time from one state to another, as a function of a transition variable. Differently, the RSDC asserts a regime switching correlation structure driven by an unobserved state variable where the correlation matrix is constant in each regime but may vary across regimes, and a Markov chain governs the switch.

Specifically, Billio *et at.* focused on daily data of Italian stock indexes divided by subsectors and compared the CCC, DCC and the Flexible DCC models to describe the dynamic correlations, evidencing the superiority of FDCC as the parameters describing the correlations display dissimilarities among macro sectors.
Authors assumed that k groups of variables follow different DCC models introducing an additional vector of parameters c but excluding the unconditional correlation matrix R considered by Engle (2002) in the equation (22) as:

$$\mathrm{Q_t} = cc' + aa' \odot \varepsilon_{t-1}\varepsilon_{t-1}' + bb' \odot Q_{t-1} \tag{16}$$

where c is a k-partitioned vector with a number of elements in each group equal to n_i $(i = 1 \ldots k)$, $\odot$ indicates the Hadamard product and $a_i a_j + b_i b_j < 1$ with $i, j = (1 \ldots k)$ the set of constraint added to avoid explosive patterns:

$$c = \left[c_1 i_{n_1}, c_2 i_{n_2}, \ldots, c_k i_{n_k}\right]' \tag{17}$$

According to Capiello, Engle, Sheppard (2006), the dynamics of the conditional correlation do not account for asymmetric effects, implying that although the model considers the magnitude of past shocks' impact on future conditional volatility and correlation, it does not differentiate between positive and negative shock effects.
Therefore, authors modified the DCC to take into account potential structural breaks in the unconditional correlations among the variables by analyzing the behavior of international equities indices and government bonds of European countries, Australasia and Americas. They documented strong asymmetries in both equities and bonds conditional correlations, with equities stronger responsive than bonds to joint bad news; moreover, they found an increasing linkage in conditional equity correlations among regional groups during financial turmoil. Their proposed correlation equation is given by:

$$Q_t = (\bar{Q} - A'\bar{Q}A - B'\bar{Q}B - G'\bar{Q}G) + A'\varepsilon_{t-1}\varepsilon_{t-1}'A - B'\bar{Q}_{t-1}B - G'\eta_{t-1}\eta_{t-1}'G \tag{18}$$

where A, B and G are $k \times k$ parameter matrices and $\eta_{it} = \mathrm{I}[\varepsilon_{it} < 0] \odot \varepsilon_{it}$ where $\mathrm{I}[\varepsilon_{it} < 0]$ is the indicator function which takes on value 1 if $\varepsilon_{it} < 0$ and 0 otherwise.
The following equations, instead, drive the correlation dynamics in the STCC:

$$R_t = R_1 f_t + R_2(1 - f_t) \tag{19}$$

$$f_t = \left[1 + e^{-\gamma(s_t - c)}\right]^{-1}, \gamma > 0 \tag{20}$$

where R_1 and R_2 are correlation matrices corresponding to the low and high regime, s_t is the transition variable, c locates the transition and γ is the slope of the function or the speed of transition. A crucial aspect is the choice of s_t which could be an exogenous variable and it could be taken as a function of the lagged values of y_t or as the market volatility. Just for example, Silvennoinen and Teräsvirta (2005) applied the model to the daily returns of five frequently traded stocks (Ford, General Motors, Hewlett-Packard, IBM, and Texas Instruments) included in the S&P index and chose the transition variable as the lagged absolute S&P 500 index returns averaged over seven days.
Pelletier (2006) noticed that the correlation matrix could be subject to regime shifts. Formally, the covariance matrix in a RSDC model follows the decomposition into correlations and standard deviations used by Engle (2002) with the correlation matrix given by:

$$R_t = R\lambda(s_t) + I_k(1 - \lambda(s_t),$$

$$s_t \,\epsilon\{h, l\} \tag{21}$$

where s_t is a discrete unobservable variable following a Markov chain and representing the regime at time t (it assumes the low correlation regime if $s_t = l$ and the high correlation regime if $s_t = h$), R is a fixed correlation matrix $k \times k$, I_k is a $k \times k$ identity matrix and $p_{ij} = pr(s_t = j|s_{t-1} = i)$ the probability to switch from one state to the other.
Alternative parameterizations have been developed even for the above DCC extensions models.
For example, in the above STCC and RSDC model, Bauwens and Otranto (2016) inserted the effect of volatility as a major determinant of correlations, introducing the so-called Volatility Dependent Conditional Correlation (VDCC) models. They considered two measures of v_t, the one-step-ahead forecast of the VIX index and the one-step ahead forecast of the regime indicator ξ_t of v_t, and assessed the significance of the volatility impact on correlations in the different models applied to a tri-variate and a 30-variate dataset.
In Markov Switching VDCC, volatility drives the transition probabilities:

$$p_{ii,t} = \frac{\exp\left(\theta_{0,i} + \theta_{1,i} v_t\right)}{1 + \exp\left(\theta_{0,i} + \theta_{1,i} v_t\right)},$$

$$p_{ij,t} = 1 - p_{ii,t}$$

$$i = h, l; i \neq l \tag{22}$$

4. DCC or MS-DCC: which model can better explain variables correlation behavior?

4.1. Preliminary analysis

Against this methodological introduction, our empirical analysis starts with an examination of each series separately to verify the basic requirement for the DCC-GARCH estimation: the specification of each individual conditional mean though the autoregressive moving average ARMA (p, q) process and the presence of conditional heteroscedasticity or ARCH effect.
The univariate models are chosen according to the information criteria.
AIC, BIC, and HQC evaluated for different processes select the same ARMA model for delinquency rate, GDP, unemployment, inflation, house price and stock market index rates.
However, information criteria pick different orders for money supply, interest and exchange rates. The AIC selects the highest among the considered orders, namely ARMA (2, 1), ARMA (3, 3) and ARMA(2,2) for quarterly M2, T-bills and REER rates respectively; the BIC findings, on the contrary, are consistent with the fact that it penalizes additional parameters more rigorously, selecting model (1,0) for M2 and T-bills and model (0,1) for REER. The HQC criterion findings agree with the AIC in the first two cases, while they coincide with BIC in the latter.
Therefore, the BIC should be preferred for its parsimony but, considering our moderate sample size, we choose the AIC criterion known to perform relatively well for small samples (Hurvich and Tsai, 1990).
The following table shows the optimal model orders for each series.

Table 3-1 Optimal filtering orders for univariate model to be used in the DCC framework.

Model selection for ARMA	
	(p, q)
delinquency tot	(3, 0)
GDP	(2, 0)
Unemployment	(3, 2)
CPI	(2, 2)
M2	(2, 1)
House price	(2,0)
S&P	(1, 0)
T-bills	(3, 3)
REER	(2, 2)

After rigorous selection of different models, the appropriate are later tested for heteroscedasticity, and the Engle's Lagrange multiplier (LM) tests[6] is performed on the residuals from the mean equations. Test statistics with the corresponding p-values are presented in Table 3-2 that briefly summarizes findings across the different data possibilities.

Outcomes do not always support the presence of the autoregressive conditionally heteroskedastic effect (ARCH), meaning that residuals are serially uncorrelated. Evidently, it is the case of delinquency rate, GDP, money supply and exchange rate variables, while the ARCH is noticed in the remaining series at a highly significant level of 1%.

Table 3-2 Test for ARCH

ARCH test									
	del tot	**GDP**	**Unempl.**	**CPI**	**M2**	**House price**	**S&P**	**T-bills**	**REER**
LM statistics	6.094	6.051	20.257	40.531	3.149	17.829	16.258	23.619	1.958
p-value	0.192	0.195	0.000	0.000	0.533	0.001	0.003	0.000	0.744

Depending on the presence of heteroskedastic effects, in the next step of the analysis, we distinguish the calculation of the standardized residuals used to estimate DCC parameters depending on whether the series were heteroscedastic or not.

In the first case, we proceed with the determination of the most appropriate GARCH equation adopting, again, traditional AIC and BIC information criteria since a GARCH model can be treated as an ARMA model for squared residuals.

We do not show such outcomes as the simply GARCH (1, 1) results as the right order of conditional volatility model and we, therefore, estimate models using a standard specification[7]. Hence, the information from the mean equations and volatility models are used to estimate the DCC-GARCH parameters.

