LIBERA UNIVERSITÀ INTERNAZIONALE DEGLI STUDI SOCIALI

Department Business and Management

Chair Quantitative Methods for Management

The term spread as a predictor for recessions. Empirical evidence on the Italian case

Professor Marco Pirra

SUPERVISOR

Professor Paolo Porchia

CO-SUPERVISOR

Nicolò Bongiovanni

CANDIDATE

Academic Year: **2019/2020**

Table of contents

1. INTRODUCTION ... 3

2. THE YIELD CURVE ... 6
2.1 - Fundamentals ... 6
2.2 - Theories behind its shape ... 7
2.3 - Italian scenario ... 8

3. LITERATURE REVIEW ... 11
3.1 - Past results ... 11

4. METHODOLOGY ... 20
4.1 - Research question ... 20
4.2 - Data collection ... 21
4.2.1 - Defining the Spread ... 21
4.2.2 - Defining Recession ... 24
4.2.3 - Other indicators ... 26
4.3 - The analysis: the probit model ... 30

5. RESULTS ... 34
5.1 - Predicting an economic crisis ... 34
5.2 - Predicting a recession ... 39
5.3 - Predicting a recession with other indicators ... 42

6. DISCUSSION ... 45
6.1 - Interpretation of the results ... 45

7. CONCLUSION ... 52

8. REFERENCES ... 53

9. APPENDIX ... 56

1. INTRODUCTION

The inversion of the yield curve is an extremely rare phenomenon. When the long-term yields fall below the short-term rates, the fear of an impending recession starts to spread among the markets, motivated by a general lack of confidence in the near future. There are many studies that show the predictive power of the yield curve, which is usually taken as a benchmark for the health of the overall economy. At the time of drafting this paper, the last inversion occurred in the United States in August 2019, therefore, if the predictive power of this event was confirmed, our economy would be approaching a recession in a few months. However, due to the worldwide pandemic Covid-19 that affected the global economy in the first semester of 2020 and that caused itself a recession, it will be difficult, if not impossible, discerning the effect of the inversion of the yield curve on the market crash. Nevertheless, a sufficient reason to develop further studies about this topic is represented by the last inversion of the curve prior to August 2019. Indeed, the previous inversion occurred in June 2007, right before the subprime crisis that later hit the market and caused the well-known damages.

Nowadays policymakers have the possibility to rely on a vast range of methods and an enormous amount of data to forecast future market conditions, but an analysis of the yield curve movements, which proved to be valid in the past, could still represent a further and useful study. However, when trying to predict a recession, it would not be correct to solely rely on the yield curve, since the history shows false positive cases. Moreover, although many studies confirmed that it is possible to predict a recession from the inversion of the yield curve, it is not possible to predict neither its severity nor its length. The severity and the length of a recession are two characteristics that vary for each recession. Therefore, it is of utter importance to provide a clear definition of this phenomenon, in order not to create confusion between recession periods and temporary economic downturns. Later on this paper, the reader will find a clarification about this distinction, which is necessary to provide a clear analysis. It is important to mention that the occurrence of the inversion of the term structure in a specific country does not necessarily act as predecessor for the same event in other nations. Indeed, there could be countries where either the inversion of the curve occurred in different time periods, or where this event did not have any consequences. This divergence may be even more manifest whether the comparison is made among countries of different continents. Nevertheless, for the western world, the curve inversion in the United States is usually denoted as a strong alert on the overall economy's health, and whenever this phenomenon occurs, financial markets start to tremble.

Before deeply explaining the fundamentals of the yield curve and the reasons that may lie behind its inversion, for the sake of clarity, it is essential to introduce this topic with a few considerations. First of all, the inversion represents a sort of "break" of the traditional theory regarding the value of time. In a traditional context, where the slope of the yield curve is positive or slightly upward sloped, investors who choose to invest their resources in long-term bonds, willing to hold part of the nation's or company's debt for a longer period of time, will earn a higher compensation. This concept seems reasonable, since extending the period of the investment, debt holders will not receive their money back for a longer time and will face higher risk. However, this reasoning is overturned when we face an inverted yield curve. In this situation, the investors do not feel safe in buying short-term securities despite their higher returns. Indeed, they prefer to invest in long-term debt, because of the shared expectation of an economic downturn in the long run. Acquiring bonds with long-term rates will ensure the investors a higher return whether the policy makers will cut the interest rates, a measure that represents the traditional reaction that central banks carry out to fight a recession time.

These few considerations let us understand that the inversion of the yield curve it is by meaning something unusual and interesting. In addition, an analysis of this phenomenon is even more worthy of attention in a modern context, characterized by high volatility, negative interest rates and increased credibility attributed to monetary policy. The main purpose of this paper is to study the historical inversions of the Italian yield curve and the curve predictive power in this country, since most of the literature concerning this topic have a United States or European outlook. To better study the predictive power of the Italian yield curve, we will conduct an econometric analysis, using the so-called "probit model". This method allows to assign a probability to the recession event, starting from the moment of the inversion of the yield curve. More precisely, this model allows to convert the steepness of the curve into a probability of a recession in the subsequent period of time, with a specific time forecast. The model is based on the definition of two important factors, whose misinterpretation could lead to different results: recession and spread. As mentioned above, providing a clear definition of the values of a recession is fundamental to correctly shape the model. The same relevance is attributed to the definition of the spread. In this framework, the spread refers to the difference between the rate of return of two government securities with different maturities. This definition is a completely different interpretation compared to the traditional one, which points out the basis point difference between two countries' government bonds yields. Choosing different combinations of maturities will lead to results which will present different levels of reliability, and hence, different predictive powers. To identify those spreads which provide the highest explanatory powers, we will compare the results with specific techniques. Throughout the model development, both recession and spread definitions will be clearly addressed. Moreover, in order to have an idea of the accuracy level of the term spreads' performance in predicting recessions, we will compare the obtained results with a bunch of alternative indicators.

The largest part of the existing literature mostly refers to the past decade and beyond, and mainly have a US or EU framework, providing an attractive space for more recent studies such as this one. This document contributes to the body of literature that looks at the relationship that exists between movements of term spreads (inverted yield curve) and the rise of a recession, providing an empirical evidence on the Italian case. The paper is organized as follows. The next section provides the basic knowledge and theoretical framework behind the yield curve and its shapes, with a look into the current Italian economic scenario. Chapter three is a dive into the existing literature, which provides evidence of the predictive power of the yield curve, across countries and time. The following section is the paper's core, where the reader will understand the probit models' estimations and the database construction. Chapter five will provide all the results of our analysis, with a comparison among the outputs. Finally, before the conclusion, chapter six will present a summary of the results obtained, a discussion of the potential interpretations and the limits of our work.

2. THE YIELD CURVE

2.1 - Fundamentals

"The yield curve is a graph that plots the yields of various bonds against their term to maturity"[1]. This definition summarizes in a few words the main concept that underlies this subject, whose characters are bonds, debt instruments that investors are willing to buy and hold for a short or long period of time, in exchange for a proper compensation defined by the interest rates. Taking into consideration the nation's debt, which is considered to be risk-free, the line represented by the combination of the various levels of interest rates and corresponding maturities defines the yield curve. The yield curve is not an historical chart; indeed, it could be also defined as a snapshot of the current yields' levels in the market. As mentioned in the introduction, intuitively, the general expectation combines longer maturity with higher yields, because long-term debt holders face potential inflation and other risks, holding the instrument for many years. Among the uses and interpretations of the yield curve, one main and distinctive use stands out: it sets the benchmark for the listing of all debt market instruments. When an entity (may it be a company or a State) comes to the issuance of new debt, issuers take into consideration the yield curve to price the debt securities. Secondly, the curve is used to anticipate future economic growth, providing market expectations about the course of future interest rates. Bank lending rates and mortgage rates are based on this curve as well, therefore, it is easy to understand the importance of the yield curve. Banks and lending institutions borrow money from central banks at short and usually low rates, then they lend it to privates or companies for higher rates and longer periods of time. This process summarizes how banks create their margins. However, if the yield curve flattens, their margins start compressing and these institutions become less and less willing to provide capital. In the extreme situation represented by the curve inversion, lending institution will hardly make loans, creating hard restrictions to access to their resources. By this way, the consumptions sink, businesses stop investing because of the higher cost of capital, and the whole economy slows down, entering into a cycle where a recession seems to represent the direct and inescapable consequence.

Short-term rates are directly affected by the actions of the central banks; indeed, the curve is fundamental for the transmission of the monetary policy. These institutions cut or raise the short-term interest rates according to the current economic situations: usually, there is a rates cut when central

[1] Moorad Choudhry, 2019. "Analysing and Interpreting the Yield Curve". Wiley Finance Series, second edition.

banks want to stimulate consumptions and investments, while a rates raise when the economy is running fast and central banks need to prevent high levels of inflation. The inversion of the yield curve is likely to occur when central banks decrease rates and market operators expect further cuts. Believing in a future recession, investors will start investing in long-term bonds, creating high demand for instruments at the long end of the curve. This high demand will lead to an increase of the prices and a decrease of the returns for long-term securities, and the opposite effect on short-term bonds, which will therefore reflect an increase in the rates. This process, whether long-lasting, will lead to the change in the form of the yield curve, which from upward sloped will turn into downward sloped. Generalizing, we may sum up the role of the yield curve with the following statement: on one side it is essential for monetary policy purposes, and on the other side it provides relevant insights regarding investors' expectations about the future outlook of the overall economy[2].

2.2 - Theories behind its shape

Evidence on historical graphs shows that four basic representations of the yield curve have been identified, analyzing different countries and across different period of times: positive sloped, flat, negative sloped and humped[3]. These shapes are as clear as their denomination. Besides these slopes, three studies have been developed to justify the yield curve shapes: the theory of pure expectations, the liquidity premium theory and the preferred habitat theory.

In the theory of pure expectations, market expectations are perfectly reflected in the form of the curve: an upward movement of the yield curve indicates an anticipated increase of future short-term rates and vice versa. This theory relies on the basic assumption that investors do not have preferences for different maturities, as long as they maximize their returns. They invest according to their expectations about future rates, creating yield differences for securities with different maturities. This process leads to changes in the slope of the curve, and therefore, an inverted yield curve would mean that market players fear an impending recession. According to this theory, the recession signals provided by an inverted curve may come significantly in advance compared to the signals provided by other indicators, since long-term expectations play such a critical role in shaping the curve. The same assumption (market expectations are perfectly reflected in the form of the curve) does not hold

[2] The yield curve may be used for other reasons as well. However, the main purposes are the two described above (example of other uses: pricing of interest rate derivatives).
[3] Humped: short-term rates rise until they reach a peak in the medium term. From this peak, they start sloping downwards for the long-term.

for the liquidity premium theory. In this case, differences in maturities are instead the only reason why investors require different returns, weighing the maturities with relative returns. In this theory, the curve reflects the future short-term rates plus a liquidity premium that depends on the maturity, which represents a sort of compensation for investors who are willing to buy securities expiring later in time. The longer the maturity, the higher the premium. Short-term securities have lower interest-rate risk [4] and higher liquidity, while long-term securities have higher prices volatility. This combination entails a higher demand for short-term instruments which consequently will provide lower returns. The liquidity premium theory is therefore consistent with an upward sloped yield curve, because an inverted yield curve would attribute a premium to short-term investors, penalizing long-term investors, who already suffer from other risks. The third and last theory relies on the assumption that different players of the market have different priorities, and therefore prefer investing in specific maturities. The distinction is purely attributed to the kind of need and relative investment horizons of several investors. For example, banks and other lending institutions usually focus great part of their investments in the short-term part of the yield curve, because of their daily cash management activities. At the same time, pension funds or insurance companies, which prefer a lower combination of risk and return, usually invest in long-term securities. This theory can explain all the shapes of the yield curve, and the spread will rise because of the different levels of supply and demand for specific kinds of instruments.

2.3 - Italian scenario

Since this paper seeks to analyze and prove a direct link between the inversion of the Italian yield curve and a subsequent recession, it is necessary to provide a glimpse about the current Italian scene, concerning both yields levels and economic cycle. As stated in the introduction, our study is based on data available at 31st December 2019, because the 2020 outlook has been deeply affected by the current situation linked to the pandemic Covid-19, and this analysis could provide misleading results if taking into consideration the most recent data. Figure 1 shows the graphical representation of three Italian yield curves at three different dates. The light blue line is the curve on the 30th April 2020, the red line is the curve on the 30th April 2016, while the green line is the curve on 30th November 2014. From the chart, we can see that the two most recent yield curves both have the short-

[4] Interest rate risk: potential of loss resulting in a change of the interest rate. Given the inverse relation between bond price and interest rate, if rates rise, bonds value will decrease.

term rates below the zero threshold, which is a feature of the current financial markets. The green line, which among these three refers to the oldest date, shows a higher slope and higher yields for most of the maturities. Indeed, between 2014 and 2020 interest rates decreased significantly. Another observation is that all of these three curves have positive slopes, with long-term rates definitely higher than short-term rates.

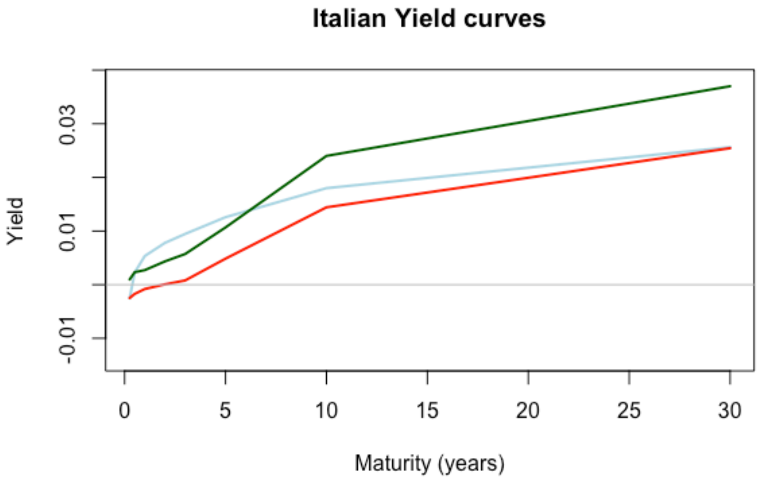


Figure 1: Author representation (data provided by ISTAT)

As mentioned in the previous paragraph, the yield curve is used as a benchmark for different purposes. For example, it is used to set the interest rates of bank loans and mortgages. Figure 2 provides the monthly interest rates imposed on loans for two different targets, privates and companies. In the graph, both Italian values (red lines), and European values (blue lines) are showed. The continuous lines represent the interest rates payed by companies to access to bank capitals, while the dotted lines represent the interest rates payed by privates when they request a mortgage for house purchase. All these values show average estimates. The first impression is that approaching more recent years, there has been an overall and sharp decrease in the rates, both in Italy and in Europe. Moreover, the spread between Italian and European values have decreased, especially at the end of 2019. From the chart it also stands out that nowadays borrowing money became cheaper compared to the past, both for families and for companies. Indeed, in the last ten years central banks often decreased short-term interest rates to stimulate the economies, increase consumptions and push companies to make more investments.

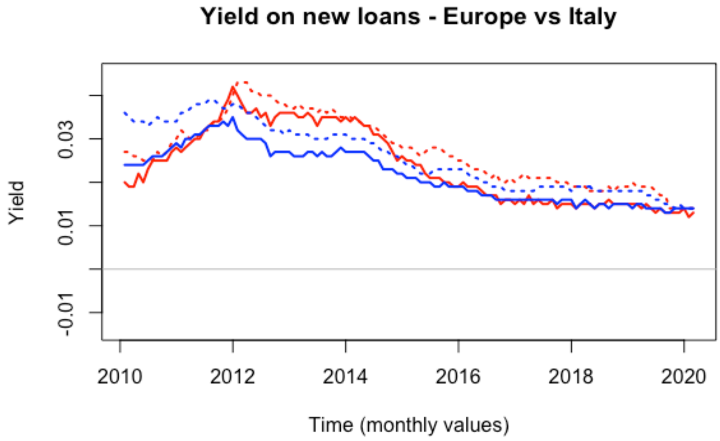


Figure 2: Author representation (data provided by ECB)

The third and last chart (Figure 3) of this section shows the growth of the Italian GDP from 1995 until last year, 2019[5]. The past few years could be described as years of uncertainty, characterized by tensions between countries, such as the US-China trade war and Brexit. The year 2019 ended with slightly positive results (GDP: 1,787.7 billions, vs 1,766.2 in 2018), sustained by an improved level of occupation (which reached historical highs since 1977) and a low level of inflation. However, this weak growth could be risky for Italy, which could fall into a recession with new eventual shocks. A driving element of Italian economy are exports, which grew in 2019 besides the tariffs imposed by Us and trade limitations. The overall favorable conditions of the labor market inducted a slight improvement of families' income, which by consequence lead to a slight increase of the consumptions. The shaded areas in the graph represent the two large crises that hit Italy, and the world, in the last twenty-five years. It is evident how the GDP started to fall in correspondence of the beginning of both crises.

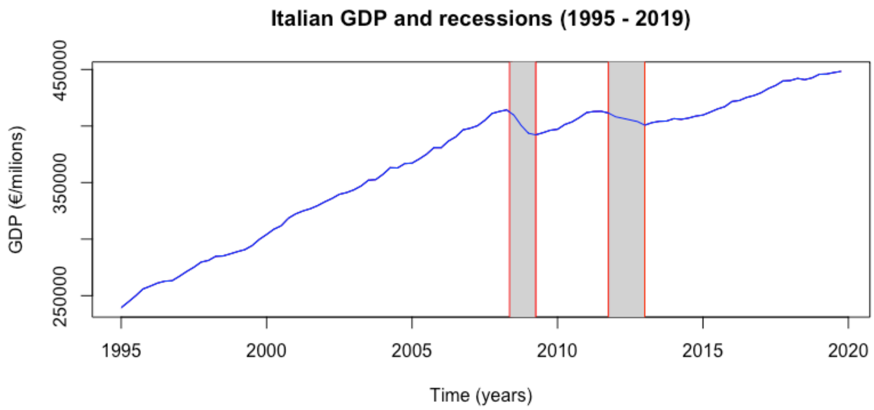


Figure 3: Author representation (data provided by ISTAT)

[5] The chart shows the GDP calculated with current prices. Therefore it has not been adjusted for the level of inflation.

3. LITERATURE REVIEW

3.1 - Past results

Recessions are difficult to predict, both because of their sporadic nature, and because their fundamental factors usually show discrepancies across time. Economists tend to explain each recession with baskets of variables that have influence during specific periods. Therefore, using the same factors combination to study different events, is likely to show inconclusive results. However, the slope of the yield curve has revealed to be a strong indicator for economic shocks in the short-term future, across countries and time. The first qualitative reports that focus on the combined behaviors of interest rates at different maturities and the business cycles dates back to early 1900. However, at the best of our knowledge, Kessel (1965) was the first to provide a specific study about the behaviors of interest rate term spreads according to fluctuations of macroeconomic variables. Generally, in the oldest studies, researchers initially focused their attention on the relationship between movements of the yield curve and levels of inflation[6]. Then, the focus shifted on the linkage between the curve and the real activity, a topic deeply analyzed by several authors, including Harvey (1988), Laurent (1988, 1989), Chen (1991), and Estrella and Hardouvelis (1991). For instance, these last two researchers analyzed the spread between long-term and short-term rates, in comparison with subsequent growth and future recession. They found a strong connection among these events, considering a sample composed by United States data. Indeed, the majority of available reports focus on the United States and provide evidence on prediction of real activity with lead times from 1 to 8 quarters. In particular, Estrella and Hardouvelis (1996) analyzed which variable among the stock price index[7], the Stock-Watson[8] index, and the yield curve term spread was the most accurate to predict recessions according to the time horizon considered. Their research showed that the Stock-Watson index was the best instrument to predict recessions one quarter in the future, although the other two variables presented predictive power as well. Nevertheless, when extending the horizon two or more quarters in the future, the term spread of the yield curve proved to be the best tool, while the stock price and Stock-Watson indexes showed more false-positives cases, and failed in

[6] Among the others, Mishkin (1990). He derived a framework from the Fisher equation, which expresses the nominal interest rate in terms of real rate and expected inflation. This equation is mainly used to calculate the Yield to Maturity of a specific asset. The Mishkin approach is based on the Fisher decomposition.
[7] In this research the NYSE (New York Stock Exchange) was taken into consideration.
[8] The Stock-Watson index is an alternative index of leading indicators.

anticipating large downturns. The dominance of the yield curve spread was even more confirmed when considering longer horizons (four to six quarters ahead).