Conversely, in the case of no ARCH effect, we overcame the "limitation" of homoscedasticity of variables by considering ARIMA residuals without GARCH in the second step of the DCC estimation procedure; it is a constant time series equal to the estimated variance of the homoscedastic residuals.

[6] Testing for ARCH effects requires regressing squared residuals on a constant and q chosen lags. Specifically, each series is, at first, regressed on its own lags in order to retrieve the residuals series. The residuals series is then squared and regressed on its own squared lags. Therefore, the test can be seen as one for autocorrelation in the squared residuals. The null hypothesis is that ARCH effect is not present, or equally, that all q lags of the squared residuals are not significantly different from zero.

[7] "The inclusion of more lags does not significantly help explain the model" (Engle, 2001, p.160).

4.2. Estimated models and empirical results

In this paragraph, we present the empirical findings of the DCC model applied to our variables along with discussion and interpretation.

However, as pointed by Pelletier (2006), a limitation of the DCC approach is that the dynamics of the conditional correlation do not represent possible asymmetric effects due to the presence of regimes. This implies that although the model accounts for the magnitude of past shocks' impact on future conditional correlation, it does not differentiate between the regime of high or low correlation.

Consequently, in this section, the traditional DCC is also compared against the MSDCC with two different regimes to see which model delineates the estimated correlation process more in line with reality. Substantially, both of them are time-varying approaches with the difference that MSDCC checks whether a discrete level shift may exist in the dynamic conditional correlation process such that correlations stay constant within a regime but switch across them according to the first order Markov chain.

Accordingly, we start by considering the DCC and to select the best model for each pair of series we have run the following procedure:

1. perform the DCC for all pairs of variables to get evidence of dynamic correlation;
2. perform a formal test to check and confirm the robustness of our results;
3. perform the MS-DCC just for variables we can reject the null of constant correlation in favor of a dynamic structure;
4. compare the selected models with a loss criterion.

Table 3-3 summarizes the parameter estimation outcomes specified in equation (11) for the DCC. There are reported the estimates of the parameters α and β produced when the correlation targeting matrix is the sample correlation matrix.

The graphs of the estimated quarterly time-varying correlations over the sample period are plotted in figures 3-2, 3-3, 3-4.

Results include the lagged forms. In fact, since loan portfolio can be influenced by macro environment after a certain period of time, models are also fitted in delayed forms between delinquency rate and lagged macroeconomic indicators up to lag 1.

The pairwise estimated of contemporaneous coefficients are all positive to ensure positive definite unconditional correlations. In addition, in all cases, the condition $\alpha + \beta < 1$ is satisfied. As mentioned, a value of the sum close to one indicates highly persistence in the dynamics of conditional correlation, while the closer the sum to zero, the quicker the conditional correlation converges to its long run value of the unconditional correlation.

Consequently, values suggest that pairs with the lowest persistence in our DCC are "delq-unemployment" (0.198), "delq-M2" (0.047), "delq-house price" (0.096) and "delq-S&P" (0.040).

Specifically, these relationships exhibit minimal time-varying movements that depend just on the sensitivity of correlations to the arrivals of news $\hat{\varepsilon}_{t-1}$. Hence, correlation dynamics can change only for large external shocks.

On the contrary, "delq-GDP", "delq-CPI", "delq-T-bills" and "delq-REER" seem to co-move more powerfully. However, the correlation between delinquency and inflation and exchange rate and the interest rate is indeed constant (as proven graphically) as they depend just on the news parameter α of the DCC model that is equal to zero and hence correlations are insensitive to shocks.

Consequently, just the correlation between delinquency and GDP is indeed time varying because just this bivariate combination has highly significant parameters α and β that are greater than zero.

Moreover, the correlation between "delq-GDP" has the highest value of the conditional correlation mean value, and it further confirms extreme movements.

Quite the opposite, the correlation of delinquency rate with CPI, T-bills and REER exhibit the lowest ones.

The same reasoning applies to lagged variables. There are not remarkable difference for GDP, money supply, house price and interest rates. On the contrary, parameters estimates differ when considering the DCC between delinquency rate and lagged unemployment, lagged inflation, lagged stock market and lagged exchange rates. They highlight significant varying conditional correlation.

Table 3-3 Models outcomes.
The table summarizes the estimated coefficients of the basic DCC model in a bivariate approach. Correlation is between total delinquency rate and each macro variable. UC is the unconditional correlation and values in parenthesis refer to the Standard Errors

Parameter estimates DCC	α	β	UC
Delq-GDP	0.027 (0.028)	0.805 (0.099)	-0.390 (0.117)
Delq-GDP_1	0.011 (0.041)	0.786 (0.254)	-0.204 (0.113)
Delq-unemploy	0.198 (0.162)	0.000 (0.021)	0.177 (0.156)
Delq-unemploy_1	0.029 (0.045)	0.692 (0.331)	0.264 (0.131)
Delq-CPI	0.000 (0.000)	0.000 (0.000)	
Delq-CPI_1	0.068 (0.096)	0.636 (0.199)	0.099 (0.182)
Delq-M2	0.047 (0.042)	0.000 (0.025)	0.073 (0.109)
Delq-M2_1	0.000 (0.000)	0.000 (0.000)	
Delq-House price	0.096 (0.079)	0.000 (0.069)	-0.087 (0.113)
Delq-House price_1	0.000 (0.014)	0.000 (0.000)	
Delq-S&P	0.040 (0.028)	0.000 (0.015)	-0.200 (0.119)
Delq-S&P_1	0.210 (0.069)	0.466 (0.105)	-0.145 (0.088)
Delq-T-bills	0.000 (0.000)	0.000 (0.000)	
Delq-T-bills_1	0.000 (0.000)	0.000 (0.000)	
Delq-REER	0.000 (0.000)	0.000 (0.000)	
Delq-REER_1	0.170 (0.099)	0.025 (0.255)	-0.012 (0.258)

Nevertheless, to validate our results we formally test whether the dynamic correlation is appropriate. As the constant correlations model can be interpreted as a restricted version of the standard DCC where the restrictions imposed are $\alpha = 0$ and $\beta = 0$, we run a Wald test of the null hypothesis of equivalence or indistinguishability of all coefficients of the model to confirm the results:

$$H_0: \alpha = \beta = 0$$

$$H_1: \text{almost one of } \alpha \text{ and } \beta \text{ is different from } 0$$

Results shown in Table 3-4 reveal that the hypothesis of constant correlation holds in the US economy for all linkages except the GDP, implying that a dynamic version can be applied to quantify its conditional correlation and to account for asymmetries.

Specifically, the test rejects the hypothesis that the parameters are both equal to zero implying that the dynamic correlation specification is appropriate, while the assumption is incorrect for the remaining variables; we have to accept the hypothesis of constant correlation for connections with unemployment, money supply, house price index, S&P.

The empty spaces in correspondence of correlation between delinquency and interest rate, delinquency rate and inflation, delinquency and exchange rate are because the test is not performed. Indeed, these correlations are constant as previously explained.

Therefore, regarding all macro variables except GDP, we can globally claim that results from the study do not show a significant time-variability in conditional correlation. Results of dynamism also apply to correlations with lagged values of GDP, of unemployment and of inflation and S&P; they do not regard the correlation with REER.

These first findings are already a novelty and can be viewed as surprising in our research areas as no previous studies tested and relaxed the assumption of constant correlations reaching conclusions of time variability in the connection between macro-environment and loan portfolio quality, and capturing dynamic behavior in reaction to news and innovations.

Table 3-4 Table reports results of the Wald test performed between total delinquency rate variable and each macro variable considered with and without lag. Values specify the p-values whereas empty spaces are indicative of the fact that for the specific pair of variables the test is not performed given the evident constant dynamic.

Wald test under the null of constant correlation	**Unlagged**	**Lag_1**
Delq-GDP	0.000	0.007
Delq-Unemployment	0.463	0.029
Delq-CPI		0.006
Delq-M2	0.526	
Delq-House price	0.450	
Delq-S&P	0.306	0.000
Delq-T-bills		
Delq-REER		0.217

Established that correlations are globally constant, we therefore proceed with the estimation of the MS-DCC and its comparison with the DCC just for contemporaneous and lagged GDP, lagged unemployment, inflation, and S&P.

Full results of the estimated coefficients of the MS-DCC model in a bivariate approach are shown in table 3-5, while the comparative table 3-6 reports in the detail model selection results.

Table 3-5 Models outcomes of MS-DCC

Parameter estimates MS-DCC Correlation Regimes					
	GDP	**GDP_1**	**Unemployment_1**	**CPI_1**	**S&P_1**
P11	0.953	1.000	0.907	0.968	0.988
SE	0.089	0.000	0.123	0.018	0.086
P22	0.921	0.431	0.916	0.984	0.000
SE	0.119	0.006	0.109	0.020	0.000
$\lambda 1$	1.778	1.125	3.217	21.51	1.017
SE	0.317	0.070	0.315	0.008	0.319
$\lambda 2$	0.337	0.975	0.029	0.000	0.999
SE	0.394	0.021	0.362	0.000	0.002
UC	-0.310	-0.191	0.168	0.029	-0.142
SE	0.078	0.066	0.051	0.003	0.035
UC= unconditional Correlation					

While hypothesis testing based on the likelihood function works for nested models, a non-nested model selection like the case of DCC and MSDCC is carried out through statistical information criteria.

To determine whether the DCC is superior to the MS-DCC model in the GDP case, we refer to the maximized log-likelihood value and the AIC. The criterion is chosen referring to Psaradakis and Spagnolo (2003) that commented the properties of the selection procedures in choosing the number of states in Markov Switching Autoregressive process and expressed in favor of the AIC, seemed to have the best performance. Moreover, we calculate the Schwarz Bayesian information and the Hannan–Quinn information criterion to give robustness ate the results. Doing so, DCC significantly outperforms the other in terms of lower information criteria.

Therefore, DCC is more suitable for our series.

Model	Log-Likd	AIC	BIC	HQC
DCC –GDP	-101.423	**208.847***	**217.472***	**212.351***
MSDCC –GDP	-100.287	210.574	224.950	216.416
DCC-GDP_1	-106.010	**218.021***	**226.623***	**221.516***
MSDCC –GDP_1	-106.037	222.074	236.412	227.900
DCC-Unemployment_1	-104.907	**215.814***	**224.417***	**219.310***
MSDCC-unemployment_1	-103.520	217.040	231.378	222.866
DCC-CPI_1	-108.139	**222.278***	**230.881***	**225.774***
MSDCC-CPI_1	-106.411	222.822	237.160	228.648
DCC-S&P_1	-113.970	**233.940***	**242.543***	**237.436***
MSDCC-S&P_1	-115.546	241.092	255.430	246.918

Table 3-6 Log-likelihood value and information criteria to compare simple DCC against MS-DCC. Bold values indicate the better performance

5. Main findings and discussion

First, as indicated by model parameters and displayed in Figure 3-2, the dynamics of correlations estimated by the simple DCC between delinquency rate and GDP are always negative but do not remain constant over time.

On average, the degree of the relation between delinquency and GDP appear to have oscillated mildly, mainly in the range between -0.3 and -0.560 with significant spikes reached during or shortly after two of the major USA episodes such as the escalation of the 1990 recession and the burst of the worst global financial and economic crisis, that of 2008[8]. Both events acted as a structural break in correlations.
Thus, the correlation has strengthened considerably increasing in absolute values; it just stabilized and weakened before and after the economic and financial crisis, as soon as the economy improves.
Results corroborate the idea that when growth rate decays, household cash flows reduce and therefore expenditures priorities go to consumptions rather than on meeting debt obligations. Quite the reverse, the positive economic environment is associated with the better capacity of respecting debts.
A detailed numerical analysis of data can enable us to understand the correlation further. For example, from October 1990 to March 1991, correlation rose or better became more negative with values from -0.369 to -0.458. The same holds in the post 2008s, with values that go from the -0.371 of October 2008 to -0560 of June 2009. As it is evident, such periods mostly coincide with the NBER recession periods (the shaded areas in the graph). There is no coincidence only with the 2001-2002 NBER recession; we found a relative modest implication of the dot bubble on the rise of correlation among GDP and problem loans. Indeed, in 2000s problem loans were concentrated mainly in Internet-based industry such that, probably, industry-specific factors and not macro trends were the main drivers of higher default rates.
The dependence on time is also recognized when considering the impact of an increase in real GDP a quarter earlier on delinquency rates. In such case, the behavior is identical to the previous one except for the lower intensity in the absolute value of the correlation.
Indeed, despite the lowest values are again identified in correspondence of the first and third recession, the different mean levels for the conditional correlation confirm that the contemporaneous DCC estimates vary more during the periods of financial crisis. A decrease of one percentage point of GDP leads to an immediate average increase of about 0.46 in the delinquency ratio and to an increase of about 0.22 during the following quarter.
Figure 3-2 is specifically depicted to provide more insights into the rather heterogeneously behavior over time and with different time lags. The series in the blue line indicates the unlagged bivariate relationship, while the red line shows the time path of correlation between total delinquency rate and GDP considered with lag of 1. Evidently, both exhibit negative values but the average correlations reveal swings that are even more considerable for unlagged terms.
Findings are in line with Salas and Saurina (2002), Quagliariello (2006) and corroborate the intuition of a quick transmission of economic shocks to banks bad debts and balance sheets.

[8] Both 1990s and 2008s recessions were based on financial crises (remember the 1989 Saving and Loans Crisis) but the former was also related to the invasion of Kuwait at the hands of Iraq, which caused a peak in the price of oil in 1990 and caused a decline in sales in the manufacturing sector. The latter crisis and subsequent recession, on the contrary, had its primary causes in the US housing market. Later, it rapidly spread to the US financial sector, and then to the global financial sector, bringing down investment and commercial banks, insurance companies and numerous companies built upon credit.

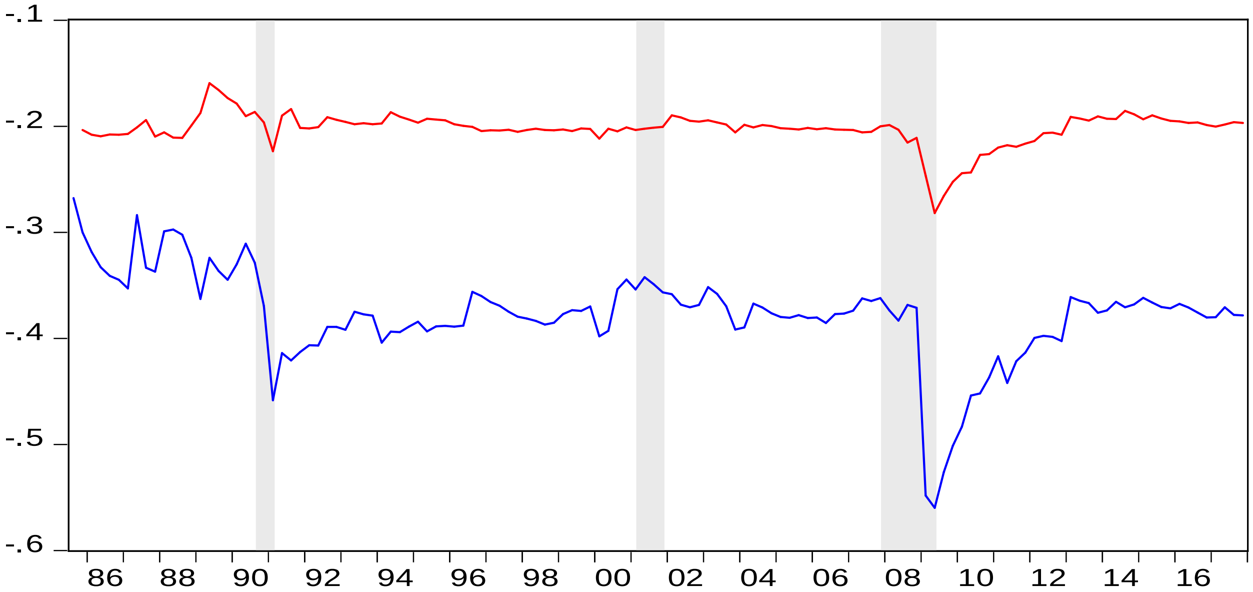


Figure 3-2 Time-varying behavior of the linkage "delinquency-GDP"

However, just the correlation between delinquency and GDP is significantly time-varying meaning that the real impact of times of crisis such as loss in the output growth has a preponderating effect on the levels of bad debts.