To have a more recent outlook, it is possible to refer to reports published in the last few years, which therefore provide innovative point of views compared to older literature. In these next paragraphs, researches conducted by David Miller (2019), Johansson and Meldrum (2018), Rudebusch and Williams (2008), O. Emre Ergungor (2016), Bernard and Gerlach (1996) and Fabio Moneta (2003) will be addressed. Each of them provides different but inherent perspectives, useful to gain concrete insights to understand the underlying dynamics behind the predictive power of the curve.

As mentioned in the introduction, in order to correctly evaluate the predictive ability of the yield curve, it is essential to define the boundaries of the term spread, the main character of the analysis. Deciding which spread to consider often conducted to divergent opinions. For instance, David Miller (2019) demonstrated in his research that it is hard to find a unique best predictor, and that many spreads may be explicative. Using the AUROC[9] model, he found out that the time horizon of the forecast is determinant in analyzing the accuracy of the spread (Figure 4). The graph below studies the accuracy of different combinations of term spreads in predicting economic downturns, and it is based on United States data between 1976 and 2010 (the author analyzed the accuracy of 20 different spreads[10]). On the left side of the chart, it is evident that different combinations of terms show different levels of reliability (y-axis, values are between 0 and 1, and 1 defines the highest accuracy) according to the time horizons considered (x-axis). Instead, the right side of the graph highlights the best combination of terms according to the time horizon. The highlighted lines create the most accurate combination, which is therefore composed by the segments that are closer to 1, according to the x-axis. From these evidences, it is possible to draw some conclusions: the spread between short and very short Treasuries was optimal to predict downturns at short horizon, while the spread between medium and short rates was best suited to predict recessions at longer horizons. However, especially considering the timeframe between eight and eighteen months, many spreads showed similar predictive abilities, almost identical in some traits.

[9] AUROC Model: It studies the accuracy of a model by measuring the ratio of true positives to false positives that the model predicts on the sample. Auroc = 1 defines the most accurate model.

[10] Considering the term spread, the general assumption, confirmed by other several authors, is to focus on the spread between long and short rates (usually, the spread between 10-year and 2/3-months rates). Miller instead did not rely on this assumption, and tested other combinations, such as the short-short or medium-short spread.

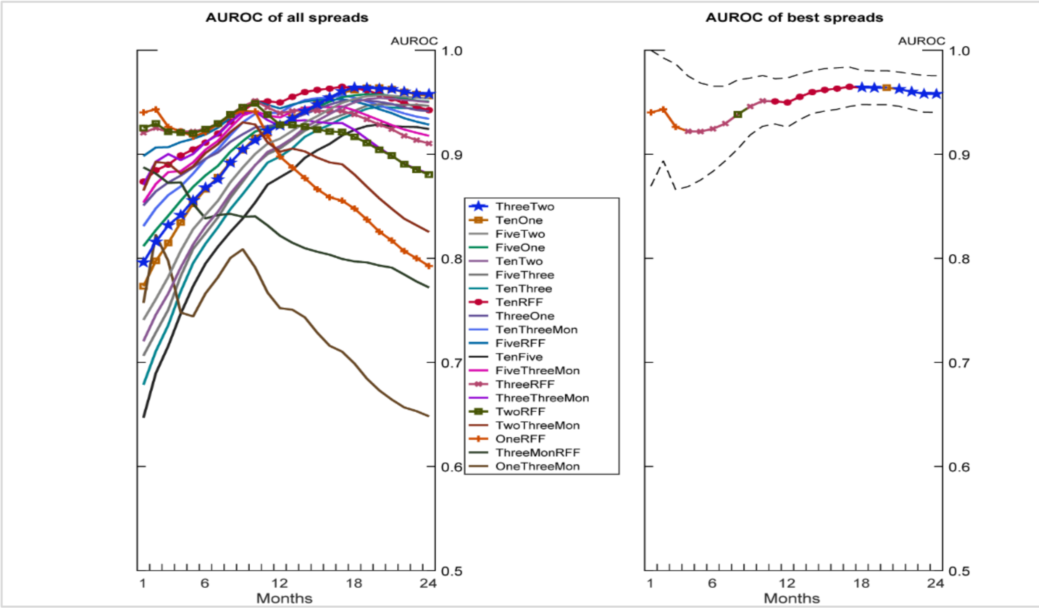


Figure 4: David Miller (2019), Federal Reserve System

The flattening of the difference between short and long United States Treasury rates dating December 2017 induced the research about the predictive power of the yield curve by Johansson and Meldrum (2018) (this paper, together with the analysis of Bauer and Mertens (2018), are two among the latest versions which provide evidence about the predictive power of the curve). In this paper, the authors took into consideration the spread between the ten-years and three-months interest rates, and the latest recessions in the United States. To define these events, they considered the definition of the National Bureau of Economic Research (NBER)[11], which identifies six[12] recessions in the years between 1970 and 2018. Comparing the time of the recessions, and the difference between long and short rates, it is easy to see that negative term spreads have been surprisingly efficient in anticipating disastrous economic situations. In the graph below (Figure 5), the grey areas define the recession periods and the blue line represents the spread between the above-mentioned securities. It is possible to notice that in only one situation (the 1990-1991 crisis caused by the Gulf War) the term spread remained positive before the year of the recession (positive, but close to zero). To test the predictive power of the spread, the authors applied three different variants of the probit model: the first considering the standard term spread as explanatory variable, the second replacing the slope of the term structure with the first three principal components of yields, and the third adjusting the term spread for the term premium. These different models proved to be best fit for different periods, and

[11] The NBER defines the recession as: "a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales”.
[12] The six US recessions date back: 1973-1975, 1980, 1981-1982, 1990-1991, 2001, 2007-2009.

hence, could predict more precisely recessions with different forecasts. Indeed, as general result, the term structure performed as a great predictor for future economic activity. Comparing these results with those obtained by older studies, among the others Estrella and Hardouvelis (1991), the predictive ability provided by this analysis is slightly weaker. This difference is mainly attributable to the recent conditions that have been affecting the markets in the last few years, such as very low (even negative) interest rates, or the decline of GDPs growth rate expectations across developed countries.

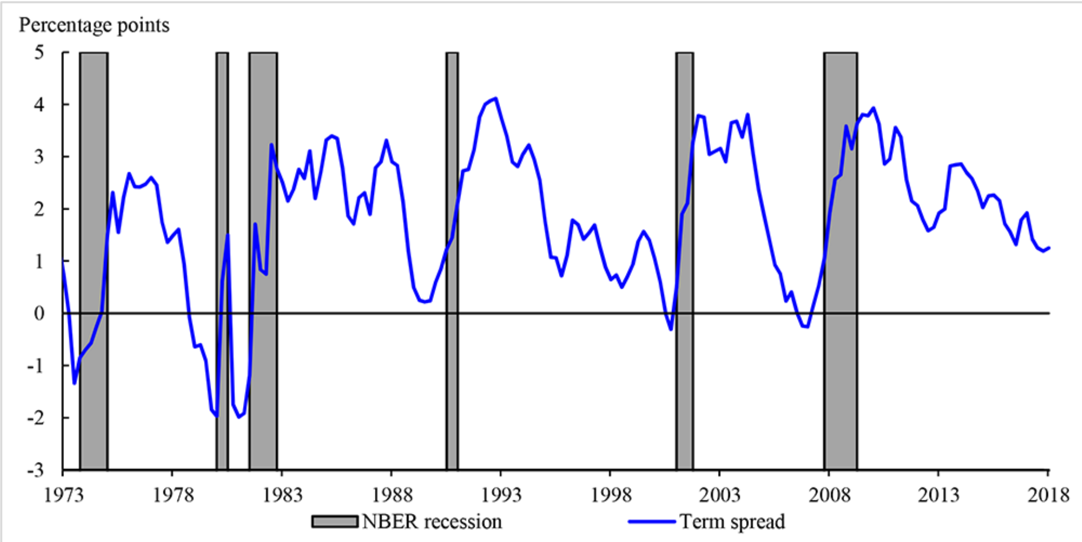


Figure 5: Johansson and Meldrum (2018) – FEDS Notes

Two other economists, Rudebusch and Williams (2008), compared the predictive power of the term spread with the results obtained by the so-called SPF, the Survey of Professional Forecasters[13]. This survey is a quarterly analysis conducted by professionals of the economic sector, such as members of financial or non-financial institutions and statistical researchers. The document put together the opinions of these individuals, who are asked to provide their thoughts concerning several subjects, such as expected inflation, GDP growth and unemployment rate at different horizons. In the paper, the model used is a probit model, which took into consideration SPF forecasts results between 1968 and 2007, and the yield curve term spread between the 10-years Treasury note and the 3-months Treasury bill in the same time horizon. The main result showed that, in several occasions, the term spread proved to outperform the SPF forecasts, demonstrating a better ability in predicting recessions. To be more specific, the paper showed that SPF forecasts provided slightly better results

[13] The SPF considered in the paper is the US version. It is conducted by the American Statistic Association and National Bureau of Economic Research. Concerning Europe, the SPF is made by the European Central Bank. Lahiri and Wang (2006) studied the forecast performance of this periodical studies, finding that they show a quite remarkable level of accuracy in predicting recessions in the very short-term.

when predicting downturns in the current or in the quarter ahead, while the yield curve inversion was definitely more accurate with forecast horizon of three or four quarters. This conclusion seemed to suggest that economists "failed" to learn about the strong predictive power of the yield curve inversion, which in past years proved to predict recessions with great accuracy. The authors provide one possible explanation for this "mistake": economists and professionals who live in different periods and face different macroeconomic scenarios, although without denying the predictive power of the curve, tend to underestimate its usefulness. They are used to consider the overall conditions of each period unique and incompatible with previous events. In the graphs below (Figure 6), the black line represents the yield spread probability forecasts, while the dotted line represents the SPF forecasts. In the chart on the right, it is evident how the yield curve term spread dominated the other variable, predicting most of the recessions with greater accuracy and better timing.

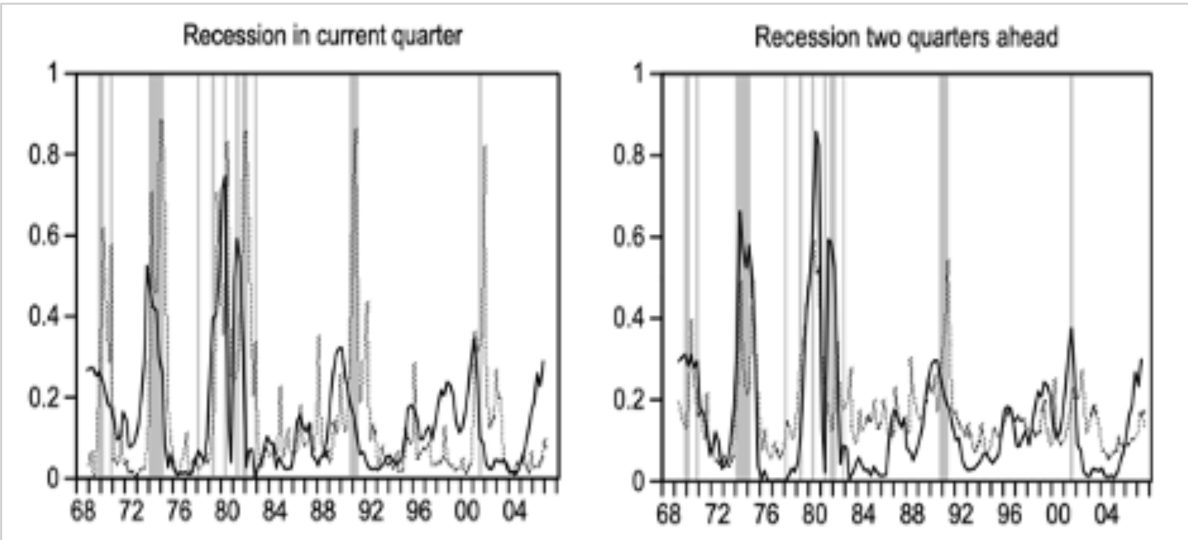


Figure 6: Rudebusch and Williams (2008)

Following the idea that current economic conditions identify specific and unique scenarios, Emre Ergungor (2016), who also analyzed the term spread power in predicting recessions, conducted a slightly different approach in his research. He started from the concept that the reliability of the term spread has changed over the time, because of the impact of recent levels of interest rates, which in the last years approached (and then, overcome) zero at short-term. When Treasuries with very short maturities have zero interest rates, in order to create an inversion of the yield curve and reproduce a negative term spread, the long-term yields should fall deeply below the zero level. However, this event is incompatible with the investments of pension funds, insurance companies and institutional investors, whose goal is to lock the money into safe and long-lasting assets. For this kind of investors, which typically buy securities with very extended maturities, positive interest rates are essential. Besides considering the traditional differential between long-term and short-term rates, the author added two new variables to test if these new scenarios affected the predictive power of an inversion of the curve: the credit spread and the growth of corporate profits. The credit spread defines the difference in yields between the rates of securities that corporate borrowers pay, while keeping the maturity constant. These borrowers are considered to have different creditworthiness (for example, the difference between rates with same maturities of a high-quality borrower, rated AAA, and a bad-quality borrower, rated BBB). In the corporate bond market, the credit spread should reflect market expectations about default risk. Therefore, an increase of the perceived default risk (meaning an increase in the credit spread) should be predictive of an incoming economic downturn[14]. The combined analysis of the term spread and the credit spread lead to better results compared to previous studies, which only based on the traditional difference between short and long-term yields. This model, with two variables, had better ability in predicting recessions twelve months ahead. Nevertheless, since the credit dynamics are based on the same fundamentals of the yields' dynamics (the fixed-income market), the implicit errors showed up again. Indeed, the yields imposed to high-quality borrowers, such as AAA company, presented the same problem as before: the zero bound limit that affected the short-term maturities. This phenomenon therefore created a redundant issue for the purpose of the analysis. The third element considered, the change in growth of corporate profits, refers to inflation-adjusted quarterly profits data between 1970 and 2016. The author decided to consider profits movements because the correlation between those and the economic outlook, defined as a result of the mix of investments and industrial production, should present high values. In order to provide more detailed results, the author also considered the changes in values of a financial index, the US S&P500 total return index (SNP). The results of the analysis provided positive relations between the variables listed above and the ability to predict recessions in the short future; once again, the term spread, this time combined with other elements, proved to be a good predictor. Model 3 of Figure 7 has as explanatory variables: the change in profits, the credit spread, the term spread and the quarterly change in the S&P 500 total return index. Model 4 includes all of these except the change in profits. It is surprising how well these variables have performed in the past to predict recessions with almost a 90% of probability in each situation (some false positive cases are however present).

[14] E. Ergungor refered to Gilchrist and Zakrajšek when dealing with credit spread. In their opinion, the credit spread has to be decomposed in two elements, the Expected Default (ED) and the Excess Bond Premium (EBP). Only the EBP, whose variations is not induced by variations in default expectation, is the element with predictive power for future economic activity. For a deeper analysis, please refer to Gilchrist and Zakrajšek (2009).

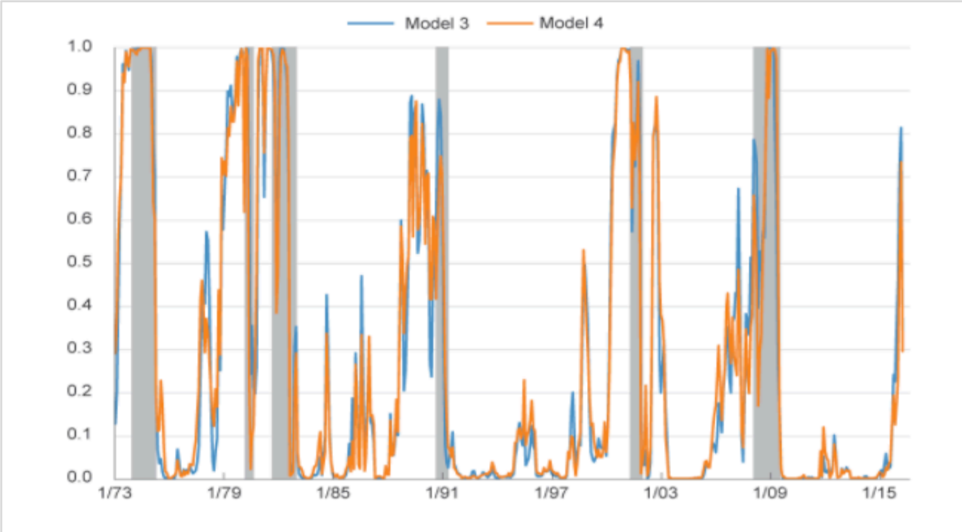


Figure 7: Emre Ergungor (2016)

Switching the focus on the European area, it is a bit more difficult to find studies that analyze the predictive power of the yield curve across European countries. Probably, the main reason lies in the structure of the European bank system compared to the United States one. In the US, the Federal Reserve System sets the interest rates periodically, and all the banks around the country take them as reference. In Europe, once the European Central Bank (ECB) announces the decision to cut or raise the level of rates, a transmission process takes place, and the central banks of each nations of the Eurozone will adapt their own rates accordingly. Therefore, to compare the predictive power of the curve in the two continents, it is necessary to analyze all the yields level in UE countries, which differ from nation to nation, according to the specific economic environment. This topic is firstly covered by authors such as Estrella and Mishkin (1996). After proving the power of the term spread in the United States, they analyzed its accuracy in France, Germany, United Kingdom and Italy. Their research, based on data between 1974 and 1994, provided evidence that the term spread revealed to be a good indicator for future recession also in Europe. More specifically, the probability forecast in Germany presented the highest value calculated, even higher than the United States'. The United Kingdom showed good results as well. France and Italy instead were the two countries where the term spread predicted recessions with lower accuracy, also revealing a few false-positive cases. The authors of the paper justified these weaker results affirming that it was more difficult to apply the model to France and Italy, due to the potential measurement error in recession dates, on which disagreements about the correct dates are evidenced[15].

[15] The authors relied on the recession dates identified by Allard (1994). For example, these dates are different from those provided by the Center for International Business Cycle Research (CIBCR) at Columbia University.

Bernard and Gerlach (1996) expanded the sample of the research, analyzing the predictive power of the term spread in eight countries: Belgium, Canada, France, Germany, Japan, Netherlands, United Kingdom and United States. To analyze such a wide sample and to provide cross-country evidence on the usefulness of the term spread, it is of fundamental importance to use the same criteria to define data, such as the recession periods. To obtain this objective, they relied on Artis et al (1995)[16], who dated recessions in twelve countries, starting in 1960, using equal criteria. The results of this analysis showed that the negative term spread acted as a fairly good predictor in each of the eight countries, with better accuracy in Canada, Germany and United States, and lower reliability in Japan. The difference among the results of different countries were mainly motivated by the variations of the financial markets regulations (especially between nations of different continents). The authors analyzed one other interesting factor as well: the impact of foreign spreads in predicting domestic recessions. Indeed, the United States and German spreads were used as regressors for the other countries of the sample, in addition to the specific country spread between short and long-term rates. Only these two variables were taken into consideration because of the relevance of these two nations on the global macroeconomic scenario. This decision was based on the high correlations that connect economic cycles across countries on global scale, and on the subsequent expectation that, adding these foreign spreads to the analysis, would have provided more precise results. However, the model developed, showed that little information was added when considering foreign spreads, resulting in a slight improvement of the yield curve ability to predict recessions (the only exception is Japan again, for which the German spread added relevant predictive content).

One last noteworthy research, which focused on European data, is Fabio Moneta's (2003) working paper for the European Central Bank. In the wake of previous works, especially the Estrella and Mishkin (1998) analysis, the author tried to test the predictive power of the term spread across Europe, but with an innovative technique. In order to have a comprehensive look of the rates, he retrieved each national data (of the most relevant European countries[17]) and then aggregated them to obtain a Euro area series, from 1970 to 2002. The predictive ability of the spread was compared with other economic variables, such as the quarterly growth rate of share price index (Eurostoxx) and the OECD (Organization for Economic Co-operation and Development) Composite Leading Indicator[18].

---

[16] Artis et al (1995) provided a methodology similar to the NBER way to define recessions. The difference is that the only series considered is the one inherent to changes in industrial production. Despite this difference, the dates identified by these authors were very similar to the NBER dates.

[17] The countries considered are Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Spain and Portugal.

[18] OECD Composite leading indicators are aggregate time series which show a leading relationship with the growth cycles of key macro-economic indicators.

The results showed that the term spread was once again the best predictor among the variables considered. The graph below (Figure 8) is extracted by the paper and shows the probability levels (y-axis) that anticipate recessions using the 10-years and 3-months term spread, based on data between 1970 and 2002 (x-axis) across Europe. The shaded areas describe the recession periods identified by the author. It is evident that an inversion of the yield curve acted as a strong predictor in all the four recession events, reaching very high probability levels. The only "mistake" made by the spread is in the 1992's recession, where it provided a slightly late warning.

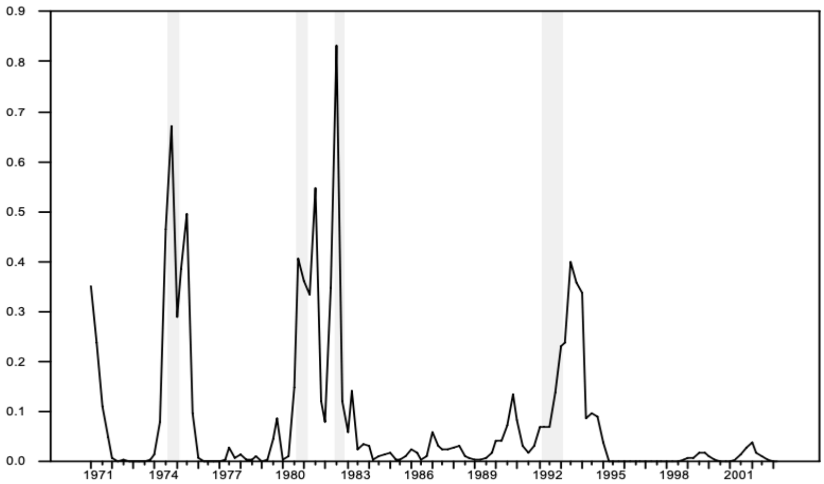


Figure 8: Fabio Moneta (2003), ECB

This section examined several works, providing evidence that a strong relation between the inversion of the yield curve and the beginning of a recession exists, both in the US, and with aggregated Euro data. The next chapter covers the core of this paper, providing the construction of the database and the model used to determine if this relationship hold true also when considering exclusively the Italian scenario.

4. METHODOLOGY

4.1 - Research question

As far as we are aware, studies or researches which focus exclusively on the Italian landscape and Italian yield curve are not present in the literature. To study the effects that different yield curve spreads have on the forecasted GDP, we propose a model whose objective is to translate the steepness of the yield curve into a probability of a recession in the future. To obtain this result, it is essential to explain the selection of two distinct elements, the components of the model that will be analysed: the spread, and the recession. Our first objective is to find the most effective spread, the one which proves to have the highest explicative power regarding a future recession. In order to understand which spread provides the highest explicative power, which will be higher or lower according to different forecasts, it is necessary to analyse the impact on GDP in different time horizons. This topic will result clearer reading the next few sections. The model that we propose is the "probit model", and will help us in testing the hypothesis of our work, listed below:

- Hypothesis: The term spread, based on the Italian yield curve, may be considered as a reliable predictor for impending national recessions or economic downturns.

The methodology and the approach that we conducted is partially based on previous works which focused on the same or similar subjects. Linear and non-linear statistical regression are the main tools used for most of economic and econometric research, because they provide quantitative guidelines able to specify the relationship that exists between variables, in our case between the term spreads and GDP decline, with a precise lag. However, besides applying the model that will show clear estimates, a general rule of thumb[19] has often been applied: the fact that recessions are anticipated by a negative term spread (without considering neither the exact lag nor the severity of the curve inversion). In the paragraph 4.2.2 'Defining the recession', an easy chart will show the rule of thumb applied to one term spread.

[19]The term "rule of thumb" refers to a principle with broad application. It refers to an easily learned and easily applied procedure or standard, based on practical experience rather than theory. This rule is not intended to be strictly accurate or reliable for every situation.

4.2 - Data collection

4.2.1 - Defining the Spread

The first step that we need to take for the implementation of the model is to define the yield spread. For the purpose of our work, the issuer considered is the Italian Government, which issues securities for national accounts' reasons. Although there is a general acceptance that the spread between the 10 years and 3 months interest rates may represent the most accurate tool to predict recessions, we have decided not to rely on this hypothesis, since most of the literature refers to US or EU studies. Therefore, the number of instruments considered in our analysis is eight, which are listed below. The first date in which there is availability for all these instruments is January 1995, hence our study covers the time frame between January 1995 and December 2019, providing a focus on more recent data and economic events.

- Interbank rate – 3 months: this is the instrument with the shortest maturity among our sample[20]. It is the only one not directly managed by Banca d'Italia (which issues all the other Italian treasuries), because it is represented by the EURIBOR[21]. EURIBOR values are only available from the beginning of 1999 (since the constitution of the European Union). Before the introduction of this European rate, the interbank rate was managed by each national central bank. Therefore, to fill the gap between January 1999 and January 1995, which is the starting date of our analysis, the values are provided by the 3-months interbank Italian rate (whose values have been used, together with the national interbank rates of all the other European countries, to calculate the EURIBOR rates in January 1999);

- BOT – 6 months, 12 months: "Buoni ordinario del Tesoro" (BOT) are the shortest debt instruments issued by Banca d'Italia[22]. Very low risk and very low return are the main characteristics of these tools, that are mainly used as a liquidity instrument instead of as investment opportunity. One important feature of these securities is that they are classified as

[20] To be more precise, the overnight, the daily and the 1-month interbank rates have shortest maturities. However, the 3-months interbank rate is the instrument with shortest maturity among the sample of this analysis.

[21] Euro interbank offered rate (EURIBOR) is a benchmark rate, daily calculated, which points the medium rate of the financial transactions between European banks.

[22] Banca d'Italia also issues 3-months BOT, however, very spare data were available about this instrument. For this reason, the interbank rate (EURIBOR) has been used as the instrument with the shortest maturity (the interbank rate was indeed used by other authors including Estrella, 1998 and Moneta, 2003).

zero-coupon-bond, no coupon is paid between their issuance and their expiration[23]. The rates of return provided by Banca d'Italia for these bonds are the gross allotment rates[24];

- CTZ – 2 years: "Certificati zero-coupon" (CTZ) are securities that have the same characteristics of the previous BOT. The only difference is represented by their longer maturity (hence, they show a slightly higher risk and higher return). The rate of return provided by Banca d'Italia for these bonds are the gross benchmark rates;

- BTP – 5 years, 10 years: "Buoni del Tesoro Poliennali" (BTP) are the medium and long-term securities issued by the Government. There are several BTP issued, including BTP with 5, 7, 15, 20, 30, and 50 years of maturity. These bonds are a concrete investment vehicle used by medium and long-term investors, who will receive coupon payments each semester with a fixed rate of return. BTP are traded on regulated markets. In our analysis only the BTP with 5 years and 10 years of maturity have been considered, because these are long-lasting instruments and because they are usually considered to be a good benchmark for medium and long-term projections. The rate of return provided by Banca d'Italia for these bonds are the gross benchmark rates.

Deciding which long-term debt instrument to select for the analysis may result easier, because of the availability of data and because of the fact that it is highly improbable to have a negative term spread when considering very long maturities (15, 20 years or larger). Instead, deciding the short-term instrument may result a bit more difficult, because of the wider variety of possibilities. In this work, we decided to take the 3-months interbank rate as shortest instrument, because both the overnight rate and 1-month rate could not represent proper indicators for market expectations, due to their direct control by European Central Bank. The fifteen combinations of spreads that we tested are the following: 6M - 3M, 1Y - 3M, 2Y - 3M, 5Y - 3M, 10Y - 3M, 1Y - 6M, 2Y - 6M, 5Y - 6M, 10Y - 6M, 2Y - 1Y, 5Y - 1Y, 10Y - 1Y, 5Y - 2Y, 10Y - 2Y, 10Y - 5Y (Y = years, M = months). As this list of combinations suggests, we decided to test also those combinations that are usually not used in these kinds of analysis (i.e. long and long-term spread and short and short-term spread).

The graphs below show two examples of the above-mentioned combinations. Figure 9 represents the spread between the 5-years BTP and the 6-months BOT, while Figure 10 represents

[23]The profit made by an investor buying these BOT derives from the difference between the price paid for the security at its issuance and the amount of money received at the expiration (the face value).
[24]Every return provided by Banca d'Italia is gross of allocation fees, administration fees and taxes.

the spread between the 10-years BTP and 3-months interbank rate. When the blue line overcomes the zero boundary, represented by the red horizontal line, the return of the short-term bond overcomes the return of the long-term bond of the combination, thus the spread turns negative (those points represent an inversion of the yield curve). For the graphical evidence of all the other spreads combinations, please refer to the Appendix attached at the end of the paper.

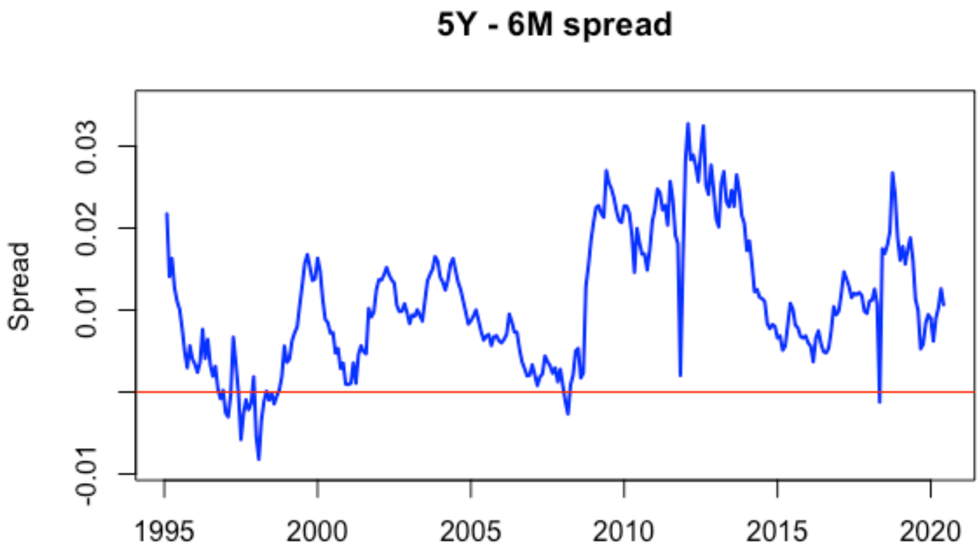


Figure 9: Author representation

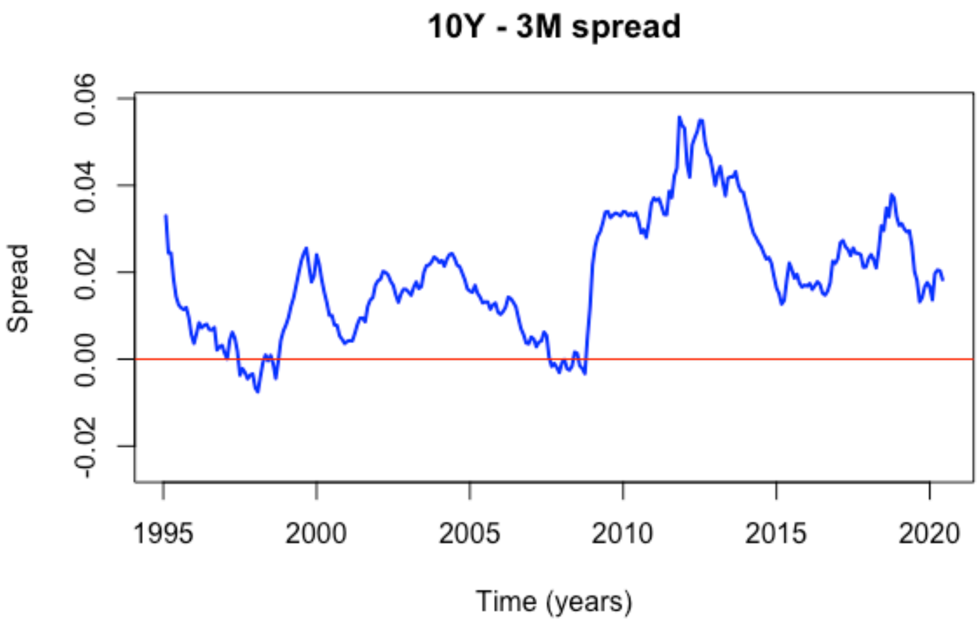


Figure 10: Author representation

4.2.2 - Defining Recession

Recessions' boundaries are always difficult to identify. Movements of national Gross Domestic Product are based on data that are often revised according to different calculations, hence finding exact dates that mark the beginning and the end of a recession is a challenging task. While the general agreement links the idea of a recession with an overall decline in economic activities, the European system does not provide a unique and unmistakable definition of this phenomenon. Some authors that conducted similar studies decided to rely on the definition provided by the National Bureau of Economic Research (NBER) of the United States[25]. However, given the Italian shade of this study, the definition provided by Borsa Italiana was taken as reference[26]. This institution provides a financial glossary where the major definitions of financial related topics and terms are described. Its proposed definition of recession states that these events: "represent a condition in which the national levels of production are lower than those that could be achieved by using completely and efficiently all the available production factors". Besides this general definition, that could be used with different interpretations, the so-called technical recession and the economic crisis are defined as well. A technical recession occurs when the real gross domestic product shows a negative variation for at least two consecutive quarters[27], while the latter occurs whether the negative variation of the economic output reports less severe values. For the purpose of our work, both definitions may be useful. In the next few sections of this paper the reader will be presented with the analysis of both situations. We chose to analyze both definitions according to two reasons. Firstly, because the goal of this paper is to understand if the term spreads have some predictive power in predicting an economic downturn, may it be very severe or slightly severe. Secondly, because considering exclusively the definition which identifies the recession as a decline occurred in two consecutive quarters, only few data of the sample would be taken as reference, since in the last twenty-five years few large recessions took place. The results provided by the analysis of both these phenomena will be showed and compared in Chapter five. The data used to calculate technical recessions and economic crises have been extrapolated by ISTAT (Istituto Nazionale di Statistica) database. The

[25]As seen in the previous chapter, this definition identifies a recession as "a significant decline in economic activity spread across the market, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales".

[26] Borsa Italiana S.p.A. is the national institution managing Italian stock exchange and financial markets. It is part of the London Stock Exchange Group.

[27] This definition has often been considered as the benchmark to use to identify recessions. The main reason is that short economic slowdowns may not turn into recessions and may be country specific. Recessions usually have a wider impact and wider duration (i.e. the Great Depression, the Great Recession among the others).

GDP measures provided by the database show quarterly values of Italian gross national product at current prices, and data have been seasonally adjusted[28].

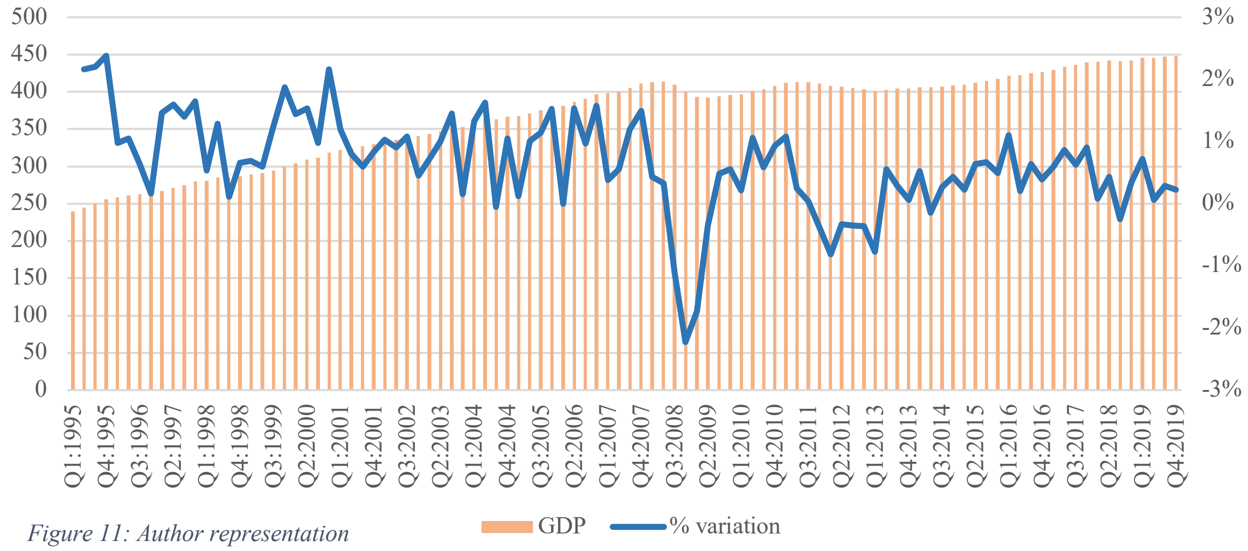


Figure 11: Author representation

Figure 11 shows the trend of Italian GDP between the first quarter of 1995 and the last quarter of 1999. The blue line represents the percentage variation year over year, and GDP data are expressed in billions of Euro. The interpretation of the graph is pretty straightforward, it is easy to point the two recessions that occurred in the timeframe of the analysis. The largest fall represents the Great Recession, which started in the second half of 2007 in the United States with the outbreak of the real estate bubble. Italy bounced back from this event only after several years, just in time to face the second severe recession deriving as a consequence of the previous one, the Sovereign Debt Crisis. These two events are the main economic shocks that Italian economy experienced in the last 25 years. The Great Recession negatively impacted the GDP for four consecutive quarters (Q3:2008 – Q2:2009), while the Sovereign Debt Crisis made the GDP fall for six consecutive quarters (Q4:2011 – Q1:2013).

Now that both the spread and the recession have been described, it is possible to apply the previous mentioned rule of thumb, simply considering one specific term spread trend and the GDP changes during the time horizon of the analysis. Figure 12 provides an example of how a negative term spread acted as a fairly good predictor for the two Italian recessions. In this case, the spread

[28] GDP with current prices refers to the so-called nominal GDP. The alternative is the real GDP, which is calculated with constant prices, selecting one year as benchmark. Seasonally adjusted data are net of seasonal fluctuations. It means they are corrected for calendar effects, when these are significant. The procedure to correct for this kind of effects is made using the Tramo-Seats procedure.

between the 2-years BTP and 6-months BOT turned negative just a few months before the start of both recessions (marked by the shaded areas in the graph). Before the outbreak of the Great recession, the spread reached a negative value of almost 0,5%, while before the second economic downturn, the difference was about 0,1%. Moreover, this spread turned negative in other occasions. For example, it turned negative in early 2000s, a period that is marked by the so-called Dotcom Bubble, and in the middle of 2018, when Italian GDP experienced a strong slowdown, when it declined between the second and third quarter of 2018[29].

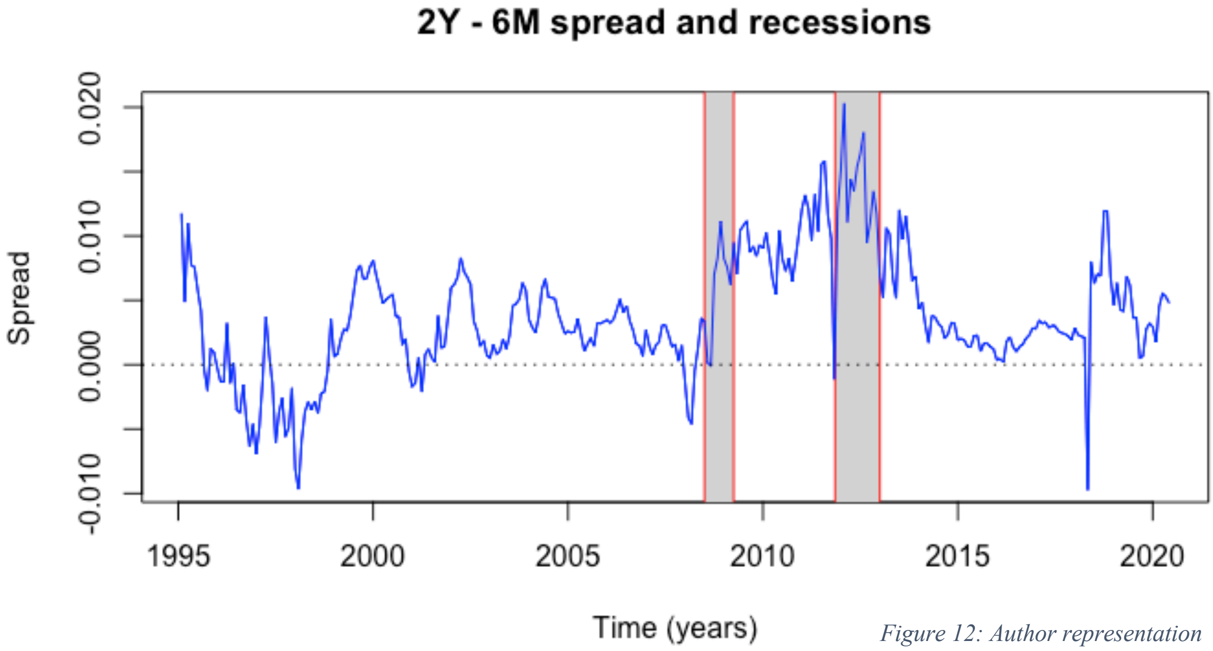


Figure 12: Author representation

4.2.3 - Other indicators

Chapter five will provide the results of the models based on the spreads combinations and recessions definitions explained in the last two paragraphs. However, before explaining the functioning of the probit model and the discussion of the results, this section describes other indicators that we selected to use alongside the analysis of the term spread to predict recessions in the near future. Therefore, our analysis will be conducted also on the predictive ability of these measures, in order to understand if Italian term spread has more, or less explanatory power compared to these variables. The measures analyzed are the Composite Leading Indicator, the Consumer Confidence

[29] The Dotcom Bubble is the market crash that occurred in early 2000s because of the bubble inherent to the developments of information technologies.