Indeed, when considering the process described by the other contemporaneous DCC models, they are immediately mean-reverting: after the occurrence of a shock, correlations return to the long-run unconditional level. Stated differently, findings demonstrate that while "delq-GDP" is more unstable and erratic over time even, the other correlations may change substantially but they recover the mean value in a very short period of time.

Considering the specific correlation equations, the persistence of constant correlation over time is evident from the parameters of the DCC as the sum of the scalars is closer to zero.

The connection unemployment – delinquency clearly highlights such trend.

Despite few much higher and lower points in the sample, conditional correlation is constant over time and the strength of the linkage averages principally between 0.100 and 0.300, with the expected positive sign congruous with the consumption smoothing behavior.

The highest values (see figure 3-3) fall between the end of the 1990 and the beginning of 1991 (when the economy collapsed into recession: 0.504; 0.523) and within two quarters from the start of 2009 (0.690; 0.708). In the post-2007, the financial crisis hit the USA labor market performance more notable than other industrialized countries. The intensity of the massive recessions, insolvencies of firms and considerable cost cuts measure like job losses spread out across a variety of industries and made it more difficult for redundant to find new job opportunities. Consequently, the unemployment rate started to increase, assuming an ascending parabola in April 2009, and in January 2010 it even doubled reaching the level of 10.4 % from the 4.6% in July 2007[9].

The first lowest point on the graph of the connection occurs in the second quarter of 1987 (-0.064) cohering with the fact that in April 1987 the unemployment rate moderately started to decline from its peak of 7 percent[10] as a consequence of the development of the manufacturing sector and more intensive use of labor resources. Subsequent points, continuing along the line of events, are dated January 2002, January 2003, January and April 2011 (-0.049,-0,032, -0.128; -0,216). Explanations of such displayed negative correlations can be different but always tentative.

[9] Source: U.S. Bureau of Labor Statistics

[10] Source: U.S. Bureau of Labor Statistics

Most likely, they are related to the unemployment benefit mechanisms that, despite enduring weak labor market conditions, reduced the unemployment indicator, with possible misrepresentation of the real job force position.
Specifically, these insurances are normally accorded only to those registering as unemployed. In the period following the recession, unemployment emergency forced government policy to finance benefit programs. They *per se*, at first, slowed and dis-incentivized the transition into the labor force, contributing to a raised unemployment and, in general, to the rise of the individual reservation wage[11] (moral hazard behavior of individuals that prefer to take time before finding and even simply searching a new job). Later, close to the time of benefit expiration[12], the cancellation of individuals from the work search registers had an opposite temporary effect on the recorded unemployed amount that seemed to be lower although much higher, demonstrating an incorrect correlation.

Figure 3-3 Correlation "delq-unemployment", lag0

The same reasoning of the constant trend, except the influence of specific shocks that slightly alter the movement (hence β=0), applies to the conditional correlation between delinquency rate and changes in house price index, delinquency rate and changes in S&P, delinquency rate and money supply.
We observe a short-term negative and significant effect of house prices on delinquency rate, with increments in absolute terms of correlations around the time of banking panics in the US, most markedly around the panic of the 1990s, however also during the housing crash of 2007, and to a certain extent in 1988q1 and 1989q2.
After the US subprime crisis in August 2007, house prices and the confidence of investors and consumers have progressively waned and, in particular, during 2008. Default rates, however, increased regularly throughout the period analyzed.
As figure 3-4 illustrates, correlation fluctuates around -0.08 prior to 2007. By the fourth quarter, it rises to -0.20. The increase continues in 2009 and correlation reaches -0.23 in 2010q3.
However, it surprising seems that despite in 2008 the pace of house decline was worse than the early 1990s; correlation grew with an intensity markedly higher during 1988-1991 periods, suggesting a significant overvaluation of house prices above other factors in the last delinquency rates trend.

[11] See Mortensen (1977) about the effect of benefit on the exit rate from unemployment. According to the author, payment benefits reduce unemployed people's job search effort or opportunity cost of search, and increase unemployment duration.
[12] Consider that insurances programs lasted from a minimum standard period of 26 weeks to a max of 77/99 weeks. During the Great Recession observed from the late 2000s to the early 2010s, they were gradually extended with the exclusion from the maximum duration of un-eligible unemployed.

Indeed, S&P, as for instance, does not abruptly affect correlation during the first economic downturn but it starts to heavily impact delinquency loans since 2008q2 (correlation -0.23) to even doubled over the next three quarter (correlation peak of 0.44).
So outcomes indicate that, although the influence of both variables may have contributed to the last high delinquency rates and financial instability, it was stock prices drop that, producing a liquidity crisis that froze up interbank lending markets, exerted a greater erosion on the Americans wealth.
Analogously, the interaction with money supply shows a clear constant positive pattern interrupted by a rapid explosion in correspondence of the first quarter of 2009 (the correlation powers 0.50). This extreme point is not an absurd fact. The strength in correlation corresponds to a period (from the last quarter of 2008 till March 2009) of a significant increase in the total stock of money available in the economy that resulted from the Federal Reserve (FED) decision to resort to quantitative easing (QE) measures in order to stimulate economic growth and increase liquidity[13].
On the contrary, regarding inflation, interest rate and exchange rates, this time, plots are totally flat. They approximate a straight line with no minimum evidence of erratic and unstable behavior indicating that only sufficiently large events in the macro variables can cause structural breaks in the series of correlations and impose correlation dynamics.
Figure 3-4 depicts the interactive linkages.
Regarding interest rates, the documented negative correlation could appear contrary to the broader literature about a positive sign. In reality, it is consistent with the fact that higher interest rates on T-bills rates are a signal of the government and banks lending policy so that the greater the rates the higher the opportunity costs of lending loans due to the lower affordability of debtors. Hence, higher rates minimize loan losses, discouraging risky and less profitable investments, and maximizes efficiency: banks realize that holding deposits instead of loans are less risky. Banks, in other words, in the short run prefer to restrict the access to credit and to alter the credit allocation versus T-bills deposits as perceived as lower risk.

---

[13] FED or better the FOMC (Federal Open Market Committee) started to buy large-scale government bonds and mortgage backed securities that expanded FED balance sheet but increased the monetary base. Such decision influenced the economy through changes in interest rates and the depreciation of the value of money due to the grater stock in circulation. Over time, all led to a rise in the prices of financial assets (because of the rise in the demand) and a consequently reduction of the yields, accompanied with a reduction of the prices and interest rates of the other obligations (such as mortgages). Moreover, it stimulated borrowing, spending, lending activity as banks became more liquid without government bonds and contributed to a reduction of unemployment (with decreasing interest rates, the cost to businesses decreases), an increase in GDP and a pre-crisis levels of Wall Street courses.

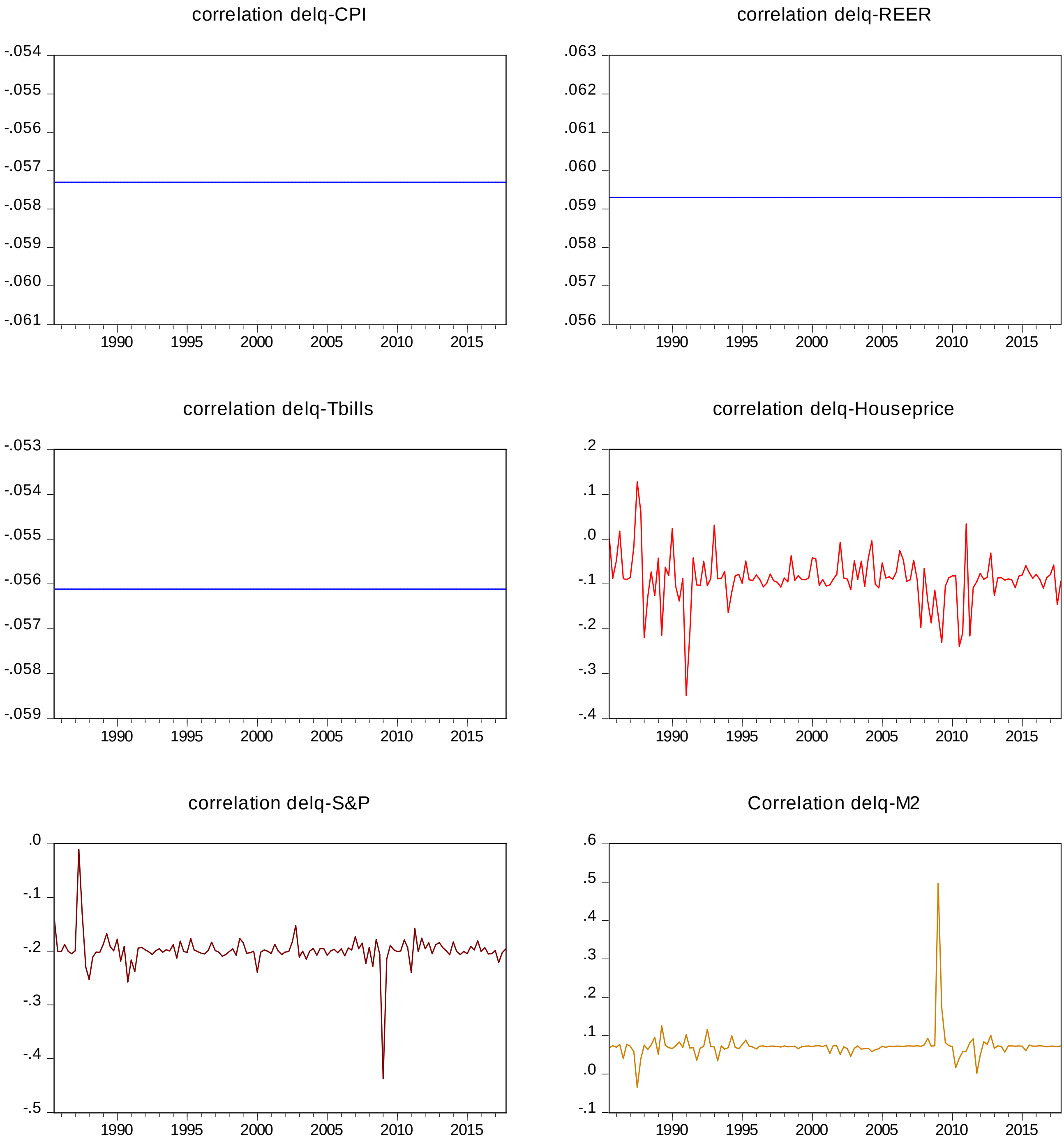


Figure 3-4 DCC with contemporaneous variables

By examining the evolution of lagged correlations, we note higher dynamic tendencies.
There exist parcels of high and low correlations in four couples “delq-macro factors” out eight that justify the time-varying nature. Again, pairwise conditional correlation coefficients seem to be more volatile with abrupt jumps immediately before or during the early 1990s and after 2007 (September 2008-December 2010).
Variables are more dynamic interdependent meaning that delinquency rate is not immediate sensitive to shocks in the economic aggregate and, every disequilibrium in macro determinants causes more considerable co-movements in delinquency rate variable with some lags.
A demonstrative case regards the link “delq-cpi” that passes from a totally constant pattern (when considering the contemporaneous impact of CPI) to a significant fluctuation during the considered period of time (when focusing on CPI lagged value).

Overall, by graphically examining the evolution of these correlations (figure 3-5), first, we observe that response to stock market shocks and, hence, the magnitude of these changes appears to be particularly important. Indeed, although periods of high negative correlations between delq and macro factors seem to coincide, in 1991 q1 and 2009 q2, the correlation among "delq and S&P growth rate" is higher (in absolute terms) than the correlation between any of the other series[14].

Outcomes are surprising because, when lagging the economic indicators, we still expected a greater influence of GDP or unemployment rather than the stock market. Consequently, such tight direct relationship has a number of implications. For instance, it may be argued that, within the periods under consideration, the deterioration of the real economy is substantial to lead a contemporaneous increase in default rate. However, in a relatively long period of just a subsequent quarter, bad debts increase more in front of the reduction of individual's asset-based wealth rather than in front of income reduction. Any news or information about stock market, the impossibility to cash stocks to make loan payments, effects on the value of pensions[15] lower consumer confidence as they are perceived as changes in the value of the financial wealth and lead to a default on loans thereby causing their negative relation with stock prices.

Proceeding, the pairwise correlation "delq-unemployment" maintains the positive sign displayed in the contemporaneous model. The higher mean value of the lagged model confirms its more decisive dynamic nature. Moreover, findings show that delinquency rate suffers more the effect of the increase in unemployment happened in the previous quarter than the increase in the same quarter; upsurges in lagged unemployment lead to higher delinquency rates suggesting that unemployment rate hits more the future purchasing power.

Just during the major periods of recession, the primary impact of unemployment on delinquencies is contemporaneous.

For instance, in both models, delq and unemployment closely track each other but changes in labor market and employment status interact considerably with the ability to pay during recessions such that rising unemployment pushes levels of delinquencies even higher. However, results of the lagged DCC suggests that, in 1991q1, a one percentage point increase in the unemployment rate in time $t-1$ leads to a 0.38 percent increase in the level of delinquencies in time t, versus the 0.52 percent registered with the contemporaneous model. Similarly, a one-percentage point increase in the unemployment rate at time $t-1$ leads to an increase of 0.44% in the level of delinquencies in 2009q2, versus 0.71% registered with the contemporaneous model.

Going on, as mentioned, the relation "delq-cpi" is indeed time varying just in the lagged case. It suggests that the impact of inflation on the standard of living and incomes is a day to day process that immediately leads to cutting the extraordinary and/or superfluous expenses but it takes time before people reach the point where they cannot pay all their expenses and loans payments.

By going to the data and into the graph (number 3-5, Appendix B), delinquency adjustments to CPI movements do not maintain the same positive sign but correlation ranges from the maximum value of 0.402 (1991q1) and two minimum values of -0.315 and -0.121 (1987q3 and 2009q2).

We find that before 1989 and in 2009q2, variables have a negative relationship. After, delinquency loans increase with inflation.

It is well known the effect of inflation on credit risk is complicated. Theoretically, it seems plausible the negative effect.

Indeed, inflation, allowing borrowers to repay lenders with an amount of money that is the same but it is worthless in terms of purchasing power, advantages borrowers and generates a wealth redistribution effects that reduce the cost of finance and thus credit risk.

[14] In 2009 q1, the value of the interaction of bad debts with S&P is -0.766 against a value of -0.159 of the interaction with GDP, and -0.195 with house price index

[15] Consider that pension funds invest in stock market and any collapse, causing considerable losses suffering by private intermediaries, can affect the retirement payout.

However, empirical evidence sometimes shows the opposite results. First, a rise in inflation positively affects the loan default rate and hampers repayment ability since, normally, wages are constant and adjust less quickly than the overall increase in prices.
Second, banks are forced to increase the interest charged. Indeed, the raised loans demand is seen by banks as an opportunity for doing business and making more money from that they have but also to cover the increased expenses (of labor costs, supplies etc.) and to compensate the credit risk caused by the reduced purchasing power of money. Therefore, customers have to pay higher interest payments, are discouraged from borrowing money, their debt serving ability is seriously affected and delinquency loans grow. Moreover, normally wages are constant and adjust less quickly than the overall increase in prices.
Whatever the effect of inflation, there are latent forces that came into play and gradually erode the buying power of the money.
Prices do not just rise on their own. For instance, oil prices are a significant concern because of the long run pass-through effect from import prices to consumption prices.
It is what happened in 1987 q3. Inflation increased[16] reflecting the improvement in economic growth and surging import prices (especially oil price, the major input in the economy used in all critical activities such as transportation), in turn, driven by the depreciation of the dollar[17].
With dollar depreciation and higher inflation, the real value of the outstanding debt tended to decline, making the servicing capacity easier and driving down delinquency rates.
Quite the opposite, the negative relation CPI-delq of 2009 derived from decreasing inflation rates.
In that case (2009 q2-2009q3), low inflation and low-interest rates may lead to an influx of less creditworthy borrowers who may easily default on their repayment obligations.
To close, although the estimated DCC indicates that delq-REER may change but a low persistence and around a straight line (beta is equal to 0.025), the Wald test results support the hypothesis of constant conditional correlation and reveal that the dynamic specification is inappropriate and incorrect. Therefore, this last relationship remains fairly stable over time.