Index and the Business Confidence Index[30]. Each of these indicators is based on a long-term average value of 100, therefore, downward and upward movements will occur accordingly to the information set on which they are built. Below, a brief description of these indicators is provided.

- Composite Leading Indicator (CLI) - Probably the most used indicator to anticipate fluctuations of economic activity at short term, the CLI is a composite index which monitors the movements of different trends occurring in a specific nation, providing early signals of turning points in business cycles. For past recessions, this indicator usually performed well in estimating periods of economic growth and decline. Different variants of the CLI exist, according to the set of variables and geographical zone considered. In our analysis, where the indicator provided by OECD was taken into consideration, the trends considered are time series which exhibits strong relationships with movements of GDP and that cover a large range of key statistics. For Italy, the series that have been used are: the consumer confidence indicator, manufacturing orders book, deflated order for total manufactured goods, manufacturing productions, imports from Germany and Consumer Price Index[31] (Figure 13.1 shows the Italian CLI vs European and US CLI between 1995 and 2019);

- Consumer Confidence Index (CCI) – This index focus on consumers' and families' behaviors. These categories are asked to provide their opinion regarding their expected financial situation, their sentiment about the overall economic health, unemployment topics and savings trends[32]. Based on a long-term average value of 100, if the indicator shows greater values, it means that the general confidence of these subjects about future economic developments is increasing. This situation occurs when these people are more prone to spend for large purchases and less willing to save money. On the other hand, if the indicator shows values lower than the average, it means that these categories hold worries about future economic outlook and they probably will reduce consumption in favor of more savings (Figure 13.2 shows the Italian CLI vs European and US CLI between 1995 and 2019);

---

[30]CLI, CCI and BCI data are provided by the Organization for Economic Co-operation and Development (OECD). This organization monthly studies these indicators for a large set of countries (40 countries all over the world). Although aggregates indicators of various nature are available and may have different predictive ability, the focus of our analysis is on Italian indicators, with data available between January 1995 and December 2019.

[31]Consumer price index is a statistic which measure the average changes in price levels that household pay for a market basket of consumer goods.

[32] For more information about the surveys conducted for both the Consumer Confidence Index and Business Confidence index, please refer to the document "The Joint Harmonised EU Programme of Business and Consumer Surveys", published by the European Commission, and available on OECD website.

- Business Confidence Index (BCI) – This indicator is built with a similar approach to the previous index, but as the name suggests, the target of the analysis is composed by enterprises and players of the industrial sector. These subjects provide their opinion regarding developments in production, order and stocks of finished goods, selling prices, export conditions, and competitive positions. This index is useful to understand which is the current trend regarding the production side of the economy, that usually represents a good benchmark to monitor output growth and business cycles. If the value of this index overcomes 100, there is general confidence regarding a future positive outlook, while values below 100 indicate a potential economic downturn in the short-term, since data about national production show a negative trend (Figure 13.3 shows the Italian CLI vs European CLI between 1995 and 2019).

These three indicators are usually defined more qualitative than quantitative. Nevertheless, they are often considered to be good predictors of future economic output, since they put together the information that different series of data provide, from market research surveys (in the case of the CCI and BCI), to constructed data series (in the case of the CLI). Lastly, one last indicator has been considered. Given that the term spread predictive ability would be partially explained by the correctness of expectations of bond investors, it would not be appropriate to exclude by our analysis the consideration of the stock market. Financial markets are good aggregators of information because of the broad participation of different players. Investors of financial markets always take into consideration their beliefs about future economic forecast when deciding investment opportunities. Therefore, their behaviors, drove by their expectations, should be reflected in shares' prices. This occurs because stock prices also reflect expectations about future dividend payments, and whether stock markets investors are worried about impending recession or economic downturns, they will be less willing to invest in stocks whose dividend may be uncertain. If all market participants have the same feeling and act as consequence, this would eventually be driving downwards the share price, turning the share price (hence, the share index) into a potential predictor for future recessions. In order to be coherent with our analysis, the share price index subject of our study will be the FTSE MIB[33], the most significant Italian index. This index refers to the basket made up by the 40 most

[33] FTSE MIB: "Financial Times Stock Exchange Milano Indice di Borsa".

capitalized and liquid companies whose shares are daily traded on the stock exchange markets (MTA and MIV[34]). Figure 13.4 provides the FTSE MIB trend between January 1997 and December 2019[35]. For each representation of the OECD indicators, the fat blue line refers to Italian values, the dotted green line refers to the average US values, while the dotted red line refers to the average EU values.

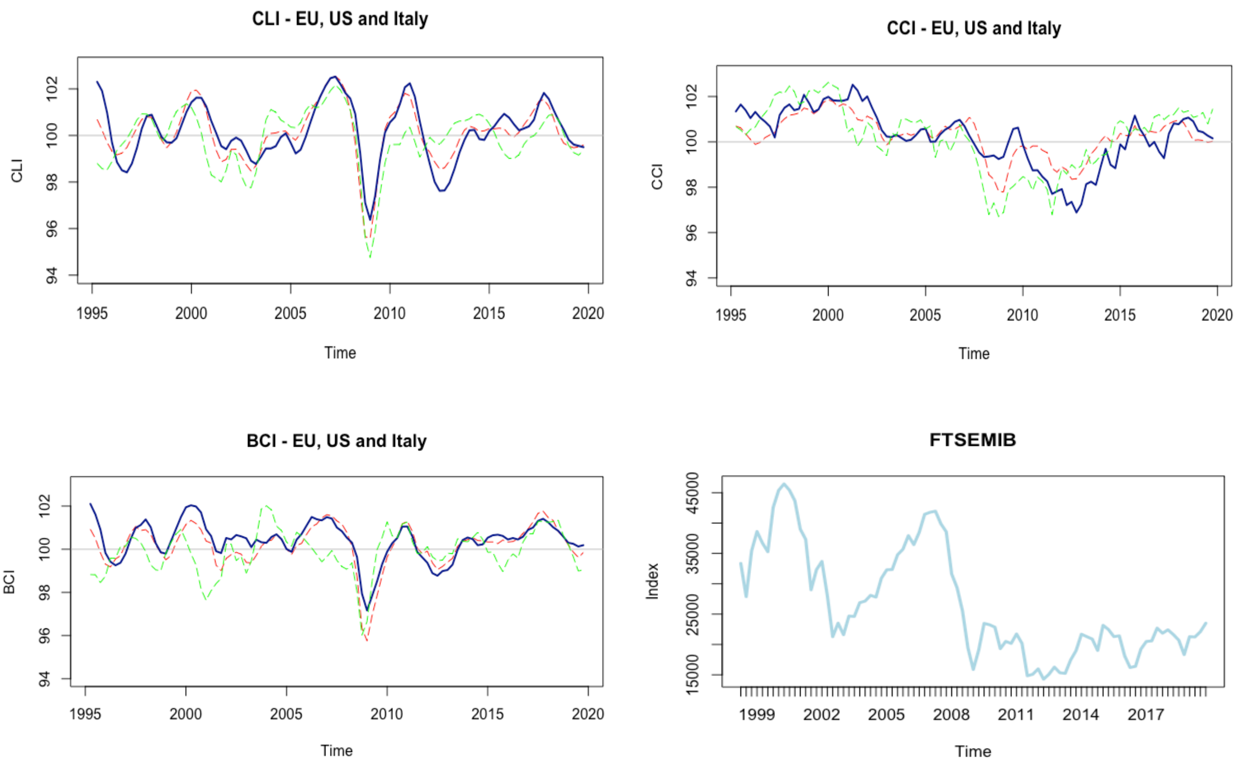


Figure 13: From top left to bottom right: 13.1 – 13.2 – 13.3 – 13.4. Author representation

---

[34] MTA: "Mercato telematico azioni", MIV: "Mercato degli Investment Vehicles". These two stock exchange markets differ for the features of the instruments that can be traded on each of them. The third Italian stock exchange is the AIM: "Mercato Alternativo del Capitale", created for companies with smaller size.

[35] In section '4.3 – The analysis: the probit model' we will explain why the time horizon considered for the FTSE MIB is slight shorter compared to the other indicators.

4.3 - The analysis: the probit model

The probit model is a binary outcome model. It is classified as a non-linear regression model specifically designed for binary dependent variables. In these models, dependent variables can only assume two determined values, usually represented by the values ‘yes’ and ‘no’ (that are generally turned into ‘one’ and ‘zero’ when it comes to statistical modeling). Since a regression with a binary dependent variable (Y) models the probability that y = 1, it is reasonable to adopt a non-linear formulation that forces the predicted values to assume values between zero and one[36]. This kind of models are among the most used in applied economics[37]. Non-linear estimation procedures are useful to estimate almost any kind of relationship between one (or a set) of independent variables and one dependent variable. Before going forward, a brief explanation of the difference between linear and non-linear regressions may help the reader to have a clearer understanding of these concepts. Linear regressions are described by equation [1]:

$$y = \ (\beta_0 + \beta_1 x) + e \qquad [1]$$

where y is the dependent variable, x is the independent variable, β_0 is the coefficient of the constant, β_1 is the coefficient of the independent variable, and e is the error. This model is defined as a linear regression because it estimates the line which can explain in the best way the relation between the variables considered. The components of this line are β_0, which represents the intercept, and β_1 which represents the slope. Although linear regressions are often used for empirical analysis, our study needs to use a different approach, since the relationship between the independent and the dependent variables is not linear. The equations of the probit model, which is a non-linear regression model, are the following. Equation [2] describes the model with one independent variable X, while equation [3] describes the model with more independent variables.

$$\Pr(y_{t+k} = 1|X_t) = \ \Phi\ (\beta_0 + \beta_1\ x_t) \qquad [2]$$

$$\Pr(y = 1|x_1, \dots, x_n) = \ \Phi\big(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p\big) \qquad [3]$$

[36] It would be possible to use standard multiple regression analysis, however, the values predicted would fall outside of the range 0-1, providing values that are not valid for our purposes (the probability levels cannot be superior than 1 or inferior to 0). Therefore, it is necessary to proceed with a restriction, to fit the predicted values inside this range.

[37] Probit models are often used jointly with another similar methodology: the logit models. These two applications represent the two most used binary outcome models. The difference between them lies in the distribution functions used to model probabilities: the standard normal distribution is used for the probit model while the logistic distribution is used for the logit model. However, results of the two models are usually very similar.

In this model, Y represents the binary variable, X represents the independent variable and Φ represents the cumulative distribution function of a standard normal distribution. Binary outcome models, instead of modeling the value of Y as a continuous variable, they will model the probability of Y = 1, since this variable can only take two values[38]. Therefore, the distinctive characteristic of this regression is that it does not model the value of Y, but the probability that Y would take the value of 1, according to the observation of the independent variable X considered. Concerning β_0 and β_1, their interpretation is not straightforward as the interpretation of the coefficients of a linear regression model. Indeed, in the probit model, an increase in the value of X leads to an increase or a decrease in the likelihood that the value of Y will be equal to 1. In other words, an increase in X, makes the outcome of 1 more or less likely. This is different from typical regression models, where higher values of X are translated into higher or lower values of Y. Another theoretical concept that differs from the typical regression models, is the evaluation of the goodness of fit measures. When estimating probit models, one of these measures is represented by the percentage of correctly predicted values. This concept provides a quality evaluation of the model calculated, because it shows how many observations were correctly predicted according to the actual values. The proportion of the true predictions to the total predictions defines this evaluation method (probit models with good-to-high explanatory power are expected to provide high values for this proportion, usually not below 70%). The second measure that it is useful to estimate the statistical goodness of our model is the so-called Pseudo R-squared[39], a different version of the traditional R-squared. This variant is defined by equation [4]:

$$Pseudo\ R^2 = 1 - \frac{L_{ur}}{L_r} \qquad [4]$$

Where L_{ur} represents the log-likelihood of the unrestricted regression model, and L_r represents the log-likelihood of the restricted regression model (restricted because it takes into consideration only β_0 as predictor - the constant, so that each estimation is predicted with the same probability of "success")[40]. The interpretation of this measure is the following: if the independent variable of our model has low or no explanatory power, the unrestricted model value will show similar (or the same) values of the restricted model, hence the ratio will be close to 1, and R-squared closed to zero (and

[38] A continuous variable may take values between [-∝,+∝]. Since the probability levels must fall in the range of [0,1], the continuous variable is not suited for the probit model.
[39] Also called the McFadden R-squared.
[40] The log-likelihood method is a logarithmic transformation of the likelihood function. It is mainly used because it is a convenient method used to simplify the calculations of the maximum likelihood estimation. The unrestricted regression model refers to the model where each parameter is kept as it is, therefore, no constraints have been applied. In this case, the probit model.

vice versa). Pseudo R-squared is useful for our calculations, because it will be one of the measures used to compare the models, and therefore to understand which spread has the strongest impact for our purposes. Values of this measure are however to be considered with a different approach compared to typical values of R-squared estimations. Indeed, there is a general consensus that the estimated model provides a good fit of the analysis if the R-squared value calculated falls into the range of [0.2 - 0.4] [41]. The predicted values computed by the models are based on a classification rule, that assigns to the binary dependent variable the value of 1, if the estimated probability for the specific observation is higher than a defined threshold. To help us in understanding the thresholds that maximize the correctly predicted values for each model, we will compare ROC (Receiver operating characteristic) and AUC (Area under the ROC curve, also called AUROC) estimations. The intuition behind the ROC curve is to repeat the classification rule for each value of the threshold. Therefore, lowering the threshold will lead to a higher number of values classified as positives. The AUROC instead represents the area under the ROC curve, and it is a useful measure to compare different models' estimations. The more the AUROC level is closer to 1, the higher is the predictive ability of the explanatory variable. Once that the percentage of correctly predicted estimations will be provided, one last step will be conducted. Indeed, to have an appropriate analysis, it is important to differentiate between false positive and false negative rates. The calculation of these rates is based on the comparison between correct and incorrect predictions according to the actual observations. These ratios are calculated from the construction of a cross-tabulation based on predictions and actual observations[42]. The false positive rate refers to the number of times that the independent variable predicted a value of 1, when in the reality the actual observation was the opposite (zero). On the other side, the false negative rate refers to the number of time that the independent variable predicted a value of zero, but in the reality the actual observation was the opposite (one). These concepts will result clearer when in the next chapter some concrete examples will be provided.

After this dutiful introduction to non-linear regression models, we can now consider the inputs of our analysis, recalling equation [2]. The input to this calculation is the value of the term spreads observed at time *t*, which represent the variable X. Variable Y, which is the binary variable, represents the recession. To be more precise, it assumes the value of 1 when a recession has occurred in time $t + k$, and 0 otherwise[43]. Since our purpose is to focus on the predictive ability of the term spread, it is necessary to apply different lags. The output of the model refers to the probability of the start of a

[41] McFadden himself suggested this range of values as a measure of a good fit. In addition, please refer to Louvier et al, 2000: "Stated Choice Methods".
[42] Please refer to the Appendix where both the description and an example of this cross-tabulation are provided.
[43] *K* refers to the lagged quarter. In our analysis, *k* will assume values from one to eight.

recession in a specific quarter, from the informative set available in a specified quarter before. In our analysis, we took into considerations term spread impacts on eight different quarters, so that the results will provide a comprehensive look over the predictive ability of these elements. Moreover, since our purpose is not to focus exclusively on recessions, but on economic crises as well, the two definitions of recessions mentioned in the previous chapter have been considered. Therefore, having at our disposal 15 combinations of term spreads, 8 lagged quarters and 2 definition of recessions, our analysis will be conducted on 240 probit models.

For what concerns the analysis of the four alternative indicators (CLI, CCI, BCI, and FTSEMIB), the framework is pretty similar to the one just explained. Regarding the Composite Leading Indicator, the Consumer Confidence Index and the Business Confidence Index, the quarterly values will represent the independent variable X, while the recession will still represent the binary dependent variable Y. For all these indicators we considered eight quarters of lag as well. The analysis of the FTSEMIB is also similar, however, the time horizon is slightly different, since historical data about this index are scarce[44]. To study its predictive power, we will take into consideration quarterly values of the index, extrapolated from the share price index's values between January 1997 and December 2019. In this case, the independent variable X will be the FTSEMIB value, while the recession will be, once again, the binary dependent variable Y. Next chapter provides a summary of the results obtained, alongside with a discussion about the results.

[44] Although the index origin dates back to December 1992, the FTSEMIB that we know today only exists since June 2009. From 1992 to 2003 the index was called MIB30 (and between 2003 and 2009, S&P MIB), and the underlying basket of companies was different compared to the current one. Therefore, considering these data would lead to an incorrect analysis. The current FTSEMIB was born with the merge between Borsa Italiana and London Stock Exchange Group. Although data used for this analysis are less compared to the study of the term spreads and the other three indicators, having the historical series starting in 1997 enable us to include the two recessions of our sample, providing anyway the possibility to study the predictive power of this indicator.

5. RESULTS

This section provides the numerical evidence of our empirical analysis with the support of a few tables and graphical interpretations, that will show the results obtained by the probit models' estimations. The first two paragraphs (5.1 and 5.2) will present the results considering the term spread as predictor, according to the two different definitions of recessions mentioned in the previous chapter. Then, the third paragraph (5.3) will provide a comparison among the results of the predictive powers obtained by the analysis of the four alternative indicators. For the sake of clarity, the software used for our purposes is Gretl, an open-source statistical software often used for econometric analysis[45].

5.1 - Predicting an economic crisis

Table 1 provides some of the relevant data extrapolated by the results of the first 120 probit models' estimations. In this first analysis six spreads have been analyzed, for a total of 15 combinations[46]. The recession was defined as one quarter of decline in GDP, therefore, as explained before, it would be more appropriate to refer to it as economic crisis. Considering this interpretation, the dependent variable assumed the value of 1 thirteen times among the sample (Q3:2004, Q3:2008 – Q2:2009, Q4:2011 – Q1:2013, Q2:2014 and Q3:2018). In the table, the rows provide the values of the pseudo r-squared and p-value calculated for each spread according to the different lags (from 1 to 8 quarters). It is important to mention p-values for each estimation because they allow us to understand the probability to reject the null hypothesis. To be clearer: the null hypothesis puts the coefficient of the spread, β_1 , equal to zero. With the coefficient equal to zero, the spread would not have any predictive power on the recession, and the analysis would be statistically not significant, because values of the term spread would not have any impact on recessions (no relationship between the independent variable X and the dependent variable Y). To decide if the analysis is statistically significant, p-values must be compared with a level of significance. In the table below, p-values are compared with three different levels: 0.01, 0.05, and 0.10. When the model shows a good fit and the null hypothesis is rejected, the symbol of the star is provided next to the p-value numbers. The

[45] This name is an acronym for "Gnu Regression, Econometrics and Time-series Library". This software was initially created to analyze historical series of data, proving to be a great tool for our kind of study. Gretl has been cited and reviewed several times among different Journal of Econometrics. Alongside with Gretl, we used R for several graphical representations.