[16] Quarter by quarter CPI growth passes from 2.19% in January 1987 to 4.47 in December 1987 (Federal Reserve Data).

[17] On September 22, 1985, the G-5 finance ministers of France, West Germany, Japan, United States and United Kingdom signed at the Plaza Hotel in New York an agreement to weaken the dollar and strengthen the yen and Deutsch mark (Plaza Accord). The planned change had the desired effect to reduce American trade deficit and support the recovery from the 1980s recession by restoring American export, domestically produced goods and competitiveness.
Because of the accord, between 1985 and 1987q3, there was a sustained decline of dollar versus British pound and Japanese yen; the dollar came back down 40%.

Conclusion

Many macro-financial phenomena are pictured best in a dynamic version. In this chapter, we raise the question whether the same holds for default rates. In particular, we ask whether USA total delinquency rate co-moves with respect to macroeconomic variables and whether they exhibit time-varying persistence in their corresponding conditional correlation. To do so, and considering the model structure that allows calculating time-varying correlations, we estimate the joint behavior of delinquency rate and macro factors by several bivariate DCC- GARCH, as set by Engle (2002).

To the best of our knowledge, there is no trend in literature for time-varying contemporaneous and lagged correlation between credit quality and macro variables. The paper wants to add the existing literature, being the first that examines the dynamic correlations. It shows not only whether variables commove but also how much. It gives a detailed photograph of the changes in correlations and faces all events that are tied with default shocks. In addition, by studying different variables, the work is able to compare the levels of the different bilateral correlations with delinquency rate over time.

All in all, findings do not suggest a significant time variation in correlation among the variables. Only the linkage between GDP and delinquency rate reveals important fluctuation over time with a more pronounced process in bad times (or when GDP is low), coherently with the idea that expected economic changes and loss of output are a leading vehicle of variation in bad debts. Consequently, such result suggests that looking at the economic trends is important to assess credit risk in the financial system and policymakers should realize robust reforms to improve and strengthen the economy and hence reduce risk and achieve financial stability.

Other variables results are more significant with lags than unlagged. Specifically, not only the lagged relationships with unemployment, inflation and change in stock market exist, additionally, the level and the direction of correlation vary more significantly lending support to the notion that credit quality deteriorates with a time lag of few months/quarter. In this sense, it is notably the case of CPI whose time-varying impact on default loans is only lagged.

Finally, from a comparison perspective, empirical findings demonstrate that conditional correlations mostly exhibit the expected behaviors in terms of signs and are consistent with the practitioner’s results. On the contrary, the regime switching DCC model does not have a better fit of the data than the DCC meaning that structural breaks do not exist or are irrelevant, correlations are not subject to abrupt changes in regime and there are not states with distinct dynamics.

Hence, as opposed to what found in the previous chapter, in this case the hypothesis of changes of states is not verified.

It does not mean that the outcomes contradict each other or that there is something wrong with the novel methodology. Simply, the identification of changes in the levels (chapter 2) does not imply that the same thing happens in the behavior of second moments and therefore in the correlations. In fact, in the financial series, correlations are more linked to increases in volatility (Bauwens and Otranto, 2016).

In terms of the broader literature, our results point to the value of applying an appropriate flexible modeling construction to accurately estimate their interaction and imply that dynamic correlation is a more successful procedure for studying variations of bad debts over the business cycle, demonstrating how analyses based on the simple correlation matrix can be limiting and incomplete.

However, it must be said that generally these models are implemented into an n-variate approach, not suitable in our case.

Indeed, the elements of the conditional correlation matrix need to share the same GARCH dynamic but our starting bivariate correlation evolutions are, reasonably, not identical and mostly constant. Therefore, an approach with many variables should not had sense.

Removing the restriction of common dynamics and allowing the individual series to behave differently in the same model through a clustering algorithm as in Otranto (2010)[18] would undoubtedly gave flexibility to the model but at the expense of an increasing number of unknown parameters to estimate, and generated a number of constraints on the parameters that could have obscured the behavior of the correlations.
Moreover, it was not convenient as our goal is to verify the correlations between macro variables and delinquency rate, not among the other variables.
Based on the research methodology, an avenue for future extension will be to use multivariate GARCH techniques to explain the conditional correlation dynamics between macro drivers and total delinquency rate components and across the different regions of US. Indeed, following Bauwens, Laurent and Rombouts (2006), "recognizing this feature through a multivariate modeling framework leads to more relevant empirical models than working with separate univariate models".

[18] An agglomerative algorithm allows to group homogeneous series by calculating the distance between DCC models.

APPENDIX A

Estimation of the DCC

The estimation is not computationally demanding and it is based on the quasi-maximum likelihood (QML). It can be represented as a two-stage GMM estimation procedure.
Let $\theta = (\omega_1, \alpha_{11}, \dots, \alpha_{1p}, \beta_{11}, \dots, \beta_{1q}, \rho_{12}, \dots)$ a vector of the DCC model parameters with p and q the order of the GARCH(p,q) model. The assumption that the innovation terms ϵ_t follows a multivariate Gaussian distribution with zero mean and covariance matrix H_t implies the following likelihood function for $\varepsilon_t = H_t^{1/2} z_t$:

$$l(\theta) = \prod_{t=1}^{T} \frac{1}{(2\pi)^{n/2} H_t^{1/2}} exp\left[-\frac{1}{2}\epsilon_t' H_t^{-1} \varepsilon_t\right]$$

(12)

By taking the log, we have the log-likelihood function:

$$\log\big(l(\theta)\big) = -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + logH_t + \epsilon_t' H_t^{-1}\varepsilon_t) =$$

$$\log\big(l(\theta)\big) = -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + \log(D_t R_t D_t) + \epsilon_t'(D_t R_t D_t)^{-1}\varepsilon_t) =$$

$$\log\big(l(\theta)\big) = -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2\log(D_t) + \log(R_t) + \epsilon_t'(D_t R_t D_t)^{-1}\varepsilon_t)$$

(13)

Engle (2002) proposed a simplification of the estimation procedure by decomposing the full log-likelihood function into the sum of two parts, namely the volatility component (first step) and the dynamic correlation components (second stage).
Let θ_v the parameters obtained in the first stage or the volatility component and θ_R those obtained in the second stage conditioning on the estimate of θ_v, the first likelihood results by substituting the conditional covariance matrix H_t with an identity matrix I_N:

$$\log\big(l(\theta_V)\big) = -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2\log(D_t) + \log(I_N) + \epsilon_t'(D_t I_N D_t)^{-1}\varepsilon_t) =$$

$$\log\big(l(\theta_V)\big) = -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2\log(D_t) + \log(I_N) + \epsilon_t'(D_t)^{-2} I_N \varepsilon_t) =$$

$$\log\big(l(\theta_V)\big) = -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2\log(D_t) + \log(I_N) + \epsilon_t'(D_t)^{-2} I_N \varepsilon_t) =$$

$$\log\left(l(\theta_V)\right) = -\frac{1}{2}\sum_{t=1}^{T}\left(nlog(2\pi) + \sum_{i=1}^{n}\left[log(h_{iit}) + \frac{\varepsilon_{iit}^2}{h_{iit}}\right]\right) =$$

$$\log\left(l(\theta_V)\right) = -\frac{1}{2}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(log(h_{iit}) + \frac{\varepsilon_{iit}^2}{h_{iit}} + c\right)$$

(14)

which is the sum of the log-likelihood of univariate GARCH equations.
Given the volatility parameters, the second stage relies on estimating the correlation matrix parameters using the correctly specified likelihood function:

$$\log\left(l(\theta_R|\theta_V)\right) = -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2\log(D_t) + \log(H_t) + \epsilon_t'(D_tH_tD_t)^{-1}\varepsilon_t) =$$

$$\log\left(l(\theta_R|\theta_V)\right) = -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2\log(D_t) + \log(H_t) + \epsilon_t'(D_tH_tD_t)^{-1}\varepsilon_t) =$$

$$\log\left(l(\theta_R|\theta_V)\right) = -\frac{1}{2}\sum_{t=1}^{T}\left(nlog(2\pi) + 2\log(D_t) + \log(H_t) + \hat{\varepsilon}_t{H_t}^{-1}\hat{\epsilon}_t\right)$$

(15)

since $\hat{\varepsilon}_{it} = {}^{\varepsilon_{it}}\big/_{\sqrt{h_{ii,t}}}$

The correlation part is then given by the simple maximization of $\log(H_t) + \hat{\varepsilon}_t{H_t}^{-1}\hat{\epsilon}_t$ as D_t is constant when conditioning on the parameters of the first step:

$$\log\left(l(\theta_R|\theta_V)\right) = -\frac{1}{2}\sum_{t=1}^{T}\left(\log(H_t) + \hat{\varepsilon}_t{H_t}^{-1}\hat{\epsilon}_t\right)$$

(16)

Summing up, the maximization of $\log\left(l(\theta_V),(\theta_R)\right) = \log\left(l(\theta_V)\right) + \log\left(l(\theta_R|\theta_V)\right)$ consists in maximizing (29), obtaining the estimates of θ_V and then maximizing (30) conditional on the estimates of the first step.
Under a sufficient set of standard regular conditions, the two-stage DCC estimator gives consistent and asymptotically normal results. A detailed proof can be found in (White, 1994).

However, (Aielli, 2013) pointed out an inconsistency problem noticing that the estimated parameters of the two-step approach suffer from no asymptotic properties; they are bias and inconsistent in large systems
Therefore, he suggested a reformulation of the DCC able to give consistent estimation crucial for adequate performance that mainly involves equation (22) as follows:

$$Q_t = (1-\alpha-\beta)\bar{Q} + \alpha\mathcal{E}_{t-1}\mathcal{E}_{t-1}' + \beta Q_{t-1}$$

(17)

where $\mathcal{E}_{i,t-1} = \varepsilon_{i,t-1}\sqrt{q_{i,t-1}}$.

APPENDIX B

Plots of DCC with lag1 macro variables

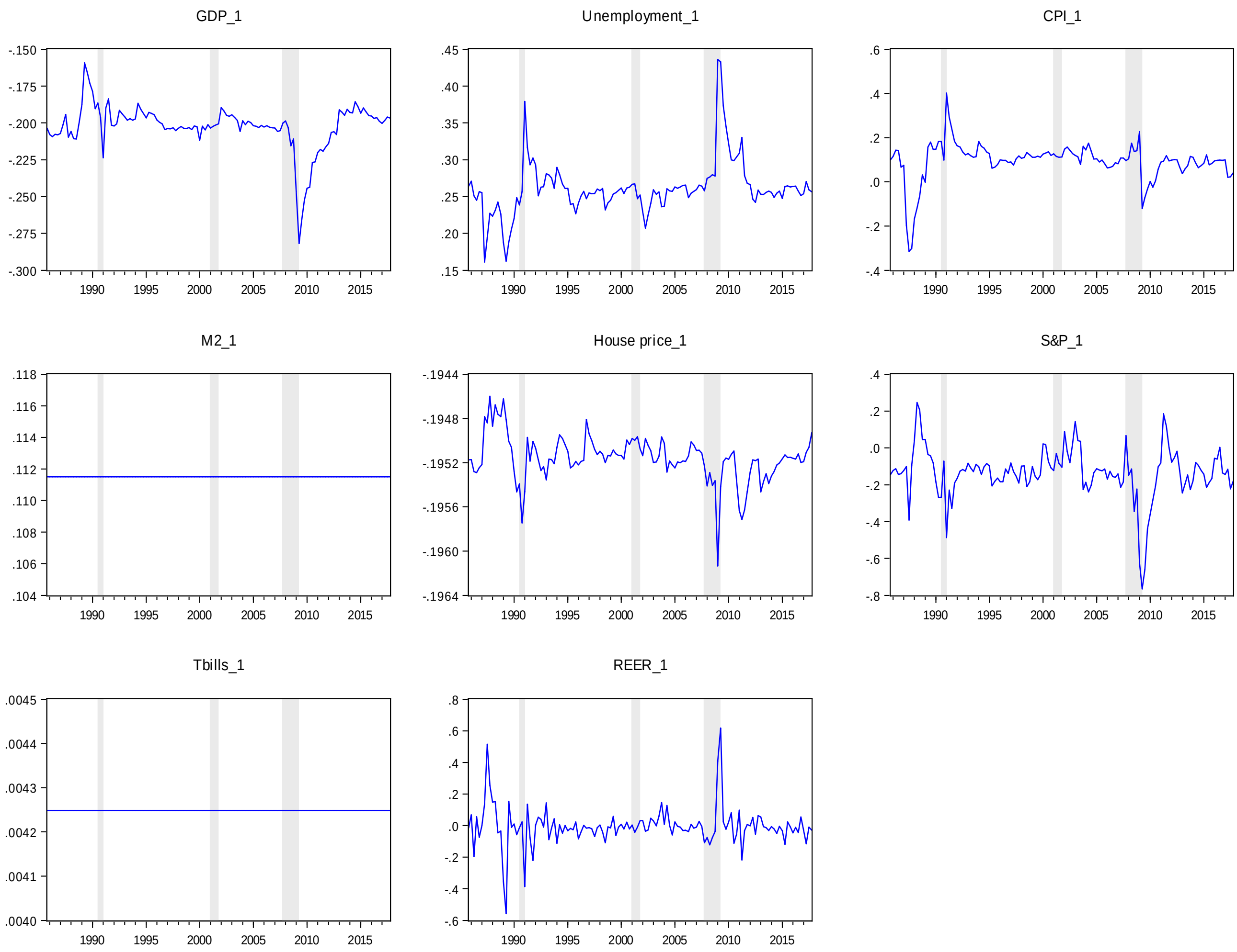


Figure 3-5 Graphs present the time-varying conditional correlation obtained from bivariate DCCs when macro variables are lagged.

REFERENCES

Adeola and Ikpesu. (2017). Macro determinants of non performing loans in Nigeria: an empirical analysis. *Journal of Developing Areas, 51*, 31-43.

Aielli. (2013). Dynamic Conditional Correlations: On Properties and Estimation. *Journal of Business and Economics Statistics, 31*, 282-299.

Anjom and Karim. (2016). Relationship between Non-Performing Loans and Macroeconomic Factors with Bank Specific Factors: A Case Study on Loan Portfolios – SAARC Countries Perspective. *Elk Asia Pacific Journal of Finance and Risk Management.*

Bauwens and Otranto. (2016). Modeling the Dependence of Conditional Correlation on Market Volatility. *Journal of Business & Economic Statistics, 34:2*, 254-268.

Bauwens and Otranto. (2016). Modeling the Dependence of Conditional Correlations on Markets Volatility. *Journal of Business & Economic Statistics*, 1537-2707.

Bauwens, Laurent and Rombouts. (2006). Multivariate GARCH models: A Survey. *Journal of Applied Econometrics*, 79-109.

Beck, Jakubik and Piloiu. (2013). Non performing Loans: What Matter in Addition to the economic Cycle? *European Central Bank Working Paper*, n.1515.

Billio, Caporin and Gobbo. (2006). Flexible Dynamic Conditional Correlation multivariate GARCH models for asset allocation. *Applied Financial Economics Letters*.

Bollerslev. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 307–327.

Bollerslev. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The Review of Economics and Statistics, 72*, 498-505.

Bollerslev, Chou and Kroner. (1992). ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics, 52*, 5-59.

Bollerslev, Engle and Wooldridge. (1988). "A Capital Asset Pricing Model With Time-Varying Covariances. *Journal of Political Economy, 96*, 116-131.

Capiello, Engle, Sheppard. (2006). Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns. *Journal of Financial Econometrics*, 537-572.

Christodoulou-Volos and Hadjixenophontos. (2017). Empirical Determinants of the Non-Performing Loans in the Cypriot Banking System. *Journal of Finance and Investment Analysis, vol. 6, no. 4*, 1-20.

De Lis, Pagés and Saurina. (2000). Credit Growth, Problem Loans And Credit Risk Provisioning" In Spain. *Banco de España — Servicio de Estudios, Documento de Trabajo no 0018*.