[46] A reminder of the fifteen combinations analyzed: 6M - 3M, 1Y - 3M, 2Y - 3M, 5Y - 3M, 10Y - 3M, 1Y - 6M, 2Y - 6M, 5Y - 6M, 10Y - 6M, 2Y - 1Y, 5Y - 1Y, 10Y - 1Y, 5Y - 2Y, 10Y - 2Y, 10Y - 5Y.

difference between one, two, or three stars derives from the significance level considered to compare the two probabilities (one, two or three stars correspond respectively to 0.10, 0.05 and 0.01 significance level). The best models are those where p-values are lower than 0.01, meaning that there is a probability lower to 1% that the term spread does not have any impact on the recession (best models are where three stars are showed '***'). Previous chapter described pseudo r-squared in details, and the values range that can be used to determine if a model provides good estimations[47]. The first impression that stands out from the table below is that Italian term spreads are fairly good predictors when the lag of the analysis is restricted to one quarter. The column corresponding to the '1 lagged quarter' is indeed the only column where each of the term spread provides good estimations. All of them, with the exception of the 10Y – 2Y and 10Y – 5Y, show p-values that are significant with at least a 0.05 level of significance, while most of them are significant with a level of 0.01. The two spreads that provide the highest pseudo r-squared are the 1Y – 6M and the 2Y – 6M spread (the former with a pseudo r-squared of 0.2275 and p-value of 0.0004, the latter with a pseudo r-squared of 0.2199 and p-value of 0.0002). Pseudo r-squared values of 1Y – 3M, 2Y – 3M, 5Y – 3M and 5Y – 6M term spreads are pretty high as well, showing numbers larger than 0.15. These first results allow us to understand that the best combinations are those combining a very short term (the 3-months interbank rate or the 6-months BOT) with a medium-long term (2-years CTZ or 5-years BTP, but not the 10-years BTP). Switching the focus on a higher number of lagged quarters, the results are way different. The only two time horizons which show a fairly good level of statistical significance are the '2 lagged quarters' and '3 lagged quarters' cases. Indeed, for both of them, the 2Y – 3M, the 5Y – 3M and 1Y – 6M spreads present low p-values, even though pseudo r-squared are quite low. Another conclusion that we may derive from these results is the low predictive abilities of the term spread when the quarters are more than four: with the exception of the two best spreads (1Y – 6M and 2Y – 6M, that show quite good estimations for all of the lags), each model provides values that are not statistically significant when the lags considered are more than four. Moreover, the spreads combining medium and long terms (such as those of the last six rows of the table) show very low pseudo r-squared values and very high p-values, thus we can exclude them from the basket of spreads with predictive power for future economic crises.

[47] The model could be defined a good predictor if the pseudo r-squared values fall inside the range [0.2 – 0.4].

Table 1: Results

		1 lagged quarter	2 lagged quarters	3 lagged quarters	4 lagged quarters	5 lagged quarters	6 lagged quarters	7 lagged quarters	8 lagged quarters
6M - 3M	Pseudo R^2	0.099678	0.07991	0.072873	0.042655	0.046492	0.012783	0.012124	0.027545
	p-value	0.0111**	0.0209**	0.0275**	0.0788*	0.0678*	0.3136	0.326	0.1444
1Y - 3M	Pseudo R^2	0.154451	0.093991	0.098761	0.050064	0.060539	0.026434	0.016136	0.049805
	p-value	0.0038***	0.0144**	0.0125**	0.0579*	0.0386**	0.1463	0.2543	0.0496**
2Y - 3M	Pseudo R^2	0.176826	0.104972	0.097695	0.057706	0.059407	0.039669	0.032078	0.056456
	p-value	0.0012***	0.0079***	0.0098***	0.0394**	0.0371**	0.0787*	0.1131	0.0402**
5Y - 3M	Pseudo R^2	0.184896	0.096502	0.074175	0.050046	0.03681	0.029188	0.027745	0.036562
	p-value	0.0005***	0.0088***	0.0202**	0.0528*	0.0964*	0.1357	0.1467	0.0997*
10Y - 3M	Pseudo R^2	0.144501	0.072544	0.046826	0.035137	0.029018	0.02264	0.023155	0.024894
	p-value	0.0017***	0.0217**	0.0621*	0.1045	0.1404	0.1905	0.1881	0.1741
1Y - 6M	Pseudo R^2	0.227536	0.082932	0.124075	0.046726	0.070768	0.055686	0.01814	0.09255
	p-value	0.0004***	0.0161**	0.0041***	0.0647*	0.0245**	0.0402**	0.2349	0.0088***
2Y - 6M	Pseudo R^2	0.219885	0.090697	0.09326	0.054426	0.052337	0.076151	0.053376	0.078541
	p-value	0.0002***	0.0107**	0.0099**	0.0454**	0.0498**	0.0198**	0.0489**	0.0191**
5Y - 6M	Pseudo R^2	0.168826	0.065062	0.044787	0.034169	0.016013	0.030979	0.0286	0.026253
	p-value	0.0008***	0.0286**	0.0679*	0.1095	0.2713	0.1281	0.1443	0.1626
10Y - 6M	Pseudo R^2	0.103189	0.037043	0.017344	0.017348	0.010264	0.019288	0.019991	0.013002
	p-value	0.0072***	0.0964*	0.2525	0.2529	0.3788	0.2294	0.2221	0.3248
2Y - 1Y	Pseudo R^2	0.060446	0.033849	0.014642	0.022206	0.007168	0.045357	0.06465	0.022193
	p-value	0.0342**	0.1118	0.2927	0.1972	0.4617	0.0683	0.0311**	0.198
5Y - 1Y	Pseudo R^2	0.078773	0.031732	0.009939	0.01598	0.000989	0.012719	0.022958	0.004813
	p-value	0.0164**	0.1219	0.3841	0.2715	0.7841	0.3256	0.1887	0.547
10Y - 1Y	Pseudo R^2	0.042933	0.014782	0.001161	0.006157	0.000745	0.00726	0.014776	0.001458
	p-value	0.0739*	0.29	0.7661	0.4945	0.8122	0.4588	0.2923	0.741
5Y - 2Y	Pseudo R^2	0.075014	0.024837	0.006177	0.010468	0.000004	0.002138	0.006526	0.000365
	p-value	0.019**	0.1696	0.4921	0.3726	0.9865	0.6871	0.4823	0.8687
10Y - 2Y	Pseudo R^2	0.030633	0.007758	0.000002	0.002274	0.000031	0.001277	0.004569	0.000001
	p-value	0.13	0.4426	0.9914	0.6779	0.9617	0.7563	0.5579	0.9923
10Y - 5Y	Pseudo R^2	0.000504	0.000151	0.009486	0.000931	0.000193	0.000335	0.001794	0.000628
	p-value	0.8441	0.9146	0.398	0.7906	0.9039	0.8739	0.7137	0.8287

These initial results let us understand which spread may be initially considered as more significant. We can now estimate the ROC and AUROC, the two methods briefly explained in the previous chapter. The ROC curve presents on the x-axis the false positive rate, and on the y-axis the true positive rate, and therefore it tracks the relationship between these two rates according to the threshold applied (the false positive rate was explained above, while the true positive rate refers to the number of times that the model forecasted a recession when it actually occurred, hence when it correctly estimated a recession) [48]. Estimating the ROC will help us in understanding which threshold

[48] For the graphical evidence of the ROC and AUC calculated for all the indicators please refer to the Appendix.

level is optimal for each spread. The threshold that will be defined optimal is the level that maximize the percentage of correctly predicted values. Table 2 provides a summary of the ROC and AUROC estimations of the best spreads provided by previous results. To select these spreads, we considered the best results in terms of pseudo r-squared and p-values. Although the results for the 10Y – 3M spread were slightly weaker compared to the others, we decided to include it in the following analysis, because we wanted to test its accuracy and understand whether the most-used spread in the literature may be considered a good estimator in Italy as well. The lagged quarters have been reduced to two, because results for more extended time horizons proved to be statistically not significant.

Table 2: Results

		1 lagged quarter	2 lagged quarters
1Y - 6M	AUROC	0.795 (0.0652)	0.644 (0.0909)
	Max correctly predicted	89.8%	89.7%%
	Threshold	0.392	0.493
2Y - 6M	AUROC	0.813 (0.0721)	0.703 (0.098)
	Max correctly predicted	90.8%	88.7%
	Threshold	0.433	0.304
5Y - 6M	AUROC	0.784 (0.0781)	0.684 (0.0955)
	Max correctly predicted	87.8%	87.6%
	Threshold	0.403	0.344
5Y - 3M	AUROC	0.753 (0.0938)	0.661 (0.108)
	Max correctly predicted	90.8%	90.7%
	Threshold	0.462	0.349
10Y - 3M	AUROC	0.744 (0.0974)	0.664 (0.11)
	Max correctly predicted	90.8%	89.7%
	Threshold	0.354	0.354

As explained in the previous chapter, AUROC refers to the area under the ROC curve, and the closer the AUROC value is to 1, the higher is the accuracy of the model. Indeed, this method represents a performance measure for the accuracy of the estimations. AUROC values can range between 1, perfect accuracy, and 0.5, no accuracy (the value in brackets represent its standard error). In this table, we may see that the different spreads reach their highest percentage of correctly predicted values with different thresholds, ranging from a low level of 0.304 (of the 2Y – 6M spread with two lagged quarters) to the medium value of 0.493 (of the 1Y – 6M spread with two lagged quarters). We can notice that for each term spread, both the AUROC values and the percentage of correctly predicted values are higher when the time horizon considered is one lagged quarter, hence confirming that best time horizons are the shortest. Comparing the AUC estimations, we may notice that the highest value is provided by the 2Y – 6M term spread (AUC equal to 0.813). However, we may also point out that the results provided in Table 2 are quite similar among the spreads considered. The percentages of correctly predicted values show numbers that never undergo the level of 87,6%[49]. Once that we provided the reader these percentages of correctly predicted values, it is necessary to shed light on the important false positive and false negative rates. The values of the previous table referred to the number of times that the independent variable (the spread) correctly predicted the value of the binary dependent variable, hence, they represented the fraction of correctly predicted estimations over the total number of observations. However, in order to provide a more appropriate analysis, it is dutiful to calculate the false positive and false negative rates. These rates are based on the comparison between correct and incorrect predictions according to the actual observations. The false positive rate refers to the number of times that a recession is forecasted, when in reality no recession occurred. On the other hand, the false negative rate refers to the number of times that no recession was forecasted, when in reality it occurred. Table 3 shows a summary of these ratios, calculated for the best predictors and best horizons. The very low false positive rates that we notice for most of the predictors are a good indicator, because it means that no recessions have been forecasted when they did not occur. At the same time, false negative rates provide larger values, meaning that in some occasions, although a recession occurred, the model did not succeed in predicting it correctly.

Table 3: Results

		1 lagged quarter	2 lagged quarters
1Y - 6M	False Positive rate	0.0233	0.0000
	False Negative rate	0.6154	0.8462
2Y - 6M	False Positive rate	0.0233	0.0349
	False Negative rate	0.6154	0.6923
5Y - 6M	False Positive rate	0.0349	0.0000
	False Negative rate	0.6923	1.000
5Y - 3M	False Positive rate	0.0000	0.0000
	False Negative rate	0.7692	0.6923
10Y - 3M	False Positive rate	0.0233	0.0000
	False Negative rate	0.6154	0.7692

[49] Probit models with percentage of correctly predicted values of over 70% are usually considered to be good models.

5.2 - Predicting a recession

The second analysis that we conducted has been carried out on the same fifteen spread combinations but with a different definition of recession. In this case, the recession was defined as the decline in GDP's growth occurred for two consecutive quarters. This is the definition that most of the time was taken as reference by other studies of the same subject, because such a large decline may not be caused by momentaneous economic trends. Limiting the recession with this constraint helped us in clearly identifying the two large economic declines that hit the Italian economy in the last twenty-five years. With this definition, the binary dependent variable assumed the value of 1 only eight times (Q4:2008 – Q2:2009 and Q1:2012 – Q1:2013) compared to the thirteen times of the previous test.

Table 4 in the next page shows the results of the second round of 120 probit model's estimations. Reading the data on this table, it is evident that the majority of the results obtained in the previous analysis is confirmed, while there are a few other considerations to be mentioned. Once again, it is clear that a lower number of lags provides higher explanatory power for future recessions, because values calculated with more than four lags show very low significance levels (the only two exceptions are the 1Y – 6M and the 2Y – 6M, for which p-values provide good fit for five, six and seven lagged quarters as well). Coherently, the first column of the table ('1 lagged quarter') is again the time horizon where almost all the spreads provide some significance level (apart from the spreads in last two rows, every estimation is significant at least with a 0.10 level with 1 lagged quarter). One comment that is worthy to point out, is that compared to the previous study, the number of spreads with a pseudo r-squared value included in the range of [0.2 – 0.4] increases significantly. In fact, in the previous analysis only two combinations fulfilled this requirement, the 1Y – 6M and 2Y – 6M spreads, both with one lagged quarter. In this analysis, instead, the results obtained when considering a tighter definition of recessions show that seven times the pseudo r-squared values reached a level higher than 0.2, hence turning these term spreads into better predictors. The five spreads that showed pseudo r-squared higher than 0.2 are: 2Y – 3M (with two lagged quarters), 5Y – 3M (with one lagged quarter), 1Y – 6M (with both one and two lagged quarters), 2Y – 6M (with both one and two lagged quarters) and finally the 5Y – 6M (with one lagged quarter). One thing in common with the previous analysis is that, once again, the highest explanatory power is provided by the same spreads: the 1Y – 6M (pseudo r-squared: 0.2672 – p-value: 0.001) and 2Y – 6M (pseudo r-squared: 0.265 – p-value: 0.008). Similarly to previous results, the worst predictive ability is provided by those spread which combine medium and long-term rates (i.e. 10Y – 2Y) or long and long-term rates (i.e. 10Y – 5Y). The widely used 10Y – 3M spread, pointed out by several authors to be the best predictor for recessions when studying movements of the yield curve, did not revealed to be the Italian spread with the highest explanatory power. Nevertheless, pseudo r-squared and p-value for this combination show quite good results, respectively 0.177 and 0.0036, proving to be statistically significant with a 0.01 level.

Table 4: Results

		1 lagged quarter	**2 lagged quarters**	**3 lagged quarters**	**4 lagged quarters**	**5 lagged quarters**	**6 lagged quarters**	**7 lagged quarters**	**8 lagged quarters**
6M - 3M	Pseudo R^2	0.10152	0.13614	0.152637	0.064302	0.064746	0.009948	0.006093	0.007439
	p-value	0.0241**	0.0108**	0.009***	0.0665*	0.0657*	0.4477	0.5517	0.51
1Y - 3M	Pseudo R^2	0.148175	0.193388	0.155843	0.08231	0.084924	0.029921	0.016203	0.021878
	p-value	0.0086***	0.0043***	0.0082***	0.0392**	0.0363	0.1845	0.3256	0.2525
2Y - 3M	Pseudo R^2	0.188247	0.205251	0.146326	0.084469	0.076513	0.041448	0.035351	0.032948
	p-value	0.0032***	0.0024***	0.0077***	0.0342**	0.0435**	0.1235	0.1527	0.168
5Y - 3M	Pseudo R^2	0.219154	0.161591	0.093984	0.054862	0.040283	0.022626	0.01939	0.022684
	p-value	0.0014***	0.0048***	0.0267**	0.084*	0.138	0.2618	0.2986	0.2619
10Y - 3M	Pseudo R^2	0.177166	0.100225	0.047986	0.030314	0.02194	0.011592	0.011182	0.015749
	p-value	0.0036***	0.0234**	0.1086	0.1982	0.2739	0.425	0.4338	0.3542
1Y - 6M	Pseudo R^2	0.212865	0.267187	0.119083	0.097697	0.105638	0.089391	0.04163	0.0602
	p-value	0.0016***	0.001***	0.0145**	0.0246**	0.0198**	0.0272**	0.1231	0.0648*
2Y - 6M	Pseudo R^2	0.264991	0.227596	0.084333	0.078381	0.06148	0.098303	0.087719	0.070587
	p-value	0.0008***	0.0015***	0.036**	0.0425**	0.0708*	0.0261**	0.0344**	0.0563*
5Y - 6M	Pseudo R^2	0.220905	0.102251	0.025416	0.024705	0.010524	0.024513	0.022946	0.025298
	p-value	0.0019***	0.023**	0.2389	0.2459	0.4469	0.2484	0.2642	0.2422
10Y - 6M	Pseudo R^2	0.135707	0.035524	0.002037	0.00578	0.001382	0.007915	0.009526	0.013655
	p-value	0.0112**	0.1679	0.7375	0.5734	0.7831	0.511	0.4715	0.3899
2Y - 1Y	Pseudo R^2	0.098954	0.032344	0.006703	0.010365	0.000825	0.03653	0.082648	0.036411
	p-value	0.0231**	0.1846	0.5438	0.4516	0.8316	0.1623	0.0396**	0.1618
5Y - 1Y	Pseudo R^2	0.117292	0.020357	0.00069	0.001755	0.001422	0.00319	0.010059	0.008547
	p-value	0.0145**	0.2906	0.8452	0.7559	0.7814	0.6752	0.4569	0.4935
10Y - 1Y	Pseudo R^2	0.063987	0.001434	0.005306	0.000491	0.0058	0.000025	0.002599	0.003774
	p-value	0.0672*	0.7784	0.5909	0.8698	0.5765	0.9704	0.7063	0.6502
5Y - 2Y	Pseudo R^2	0.108807	0.011886	0.000086	0.00001	0.005422	0.000688	0.000239	0.00094
	p-value	0.0185**	0.4177	0.9451	0.9817	0.5908	0.847	0.9095	0.8208
10Y - 2Y	Pseudo R^2	0.044083	0.000205	0.01478	0.004059	0.011176	0.003373	0.001047	0.000233
	p-value	0.1268	0.9153	0.3729	0.6383	0.44	0.6697	0.812	0.9105
10Y - 5Y	Pseudo R^2	0.000339	0.029518	0.062288	0.020237	0.01731	0.007942	0.002355	0.000026
	p-value	0.891	0.2105	0.078	0.2995	0.3373	0.5133	0.7212	0.9698

Table 5 reports the AUROC and ROC estimations for this second test. It is evident how numbers improved from those reported in the previous analysis of the same estimations (Table 2). Indeed, both the AUC values and the percentages of correctly predicted values increase. The 2Y – 6M term spread with one lagged quarter shows the highest AUC value (0.84, while in the previous analysis the same term spread showed a lower AUC value, 0.81). Considering the highest percentages of correctly predicted values, we note that in this case, the 5Y – 3M term spread with two lagged quarters is the one which provides the best estimation (95.9%). The 10Y – 3M spread, although presenting fairly good results, does not represent the best estimator.

Table 5: Results

		1 lagged quarter	**2 lagged quarters**
1Y - 6M	AUROC	0.801 (0.0887)	0.794 (0.0939)
	Max correctly predicted	92.9%	94.8%
	Threshold	0.226	0.677
2Y - 6M	AUROC	0.84 (0.104)	0.817 (0.109)
	Max correctly predicted	93.9%	93.8%
	Threshold	0.313	0.293
5Y - 6M	AUROC	0.829 (0.103)	0.722 (0.12)
	Max correctly predicted	92.9%	92.8%
	Threshold	0.307	0.288
5Y - 3M	AUROC	0.768 (0.127)	0.691 (0.15)
	Max correctly predicted	94.9%	95.9%
	Threshold	0.391	0.284
10Y - 3M	AUROC	0.767 (0.13)	0.676 (0.155)
	Max correctly predicted	94.9%	94.8%
	Threshold	0.274	0.272

Table 6 shows the new false positive and false negative rates, calculated accordingly to the new thresholds and new percentages of correctly predicted values. In this case, the two best spreads are identified by the 2Y – 6M and 5Y – 3M. Indeed, the former provides a good false positive rate, and a fairly good false negative rate (the lowest of the sample), while the latter provides a better false positive rate, but a slightly worse false negative ratio. We need to mention that, comparing these results with those obtained with the previous analysis, it is possible to notice that the value of the false negative rates improve (i.e. they decline) when considering the definition of recession instead the definition of economic crisis (hence, this second analysis proved to have better results compared to the first one).

Table 6: Results

		1 lagged quarter	2 lagged quarters
1Y - 6M	False Positive rate	0.0329	0.0000
	False Negative rate	0.6250	0.7500
2Y - 6M	False Positive rate	0.0329	0.0329
	False Negative rate	0.5000	0.5000
5Y - 6M	False Positive rate	0.0329	0.0000
	False Negative rate	0.6250	1.000
5Y - 3M	False Positive rate	0.0000	0.0000
	False Negative rate	0.6250	0.6250
10Y - 3M	False Positive rate	0.0109	0.0000
	False Negative rate	0.6250	0.6250

5.3 - Predicting a recession with other indicators

This paragraph provides the evidence of the models estimated for the four alternative indicators presented in Chapter 4, often considered to be good predictors for future economic activity. As mentioned before, the three indicators provided by the OECD (CLI, CCI, and BCI) have the purpose to study and represent the movements of different underlying trends, such as the expectations of markets participants or fluctuations in macro variables (like unemployment, industrial production, etc). Therefore, from a theoretical point of view, they all are more suited to estimate correct predictions about economic movements. The analysis of these indicators has been conducted considering the same time horizon as before (eight lags), in order to be coherent with the analysis made for the term spreads. Concerning the selection of the dependent variable, the definition of recession was preferred over the definition of economic crisis, due to the better results provided when analyzing the predictive power of the term spreads. The final part of this paragraph shows the results calculated for the probit models based on the Italian stock market index (FTSE MIB).

Table 7 shows a summary of the statistics used to compare previous results: the pseudo r-squared, p-values, false positive rates and false negative rates[50]. In these results only two lags of the quarters are shown, since those are the time horizons that proved to be more significant and which

[50] In order not to fill one table with too many data, the value concerning the ROC, the AUC, the maximum percentage of correctly predicted values and the optimal threshold level were omitted by the table. However, in the Appendix the reader will find the estimates for these values as well. Indeed, false positive and false negative rates were calculated using the best thresholds provided by the ROC and AUC estimations. This same reasoning is valid also for the estimations calculated for the FTSE MIB.

provided better estimations[51]. From an overall point of view, the results of these models are significative. Most of the measures used to compare the models show relevant values. One main evidence is that all three indicators provide way better estimations when the forecast horizon is restricted to one quarter. This finding is particularly evident for the Composite Leading Indicator, whose pseudo r-squares are respectively 0.4727 with one lagged quarter (the highest value among this series) and 0.1297 with two lagged quarters (the lowest). Similar comments may be proposed for the results of the two remaining indexes. Concerning the Business Confidence Index, the pseudo r-squared shows very high values, especially for the '1 lagged quarter' column, and p-value shows good statistical significance as well. However, the difference between the results obtained with one lagged quarter and the results obtained with two lagged quarter is large, the pseudo r-squared widely decreases when the forecast horizon is amplified. The same consideration is valid when considering the false positive and false negative rates: the former provides pretty good estimations, but the latter presents worse values, especially with two lagged quarters. Finally, the Consumer Confidence Index is the only one with an above average pseudo r-squared value when considering 2 lagged quarters. If we had to select the best predictor among these indexes, the choice would fall on the Composite Leading Indicator, both because of the highest goodness of fit highlighted, and because of the quite low false negative rates, a feature that is fundamental in this kind of analysis.

Table 7: Results

		1 lagged quarter	**2 lagged quarters**
BCI	Pseudo r-squared	0.468833	0.222375
	P-value	0.0001***	0.001***
	False positive rate	0.0219	0.0329
	False negative rate	0.5000	0.6250
CCI	Pseudo r-squared	0.420935	0.379262
	P-value	0.0002***	0.0003***
	False positive rate	0.0219	0.0219
	False negative rate	0.5000	0.5000
CLI	Pseudo r-squared	0.472719	0.129654
	P-value	0.0002***	0.0116**
	False positive rate	0.0219	0.0219
	False negative rate	0.5000	0.75

[51]For the full results (considering eight lagged quarters), please refer to the Appendix.

The FTSEMIB is the last indicator that we analyzed to understand if it may be considered as a good predictor for future economic downturns. As mentioned before, data availability for this index is slightly more restricted compared to the data used for the analysis of the other indicators. Indeed, data between January 1997 and December 2019 were taken into consideration. Although the sample is smaller, it contains the index's data relative to the two large recessions occurred in our time horizon, therefore the analysis may be conducted as well. Table 8 provides a summary of the FTSEMIB probit model's estimations, showed in the same format provided for previous indicators. The results are consistent with the results shown for the OECD's predictors. Once again, it is evident that the selection of the time horizon plays a big role concerning the explanatory power of the independent variable, indeed, one lagged quarter proves to be the best forecast also for the stock market index. False negative rates are however quite high for both lags, which it is not a positive signal. P-values show good levels of statistical significance and false positive rates provide good results as well, the models almost never predicted a recession when it did not occur. Comparing these results to the OECD's indicators, however, one conclusion that we may derive is that the FTSE MIB lagged one quarter turned out to be the predictor with less explanatory power. All the measures used to compare the estimations present worse results: the goodness of fit and the statistical significance are lower, false positive rate is basically the same and false negative rate provides a higher value.

Table 8: Results

		1 lagged quarter	**2 lagged quarters**
FTSE MIB	Pseudo r-squared	0.321516	0.180924
	p-value	0.0043***	0.0211
	False positive rate	0.0659	0.0219
	False negative rate	0.6250	0.8750

6. DISCUSSION

6.1 - Interpretation of the results

In the previous chapter the amount of data and results provided may seem vast. Before analyzing the reasons that lie behind those numbers and proposing a critical interpretation, a recap of the analysis and of the main findings may help the reader to clarify the ideas and key takeaways. Our research question was to understand if the term spread, based on the difference between Italian interest rates with different maturities, could be used as an intuitive tool to predict recessions. Since a definition of these events was missing, we defined two different scenarios, economic crises and recessions, and we analyzed the explicative power of the term spreads for both definitions. In order to reach our objective, the two definitions of recession were tested with fifteen combinations of spreads, and according to eight different lags, for a total of 240 probit models' estimations. Then, other four indicators (BCI, CCI, CLI and FTSE MIB) were described and analyzed, with the aim to understand whether their predictive ability was better or worse. A concise comparison of the most important results of our study is provided in Table 9**.** This table is divided in two sections: in the upper part, results calculated with term spreads as explanatory variables are listed, in the lower part the results shown were obtained considering the four alternative indicators as explanatory variables.

The three term spreads that proved to be the best predictors are the 5Y – 3M, 1Y – 6M, and 2Y – 6M spreads, and this result is consistent both when predicting an economic crisis, and when predicting a recession[52]. Comparing the results for the two definitions, it is easy to see that these spreads provided a better forecast when considering the definition of recession, described, just to remind, as a decline in GDP for two consecutive quarters. Each of these three spreads provided the highest accuracy according to three different measures: the 1Y – 6M spread with two lagged quarters showed the highest pseudo r-squared value (0.267), the 2Y – 6M spread showed the lowest false negative rate (0.5), and the 5Y – 3M spread showed the highest percentage of correctly predicted values. However, since the pseudo r-square value of the 2Y – 6M spread is almost the same compared to the 1Y – 6M spread's (0.265 vs 0.267 respectively), and that the percentage of correctly predicted values is almost the same compared to the 5Y – 3M term spread's (93.9% vs 94.9% respectively), we will refer to the 2Y – 6M as our overall best spread. The second section of the table provides instead

[52] Considering the second analysis (predicting a recession), the 5Y – 3M spread provided a pseudo r-squared value slightly lower compared to the 5Y – 6M (0.219 vs 0.22), however, we pointed the former as one of the best spread because the false positive rate and the percentage of correctly predicted values showed better estimations.

the results obtained by the estimations of the probit models considering the four alternative indicators. These models show better goodness of fit compared to the previous results; indeed, pseudo r-squared values present very high numbers. However, comparing the percentages of correctly predicted values, the false positive and false negative rates, the results are very much similar, if not worse, to spreads' estimations. Moreover, the results of these indicators are very good when the time horizon considered is exclusively one quarter in the future; with two lagged quarters, pseudo r-squared values largely decrease for each of them, providing even worse estimations compared to the models based on the term spreads. Among the set of alternative indicators, the Composite Leading Indicator with one lagged quarter showed the best estimations, suggesting that a composite analysis of macro variables fluctuations had a better predictive ability compared to yield curve fluctuations.

Table 9: Results

			Pseudo R2	p-value	Predicted values	False positive rate	False negative rate
Predicting an **economic crisis**	**1Y - 6M**	**1 lag**	0.227536	0.0004***	89.8%	0.0233	0.6154
		2 lags	0.082932	0.0161**	89.7%	0.0000	0.8462
	2Y - 6M	**1 lag**	0.219885	0.0002***	90.8%	0.0233	0.6154
		2 lags	0.090697	0.0107**	88.7%	0.0349	0.6923
	5Y - 6M	**1 lag**	0.168826	0.0008***	87.8%	0.0349	0.6923
		2 lags	0.065062	0.0286**	87.6%	0.0000	1.000
	5Y - 3M	**1 lag**	0.184896	0.0005***	90.8%	0.0000	0.7692
		2 lags	0.096502	0.0088***	90.7%	0.0000	0.6923
	10Y - 3M	**1 lag**	0.144501	0.0017***	90.8%	0.0233	0.6154
		2 lags	0.072544	0.0217**	89.7%	0.0000	0.7692
Predicting a **recession**	**1Y - 6M**	**1 lag**	0.212865	0.0016***	92.9%	0.0329	0.6250
		2 lags	0.267187	0.001***	94.8%	0.0000	0.7500
	2Y - 6M	**1 lag**	0.264991	0.0008***	93.9%	0.0329	0.5000
		2 lags	0.227596	0.0015***	93.8%	0.0329	0.5000
	5Y - 6M	**1 lag**	0.220905	0.0019***	92.9%	0.0329	0.6250
		2 lags	0.102251	0.023**	92.8%	0.0000	1.000
	5Y - 3M	**1 lag**	0.219154	0.0014***	94.9%	0.0000	0.6250
		2 lags	0.161591	0.0048***	95.9%	0.0000	0.6250
	10Y - 3M	**1 lag**	0.177166	0.0036***	94.9%	0.0109	0.6250
		2 lags	0.100225	0.0234**	94.8%	0.0000	0.6250
Predicting a **recession**	**BCI**	**1 lag**	0.468833	0.0001***	94.9%	0.0219	0.5000
		2 lags	0.222375	0.001***	91.8%	0.0329	0.6250
	CCI	**1 lag**	0.420935	0.0002***	94.9%	0.0219	0.5000
		2 lags	0.379262	0.0003***	94.8%	0.0219	0.5000
	CLI	**1 lag**	0.472719	0.0002***	94.9%	0.0219	0.5000
		2 lags	0.129654	0.0116**	91.8%	0.0219	0.75
	FTSEMIB	**1 lag**	0.321516	0.0043***	93%	0.0659	0.6250
		2 lags	0.180924	0.0211**	95.3%	0.0219	0.8750

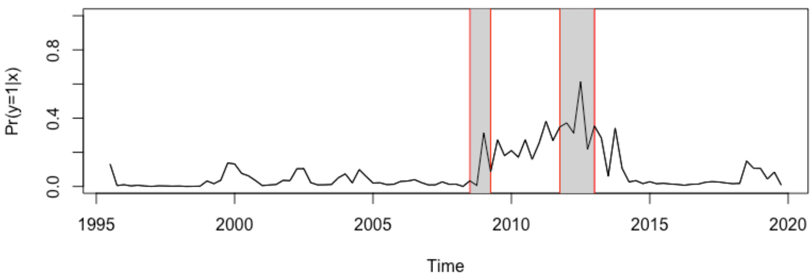
In order to provide a theoretical discussion of our results, first of all, it is essential to understand if the analysis successfully tested the initial hypothesis of our work. For the sake of simplicity, we recall here our hypothesis: 'The term spread, based on the Italian yield curve, may be considered as a reliable predictor for impending national recessions or economic crisis'. Results provided above are helpful to conclude that there at least three different spreads that may be considered as good predictors, hence it is possible to affirm that the results support the paper's hypothesis. Although our objective has been positively tested, there are a few considerations to be made. When referring to an inversion of the yield curve, the doctrine usually refers to the basis point difference between a security with a long maturity and another security with a very short maturity. Indeed, most of past literature is based on the analysis of the 10Y – 3M spread, which in the United States proved to forecast past recessions with great accuracy. In order not to limit our analysis to that specific element, we estimated the models for a large set of combinations. We found out that focusing exclusively on the Italian landscape, the 10Y – 3M spread was outperformed by other combinations of maturities. Among our results, the best estimations are provided by the 2Y – 6M spread, which is made up by the rates of return of two securities with, respectively, a short/medium and a very short maturity. Therefore, it would be quite pretentious to consider this spread as an inversion of the yield curve, since no long-term rates are part of the estimator. However, if we don't focus our attention simply on this spread, our consideration may change. Indeed, the 5Y – 3M spread with one lagged quarter is one of the three best estimators of our sample as well. A bond with a five years maturity, maybe could not be described as a long-term instrument such as the ten years government bond, but at the same time, it can't be described as a short-term debt obligation[53]. We could thus identify five years as a medium/long maturity. With this legit interpretation, it is now correct to state that the inversion of the yield curve based on the 5Y – 3M spread may be defined as a good predictor for recessions in Italy.

Since comparing statistical figures may result tricky, and it may seem that the results provided by the different indicators lack a dynamic interpretation, a graphical representation where the probabilities of recession are plotted may be helpful. In the next page the reader will find the probit models' results for the three best term spreads and for the Composite Leading Indicator. The aim of these representations is to provide a concrete vision of the actual level of probabilities estimated according to the time horizon. These graphs are quite intuitive, and they probably represent the most

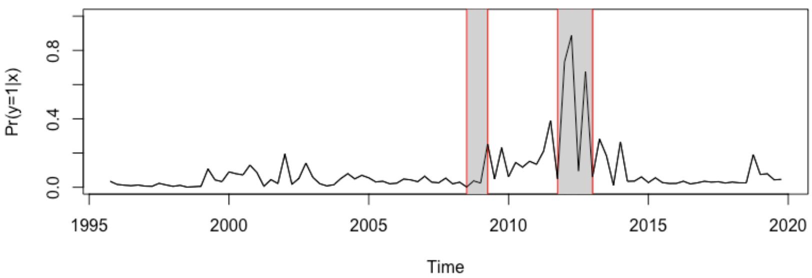
[53] The traditional classification of time horizons is the following: short-term rates are those with maturities lower or equal to twelve months (in our case, the interbank rate, the 3-months BOT and 6-months BOT), medium-term rates have maturities included between two and five years (in our case, the 2-years CTZ and the 5-years BTP) and long-term rates have maturities higher than five years (in our case, the 10-years BTP).

representative final output of the entire study. The black lines draw the trends of the probability scores, while the grey shaded areas define the two recessions occurred in our time horizon.

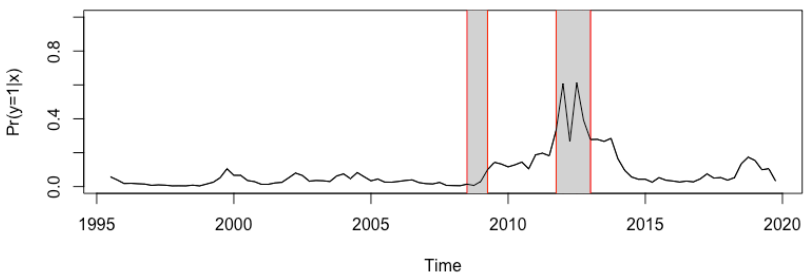
2Y - 6M spread, one lagged quarter



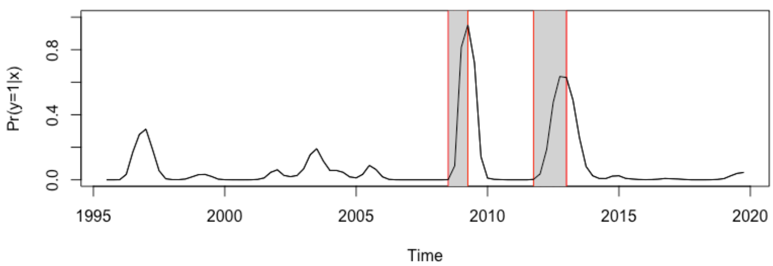
1Y - 6M spread, two lagged quarters



5Y - 3M spread, one lagged quarter



Composite leading indicator, one lagged quarter



Some conclusions may be derived from the observation of these graphs. Each term spread correctly predicted with a good accuracy the Sovereign Debt crisis occurred between the last quarter of 2011 and the first quarter of 2013. In each graph we can observe that the probability levels heavily increase approaching this specific recession period. This fact is more manifest for the 1Y – 6M spread, where the probability reaches a level of almost 80%, but it is valid for the other two spreads as well, where the estimated probability overcomes the 50% level. One aspect that catches the attention is the difference between the probability distributions of the spreads estimations. The 1Y – 6M and 2Y – 6M plots present more jagged lines, with more up and downs, while the 5Y – 3M shows a smoother distribution. This difference may be explained by the composition of the explanatory variables. The rate of returns of the instruments with shorter maturities are way more likely to be similar, and the difference among these returns may become negative more often. On the other side, it is almost a rare event to assist to an inversion of the rate of returns of two securities which have very distant maturities (i.e. 3-months interbank rate and 5-years BTP). This difference represents the basic logic that allow us to state that the 5Y – 3M spread may be considered more suited to predict recessions. Indeed, a negative spread considering these debt instruments is a rare event, and thus it may represent a signal for an unusual situation of the overall economy. Switching the focus on the comparison between different explanatory variables, the most notable difference between the results obtained with the spreads and those obtained with the Composite Leading Indicator, is the ability that the CLI proved to have in predicting the first recession, in 2008-2009 years. Each of the term spreads did not succeeded in predicting this economic crash with good accuracy, even if probability's levels slightly increase for each of them. However, we also need to mention that observing the CLI's graph with a more rigorous eye, we may see that both recessions were predicted late. Indeed, the two peaks of the probability levels are reached in correspondence of the end of both recessions. Nevertheless, this consideration should not surprise. If we think to the fundamental elements of the Composite Leading Indicator, such as industrial production, unemployment rate or import and export balance, it is reasonable to expect a good but late prediction; indeed, all these elements could be defined more as consequences of a recession than as predictors. Referring again to the CLI's graph, this indicator also provides a sort of false positive case between years 1995 and 2000, when it estimated a probability of recession with almost 40% of probability, but when actually no recessions occurred. On the other side, each of the spread's estimations provide a slight increase of the recession probabilities in 2018, while the CLI's probability line stay flat and close to zero. This sort of alarm corresponded to the decline of the GDP that occurred in the third quarter of 2018, hence the spreads showed a better capacity to predict this information.

From an overall point of view, our analysis provided results that could be described significant and consistent with our initial expectations. Nevertheless, before being able to define the Italian yield curve a great predictor, it is important to understand the limits of our work and the areas of improvement. Although the statistical side of the work provided good results, there are a few topics and considerations that need to be mentioned. For example, the lack of accuracy shown in the prediction of the first recession represents one issue concerning the robustness of our results. However, it is difficult to state whether this misinterpretation was due to the poor performance of the model in forecasting this specific recession, or whether it was due to issues related to data availability. We mention data availability because our probit models were estimated on a quite restricted amount of data. Indeed, referring to the statistical estimation, it has been difficult for our models to correctly predict the values of the dependent variable, considering that this variable was observed only eight times among the whole sample. This scarcity of data may have complicated the estimation of the recession probabilities, and therefore it probably represents the main limit of our study. Moreover, econometric models whose purpose is to forecast GDP fluctuations, usually rely on a vast set of variables. Using exclusively the term spread to predict declines of national accounts may therefore lead to correct, but improvable results.

Finally, it is important to provide a comment regarding the predictive power of the curve in the current environment. As explained in the first chapter, the signals provided by the yield curve are very sensitive to financial markets conditions. Nowadays, financial markets present peculiar characteristics, and the activity that they support is massive. Therefore, it became more and more complicated to understand if changes in financial markets are due to technical factors or economic fundamentals. The slope of the yield curve may thus be affected for short times by temporary trends caused either by an increase in demand for a specific asset class, or by the change in preferences of a specific set of investors, or still by speculation. These circumstances, whether affecting the yield curve economic fundamentals, may cause the signals provided by the slope of the curve to be misleading. The current situation of interest rates may be another topic of discussion as well. Financial markets are experiencing a period of extremely low interest rates, and this trend is not expected to end in the short term (this current situation starts to be defined as the "new normal"). Therefore, applying the study of the inversion of the yield curve, which proved to be very powerful in past macroeconomic conditions, may not be totally effective. Moreover, a recent study conducted by the European Central Bank, and published after the last inversion of the yield curve occurred in the United States in mid 2019, analyzed the recession probabilities estimated with adjusted probit models. In this analysis, the term spread was adjusted for the effect of different variables, such as the impact of the quantitative easing program undertaken by the Federal Reserve, or the spillover effect of US asset purchases by the European Central Bank. The result of this study showed that the probability estimations based on the yield curve term spread became less significant when several adjustments were made. However, this result should not surprise too much. Indeed, the quantitative easing programs and the measures adopted by central banks in the last ten years have distorted interest rates. Gillian Tett, former managing director of the Financial Times, recently wrote in one of her articles that the nature of business cycles has shifted in the last few years. This shift refers to the causes that lead to recessions, meaning that nowadays we are assisting to the shift from inflation-induced to financial cycle-induced recessions (the 2008 crisis was the first example of a financial cycle-induced recession). This analysis is a more advanced kind of study, but it could represent an interesting suggestion for future research related to our paper. Nevertheless, we need not to forget that Italian results will always be difficult to compare with other nations, in particular with the US results, due to the substantial differences of the bank systems. Moreover, although the analysis based on the Italian scenario provided fairly good results, we still believe that the inversions of the US yield curve may be considered as a good benchmark for the health of the global economy. The main reason is that with ultra-connected global economies such as nowadays' environment, and with the United States still representing one of the strongest players around the markets, an inversion of the Treasuries yield curve may reflect a pessimistic picture for many more countries.