Durafe and Singh. (2016). Cyclical Behavior of Public and Private Sector Banks: A Comparative Study of Non-performing Assets. *Journal of Business and Management Research.*

Hurvich and Tsai. (1990). The Impact of Model Selection on inference in Linear Regression. *American Statistician, 44*, 214-217.

Engle. (1982). Autoregressive conditional heteroscedasticity with estimate of the variance of United Kingdom inflation. *Econometrica 50*, 987–1007.

Engle. (2002). Dynamic Conditional Correlation. *Journal of Business & Economic Statistics, 20:3*, 339-350.

Engle and Kroner. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory, 11*, 122-150.

Engle, Ng and Rothschild. (1990). Asset pricing with a factor-arch covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics, Vol 45 - issues 1-2*, 213-237.

Filip. (2015). Non-performing loans – dimension of the non-quality of bank lending/loans and their specific connections. *Theoretical and Applied Economics, Volume XXI*, 127-146.

Filip. (2017). Lessons From The Impact Of Internal And Macroeconomic Determinants Of Bad Loans In Cee Banks. *EURINT.*

Fofack. (2005). Non Performing Loams in Sub Saharan Africa: causal analysis and macroeconomic implications. *World Bank Policy Research Working Paper 3769.*

Ghosh. (2015). Banking-industry specific and regional economic determinants of non-performing loans: Evidence from US states. *Journal of Financial Stability, 20*, 93-104.

Guidolin. (s.d.). Modelling and Forecasting Conditional Covariances: DCC and Multivariate GARCH.

Iuga and Lazea. (2012). Study regarding the influence of the unemployment rate over non performing loans in Romania using the Correlation indicator. *Annales Universitatis Apulensis Series Oeconomica, 14(2).*

Messai and Jouini. (2013). Micro and Macro Determinants of Non-performing Loans. *International Journal of Economics and Financial Issues*.

Mondal. (2016). Sensitivity of Non-Performing Loan to Macroeconomic Variables: Empirical Evidence from Banking Industry of Bangladesh. *Global Journal of Management and Business Research, Volume 16 Issue 4 Version 1.0.*

Mortensen . (1977). Unemployment Insurance and Job Search Decision. *Industrial and Labor Relations Review, Vol. 30, No. 4*.

Murumba. (2013). The Relationship between Real GDP and Non-performing Loans: Evidence from Nigeria (1995 – 2009). *International Journal of Capacity Building in Education and Management (ICBM), Vol. 2, No 1*.

Otranto. (2010). Identifying financial time series with similar dynamic conditional correlation. *Computational Statistics and Data Analysis 54*, 1-15.

Pelletier. (2006). Regime switching for dynamic correlations. *Econometrics*, 445-473.

Psaradakis and Spagnolo. (2003). On the determination of the number of regimes in Markov-Switching Autoregressive models. *Journal of Time Series Analysis Vol 24*, 237-252.

Quagliariello. (2006). Banks' Riskiness over the Business Cycle: A Panel Analysis on Italian Intermediaries. *Discussion Papers Banca Italia.*

Salas and Saurina. (2002, December). Credit Risk in Two Institutional Regimes: Spanish Commercial and Savings Banks. *Journal of Financial Services Research, vol.22, issue 3*, p. 203-224.

Sandada and Kanhukamwe. (2016). An Analysis of the Factors Leading to Rising Credit Risk in the Zimbabwe Banking Sector. *Acta Universitatis Danubius, vol 12, n.1*.

Silvennoinen and Teräsvirta. (2005). Multivariate autoregressive conditional heteroskedasticity with smooth transitions in conditional correlations. *SSE/EFI Working Paper Series in Economics and Finance No. 577*.

Socol and Iuga. (2010). Study of correlation between average interest rate and nonperforming loans in the Romanian banking system during 2006 - February 2010. *Annals of University from Oradea, Economic Sciences, TOM XIX, 2nd ISSUE / DECEMBER*.

Tse and Tsui. (2002). A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with TimeVarying Correlations. *Journal of Business and Economic Statistics*, 351-362.

Van Der Weide. (2002). GO–GARCH: A multivariate generalized orthogonal GARCH. *Journal of Applied Econometrics, 17*, 549-564.

Verbeek. (2012). *A guide to modern econometrics.* John Wiley & Sons.

White. (1994). Estimation, Inference and Specification Analysis. *Econometric Society Monographs no. 22, CUP*.

Conclusions

Adopting a non-linear perspective, this thesis discusses the ways in which macroeconomic developments impact banks loan quality and make it difficult for bank borrowers to pay their debts in full and on time.
The relevance of the topic has been highlighted by several financial crisis episodes. In fact, the latter have demonstrated how credit risk is not a separate issue but ties in with macro-economic conditions given that historically countries that have been hit by the deepest recessions recorded the highest bad debts ratios.
See, for example, the recent global financial crisis that caused a significant increase in non-performing loans almost within every developed and no country. Being related to the deterioration of the borrowers' creditworthiness as a result of an economic decline, this phenomenon is a further demonstration that credit risk does not depend only on specific microeconomic factors, but also the macroeconomic factors have an impact on it. Therefore, academics and workers in the sector commonly agree that banking stability is a key policy object and underscoring the causes of banks' credit risk is a precondition for the long-term viability of a country's banking sector.

In the context of the thesis, the relationship credit quality-macroeconomic environment has been reexamined through three different chapters: at first, it is specified an overview regarding the present literature, and then two quantitative studies are performed.

Specifically, chapter one re-evokes lots of academic publishing, especially the quantitative and methodological contributions, in order to underline potential limitations of existing international research. By noticing that the literature generally displays similar symptoms of credit risk partly because all studies adopt models such as linear regressions, models for panel data and VAR that feature stable relations expressed by fixed parameters, the chapter primary criticizes against the failure to account for the nonlinearities inherent in the relationship.
Indeed, the strong link between bad debts and economic cycle subject to changes in regimes that disturb the stability of parameters over time suggests the attempt to model the variable of bad debts with changing parameters based on different regime as quite natural.

In doing so, chapters 2 and 3 aim to provide an answer to the question of non-linearity and fill the gap in the literature.

Each chapter applies advanced time series econometric techniques to a quite long sample of data. This is a disregarded strategy in the literature about the linkage between bad debts and macro-economic environment since commonly there are available short data samples about credit quality. On the contrary, the use of long time spans of data can make it easier to discover the underlying dynamics in the data.
Therefore, a key contribution of the empirical research relates to the choice of econometric techniques used, namely, the Markov switching approach and the DCC modelling.
Although both techniques have been used in economic research, in the case of empirical work on forecasting credit risk determinants and modelling the evolution of conditional correlation across variables, the evidence is almost non-existent.
As such, the thesis addresses a gap in the existing literature.
The MS developed in chapter 2 copes with the problem of structural changes in the data and simplifies the investigation of credit risk's responses to shocks originating from a set of macro variables under different regimes. By focusing on three credit quality variables for the United States, real estate, consumer and commercial and industrial loan losses, the chapter puts the linearity assumption into question, proves the

expected non-linearity of their relationship with macroeconomic variables, and evidences a clear superiority of the MS model over linear ones in both modelling and forecasting performance.
In other words, the work suggests that linear models may work well under regular economic conditions but when switches are evident, the MS model is better in spite of possible over parameterization.

By estimating the joint behavior of total US delinquency rates with the same macro data introduced previously, chapter 3 proposes an application of the DCC-GARCH model.
One of the appealing feature of the model consists of quantifying how much stable and lasting linear relationships between real and financial sector is by capturing the time-varying conditional correlations without creating too many parameters.
Despite in the DCC setting the hypothesis of changes of states does not hold, results of simple DCC analysis document that time varying co-movements are present for the USA market just for GDP whereas, for the other variables, they appear mostly in lagged terms lending support to the intuition that credit quality deteriorates with a time lag of few quarters.

Results adds to the arguments for enhancing analysis on the determinants of loan quality.
They generally have several policy and banking management implications useful to minimize and control defaults but, generally, to promote the overall financial health and a sustainable growth. Identifying the macro predictors of increased bad debts and understanding what typically corresponds to magnified correlations could help to implement proper policies and diversify strategies depending on the economic trends to strengthen the banking sector, restrain, and limit potential contagious effects.