7. CONCLUSION

Our analysis was drove by the interesting relationship existing between the movements of the yield curve and the start of a recession. Assuming that this relationship was something certain and unavoidable, an inversion of the curve would be considered with more attention. Indeed, in the United States, where this relationship proved to be accurate several times, this event receives much more credit. The idea to test this hypothesis in Italy derives from the low attention that these events usually receive in our country. The statistical models that we used to translate the observations of real phenomena into specific values had a concrete objective: to test if these events could represent a useful signal to partially avoid the disastrous consequences that a recession necessarily brings. Our results proved that, although an inversion of the curve may not be considered as a universal principle to predict the start of a recession, it still represents an alarm for economic instability. Besides considering the term spread between the 2-years and 6-months securities, which showed the best estimations but which cannot be described as a proper inversion of the curve, the results show that an inversion between the 5-years and 3-months government bonds correctly predicted most of the observations of the sample. These results were then compared with those obtained by the estimations of other indicators, such as the CLI, the CCI, the BCI and the FTSEMIB. These variables, apart from providing better statistical goodness of fit compared to the spread's estimations, did not particularly improved the results previously obtained. Only the Composite Leading Indicator succeeded in predicting recessions with higher accuracy compared to the term spread, even if it provided slightly late predictions. Considering the weaknesses of our study, the main drawback was represented by the limited availability regarding the length of historical series and the quantity of recession periods of our sample, which may both have affected the robustness of our results. However, we believe that this analysis proved to be coherent with our initial expectations. Moreover, it may be considered as a starting point for future literature. Future research may focus for example on the impact that current macroeconomic conditions are having on the predictive ability of the Italian term spread. Finally, although our results were significant, we need to mention that in our opinion the United States' yield curve could still represent a good benchmark, because of the predominant position that the US still has among global economies, and because of the large interconnections that nowadays exist among countries on a global scale.

8. REFERENCES

- Menzie D. Chinn, Kavan J. Kucko. 2010. “The predictive power of the Yield Curve across countries and time”. National Bureau of Economic Research - Working Paper 16398;

- Arturo Estrella, Mary R. Trubin, 2006. “The Yield curve as a Leading Indicator, Some Practical Issues”. Federal Reserve Bank of New York, volume 12, number 5, July August 2006;

- Yolanda S. Stander, 2005. “Yield curve modeling”. Palgrave Macmillan;

- Moorad Choudhry, 2019. “Analysing and Interpreting the Yield Curve”. Wiley Finance Series, second edition;

- Philip R. Lane, 2019. “The yield curve and monetary policy”. European Central Bank, Public lecture for the Center for Finance and the Department of Economics at University of London, speech of Philip Lane, member of the executive board of ECB;

- Maria D. G. Rivas and Gabriel Perez-Quiros, 2014. “The failure to predict the Great Recession. A view through the role of credit”. Journal of the European Economic Association, vol. 13(3), pages 534-559;

- Philip E. Davis and Gabriel Fagan, 1997. “Are Financial Spreads Useful Indicators of Future Inflation and Output Growth in EU Countries?”. Journal of Applied Econometric, 12, pages 701-714;

- Fabio Moneta, 2003. “Does the yield spread predict recessions in the Euro area?”. European Central Bank, working paper series n° 294;

- Michael D. Bauer and Thomas M. Mertens, 2018. “Economic Forecasts with the Yield curve”. Federal Reserve Bank of San Francisco, Economic Letter 2018-07;

- David S. Miller, 2019. “There is No Single Best Predictor of Recessions”. Board of Governors of the Federal Reserve System, FEDS Notes, May 21, 2019;

- Arturo Estrella, Frederic Mishkin, 1996. “The yield curve as a predictor of US recessions”. Federal Reserve Bank of New York, Current Issues in Economics and Finance, Volume 2, Number 7;

- Peter Johansson and Andrew Meldrum, 2018. “Predicting Recession Probabilities Using the Slope of the Yield Curve”. Board of Governors of the Federal Reserve System, FEDS Notes, March 01, 2018;

- Glenn D. Rudebusch and John C. Williams, 2008. “Forecasting recessions, the puzzle of the enduring power of the yield curve”. Federal Reserve Bank of San Francisco. Working Paper 2007-16;

- Henri Bernard and Stefan Gerlach, 1996. “Does the Term Structure Predict Recession? The International Evidence”. Bank of International Settlements, Monetary and Economic Department, Basle. Working Paper n. 37;

- Fabio Fornari and Wolfgang Lemke, 2010. “Predicting recession probabilities with financial variables over multiple horizons”. European Central Bank, Working Paper Series, No 1225/October 2010;

- Emre O. Ergungor, 2016. “Recession Probabilities”. Federal Reserve of Cleveland, Economic commentary Number 2016-09;

- Jonas Dover and Christina Ziegler, 2008. “Predicting growth rates and recessions: assessing US leading indicators under real-time conditions”. Kiel Institute for the World Economy, Kiel Working Paper, No 1397;

- Arturo Estrella, Anthony P. Rodrigues and Sebastian Schich, 2003. “How Stable is the Predictive Power of the Yield Curve? Evidence from Germany and the United State”. Review of Economic and Statistic, N 85-3;

- Thomas B. King, Andrew T. Levin and Roberto Perli, 2007. “Financial Market Perceptions of Recession Risk”. Federal Reserve Board Washington D.C., Finance and Economics Discussion Series, 2007-57;

- Michael D. Bauer and Glenn D. Rudebusch, 2016. “Why Are Long-Term Interest Rates So Low?”. Federal Reserve Bank of San Francisco, Economic Letter 2016-36

- Michael Dotsey, 1998. “The predictive content of the interest rate term spread for future economic growth”. Federal Reserve Bank of Richemond, Research Publication, vol 84, n° 3;

- Sean D. Campbell 2007, “Macroeconomic Volatility, Predictability and Uncertainty in the Great Moderation: Evidence From the Survey of Professional Forecasters”. Journal of Business and Economics Statistics, 25, 191.200

- Jordan J. Louviere, David Hensher, Joffre Swait, 2000. “Stated choice methods”. Cambridge University Press;

- Reuben A. Kessel, 1965. “Explanation of the Term Structure of Interest Rates”. National Bureau of Economic Research, Chapter in NBER book “The Cyclical Behaviour of the Term Structure of Interest Rates, published by NBER in 1965;

- Marco Del Negro, 2001. “Turn, Turn, Turn: Predicting turning points in economic activity”. Federal Reserve Bank of Atlanta, Economic Review 86(Q2): 1-12;

- Arturo Estrella and Frederic S. Mishkin, 1997. “The Predictive Power of the Term Structure of Interest Rates In Europe And The United States: Implications for the European Central Bank”. European Economic Review 41(7);

- “The Joint Harmonised EU Programme of Business and Consumer Surveys”. European Commision Directorate-General for Economic and Financial Affairs, updated 4th July 2007;

- User guide published by the European Commission Directorate-General For Economic And Financial Affairs, updated on 4th July 2007. "The Joint Harmonised EU Programme of Business and Consumer Surveys;

- John C. Williams, 2017. "Interest Rates and the 'New Normal'". Federal Reserve Bank of San Francisco, Economic Letter 2017-29 (October 10);

- Article by Gillian Tett: "The importance of an inverted yield curve is overrated". Financial Times website, 14th November 2019;

- Article by Richard Leong: "U.S. Yield curve: Invert, Steepen, Repeat". Reuters website, 20th August 2019;

- Article by Jim Glassman: "Reassessing the Inverted Yield Curve". JP Morgan Chase, Commercial Banking, 21st August 2019

- Article by Vittorio Carlini and Vito Lops: "Altro che spread! Per monitorare la crisi guardate la curva dei rendimenti". Il Sole 24 Ore, 30th May 2018;

- Data bank: "Istituto Nazionale di statistica". Istat Database;

- Data bank: "Banca d'Italia". Bank of Italy Database;

- Data bank: "Monthly report on Italian economic conditions". Istat Database;

- Data bank: "European Central Bank (ECB)";

- Data bank: "Organization for Economic Co-operation and Developmnent". OECD Database;

- Website: "Ministero dell'Economie e delle Finanze (MEF)";

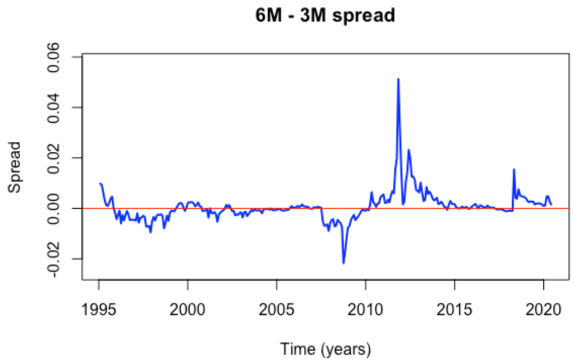
- Website: "Borsa Italiana";

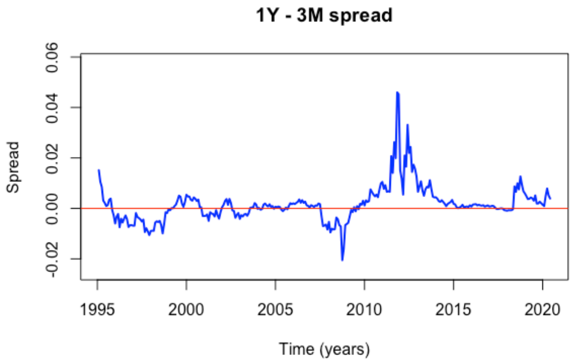
- Website: "National Bureau of Economic Research (NBER)";

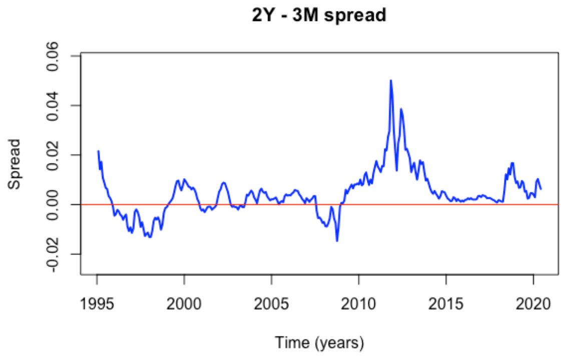
- Website: "Commissione Nazionale per la Società e la Borsa (CONSOB)"

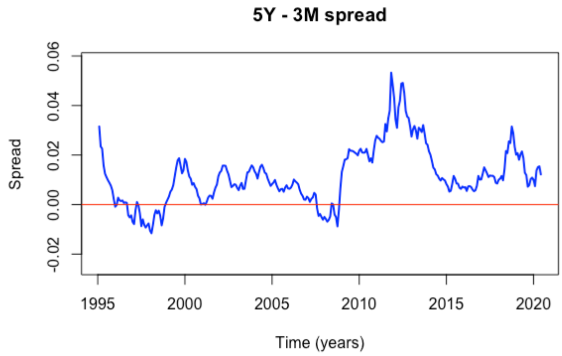
9. APPENDIX

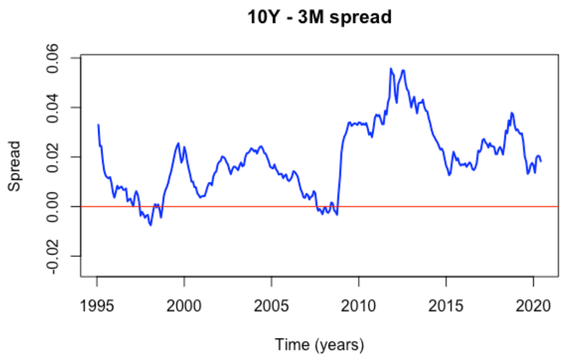
9.1 - The following graphs show the spread values calculated for each of the combinations considered in our analysis. The red line represents the 0 level, therefore, when the spread line falls below it, an inversion of the curve occurs.

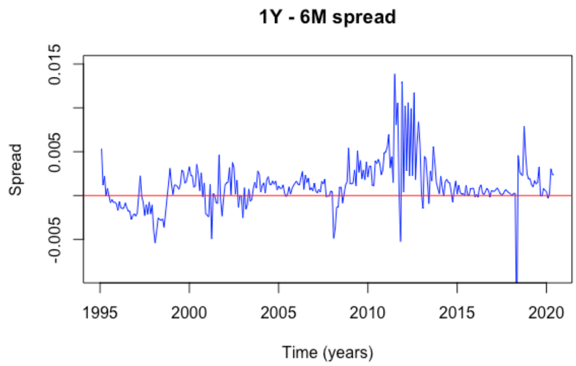


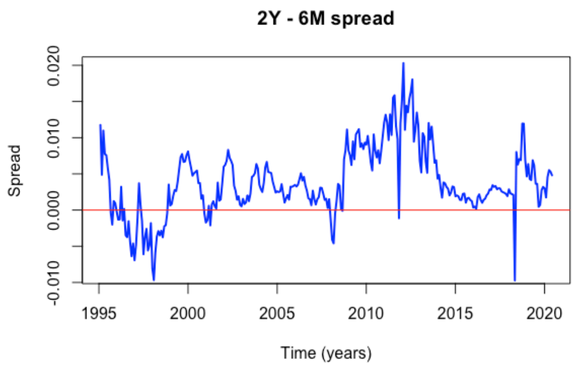


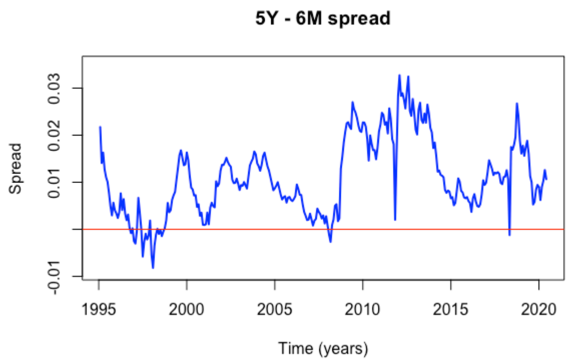


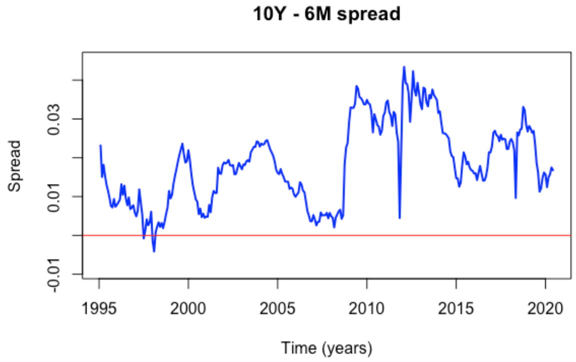


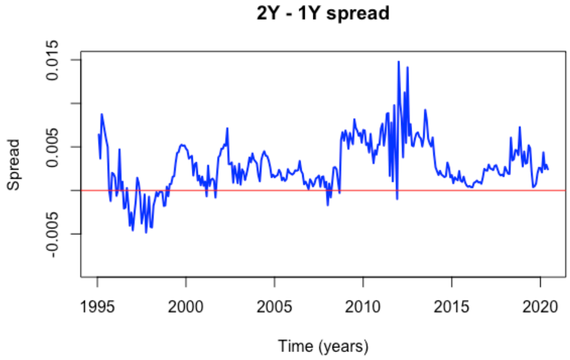


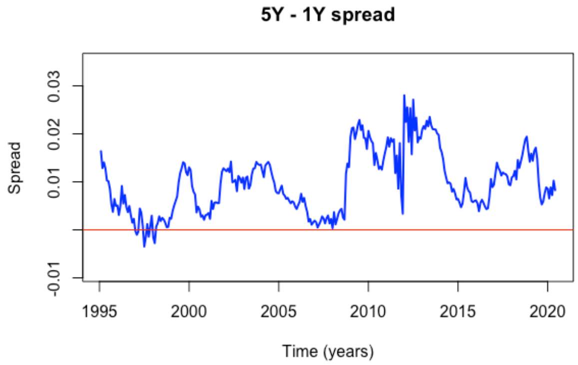


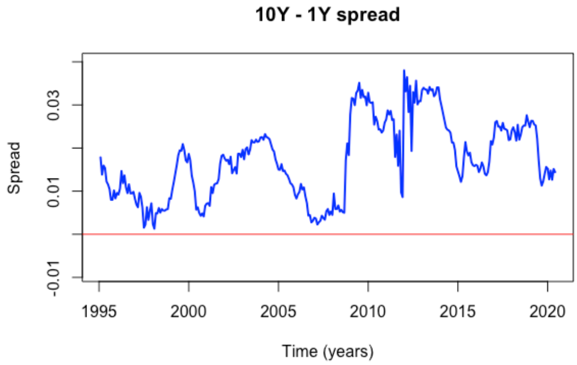


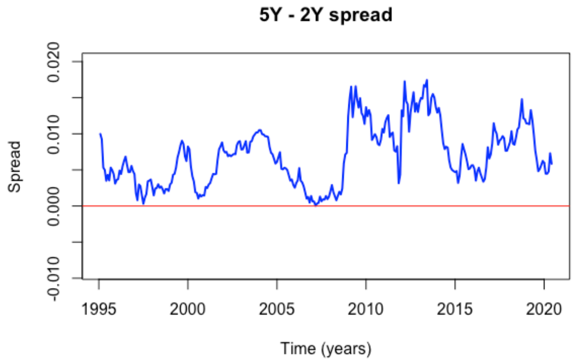


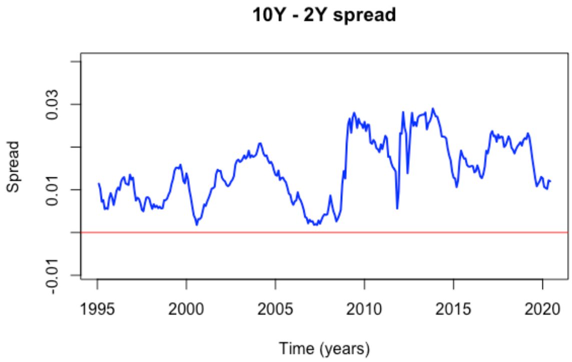

10Y - 6M spread
Spread
-0.01
0.01
0.03
1995
2000
2005
2010
2015
2020
Time (years)

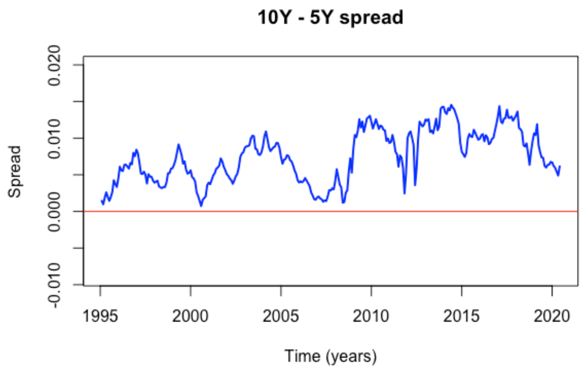

2Y - 1Y spread
Spread
-0.005
0.005
0.015
1995
2000
2005
2010
2015
2020
Time (years)


5Y - 1Y spread
Spread
-0.01
0.01
0.02
0.03
1995
2000
2005
2010
2015
2020
Time (years)


10Y - 1Y spread
Spread
-0.01
0.01
0.03
1995
2000
2005
2010
2015
2020
Time (years)


5Y - 2Y spread
Spread
-0.010
0.000
0.010
0.020
1995
2000
2005
2010
2015
2020
Time (years)


10Y - 2Y spread
Spread
-0.01
0.01
0.03
1995
2000
2005
2010
2015
2020
Time (years)


10Y - 5Y spread
Spread
-0.010
0.000
0.010
0.020
1995
2000
2005
2010
2015
2020
Time (years)

9.2 – The next representation shows the cross-tabulation of predictions and actual observations. This matrix helped us to calculate the false positive and false negative rates for each observation and estimator.

		Prediction		
		0	1	Tot.
Actual	0	n_{00}	n_{01}	n_0
	1	n_{10}	n_{11}	n_1
	Tot.	N_0	N_1	n

n_{00} : True negative n_{10} : False negative n_{01} : False positive n_{11} : True positive

Several goodness of fit measures can be obtained on the basis of this table. Overall the proportion of incorrect predictions is:

$$proportion\ of\ incorrect\ predictions = \frac{n_{01} + n_{10}}{n}$$

For the sake of clarity, an example related to the prediction of recession is provided below.

		Predictions		
		0	1	Tot.
Actual	0	*95*	*10*	*105*
	1	*8*	*15*	*23*
	Tot.	*103*	*25*	*128*

For this specific model, we calculate the false positive rate, the false negative rate and the proportion of incorrect predictions:

False positive rate: $n_{01}/(n_{00} + n_{01})$ hence, 10/(95+10) = **0.0952**

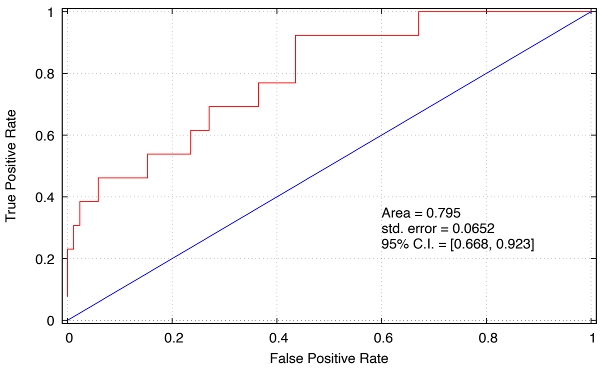
False negative rate: $n_{10}/(n_{10} + n_{11})$ hence, 8/(8+15) = **0.3478**

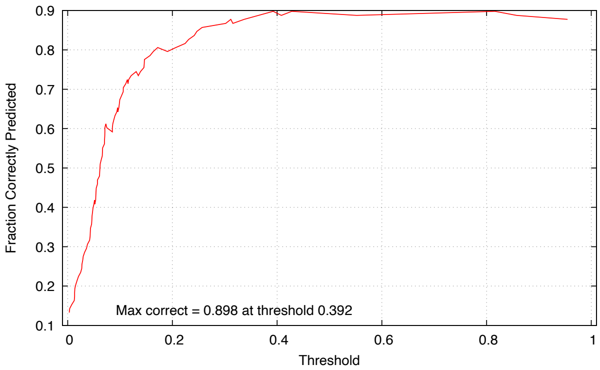
Proportion of incorrect predictions: $\frac{n_{01}+n_{10}}{n}$ hence, (10+8)/128 = **0.1406**

9.3 – Next graphs are the representation of the ROC and AUC estimations. For each combination of spread, two graphs are provided. The former plots the ROC curve, evidencing that AUC (Area under the curve, also called AUROC), its standard error, and a 95% confidence interval. This graph is cut by a red 45-degree line, which represent a classifier with absolutely zero explanatory power (the area under this curve is 0.5, hence the closest is the ROC curve to the top-right corner, the higher the accuracy of the model). The latter plots the fractions of correctly predicted values according to the threshold considered (here we have the evidence of the thresholds that maximize the percentage of correctly predicted values).

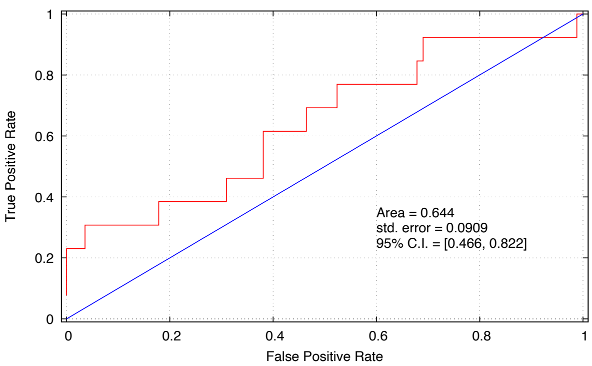
First analysis. Predicting an economic crisis (with spreads)

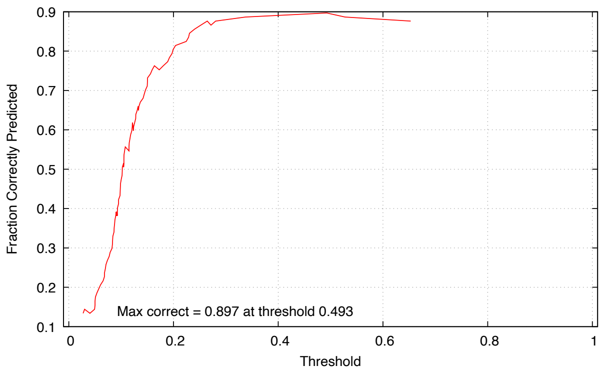
<u>1Y – 6M spread, one lagged quarter</u>



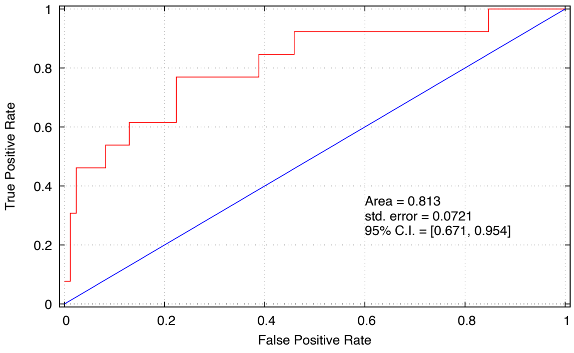


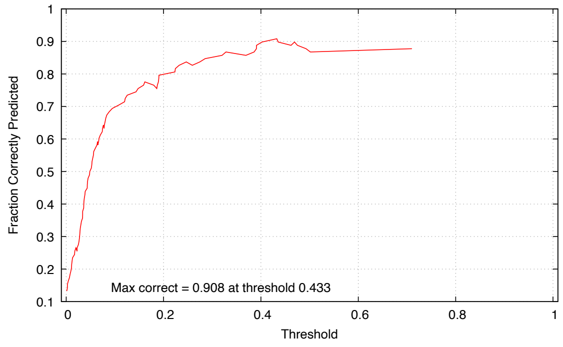
<u>1Y – 6M spread, two lagged quarters</u>



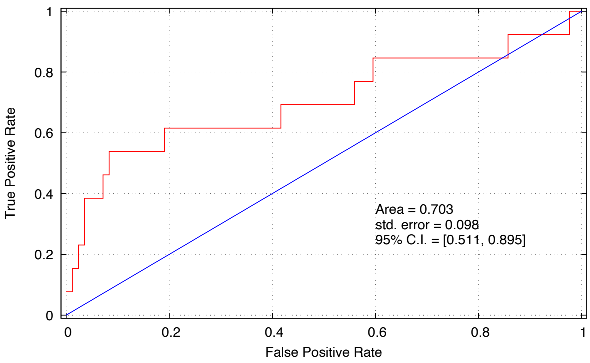


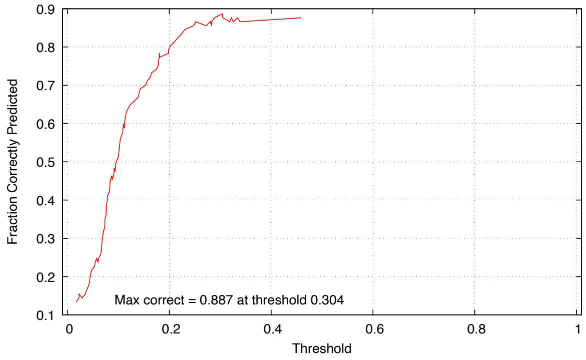
2Y – 6M spread, one lagged quarter



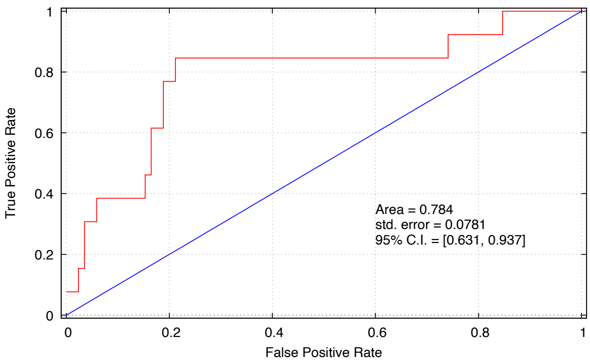


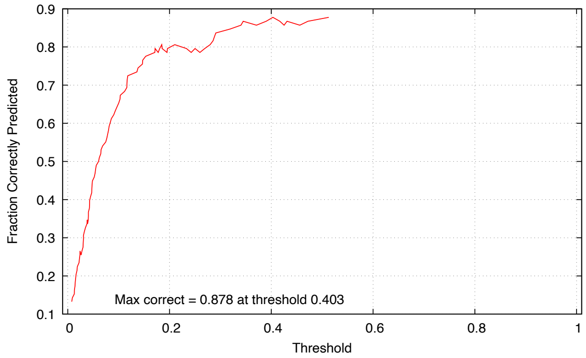
2Y – 6M spread, two lagged quarters



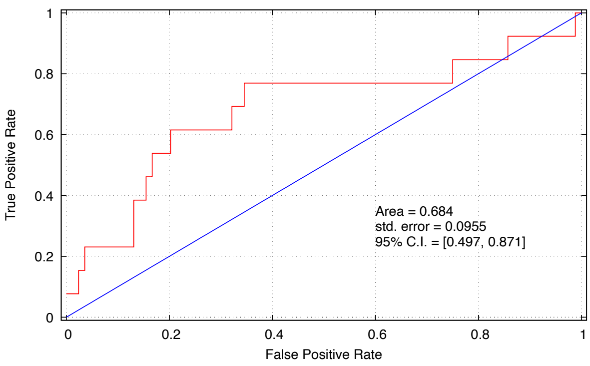


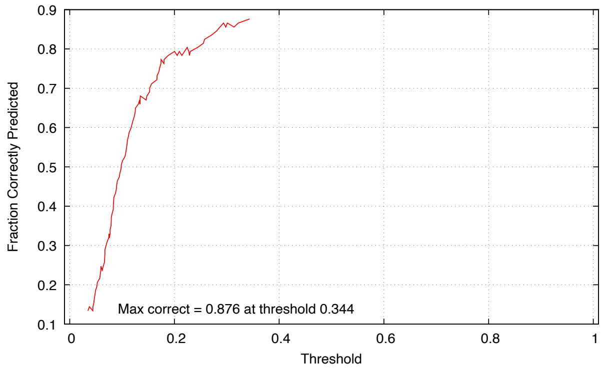
5Y – 6M spread, one lagged quarter



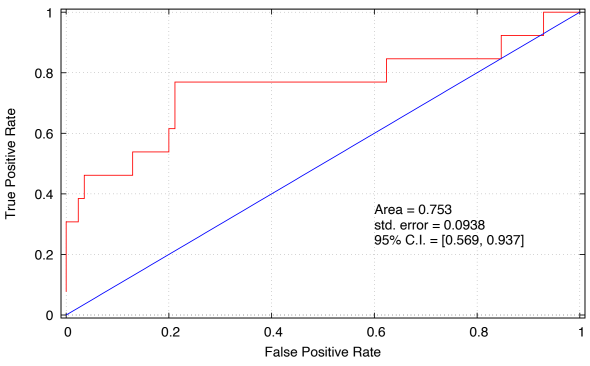


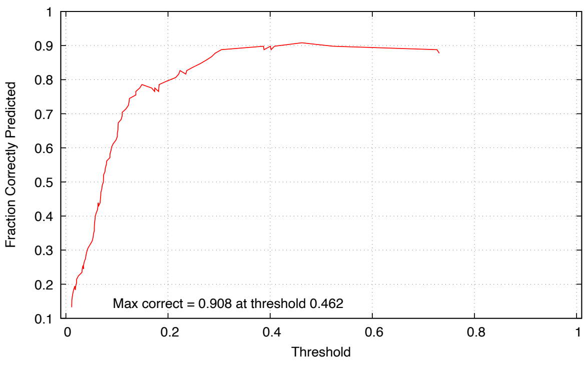
5Y – 6M spread, two lagged quarters



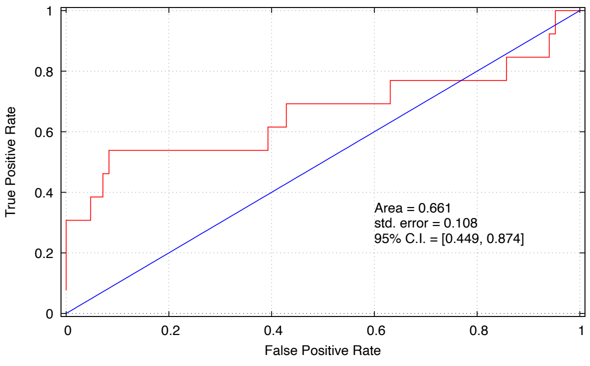


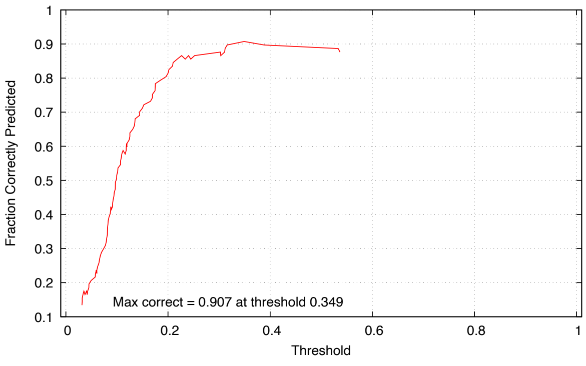
5Y – 3M spread, one lagged quarter



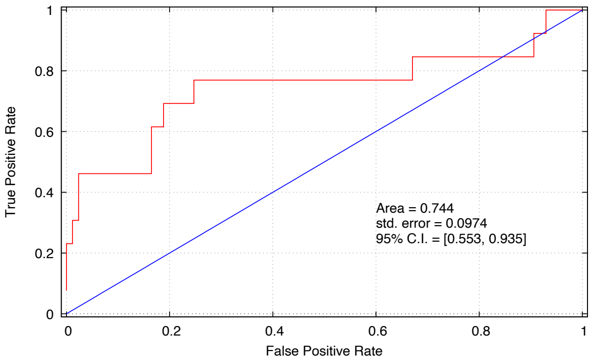


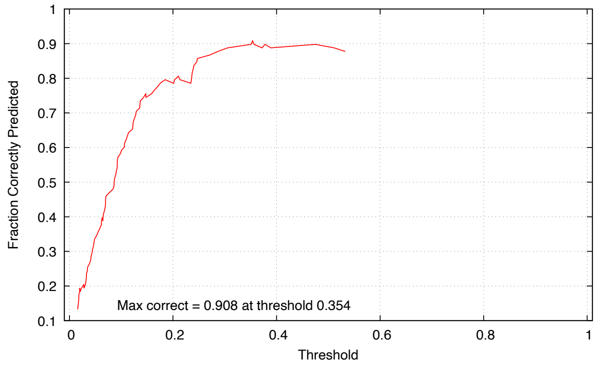
5Y – 3M spread, two lagged quarters



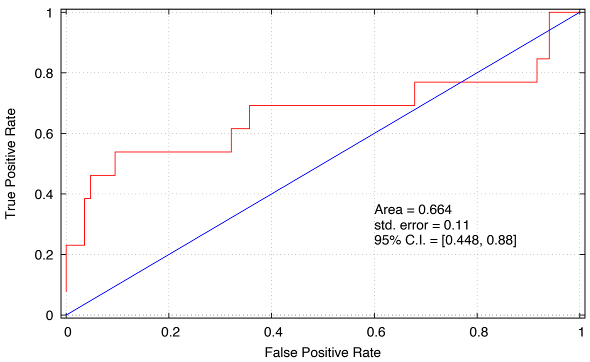


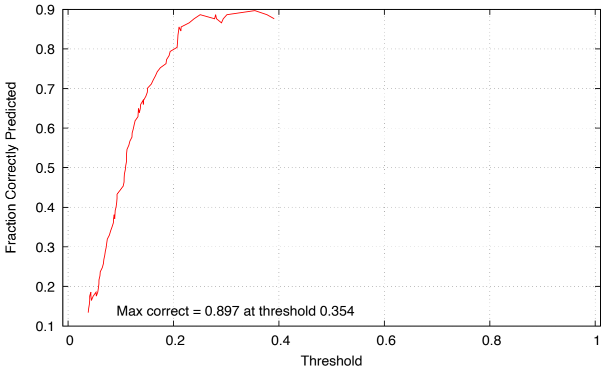
10Y – 3M spread, one lagged quarter





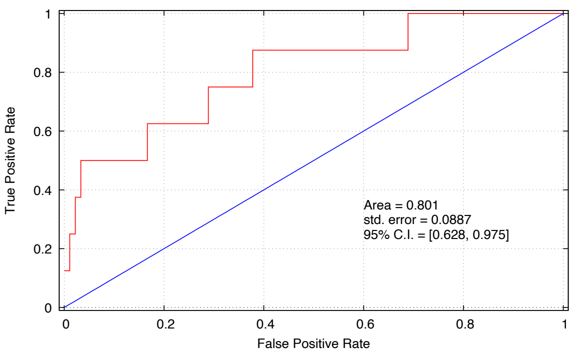
10Y – 3M spread, two lagged quarters

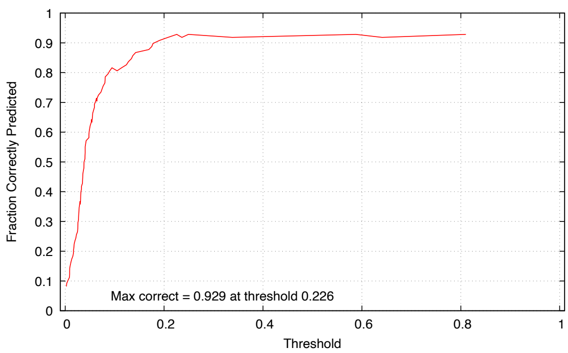




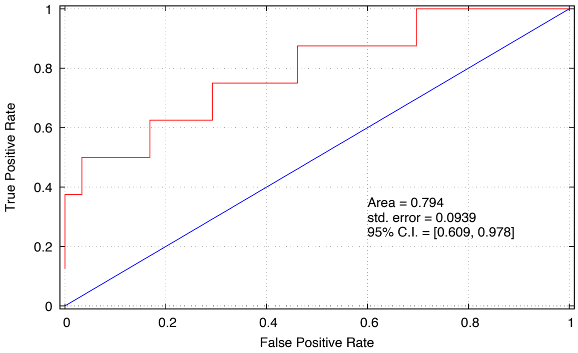
Second analysis. Predicting a recession (with spreads)

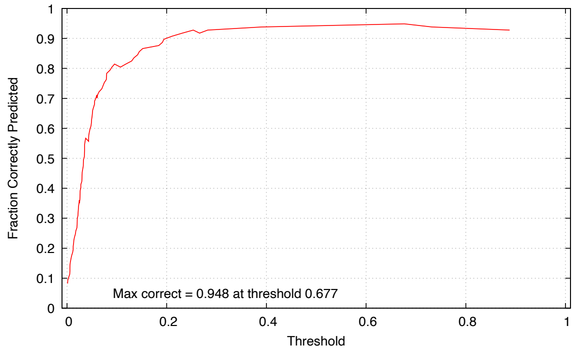
1Y – 6M spread, one lagged quarter



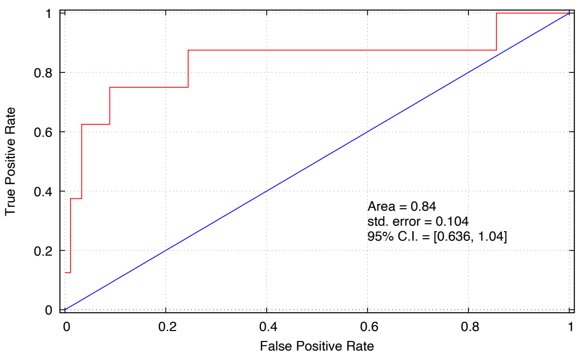


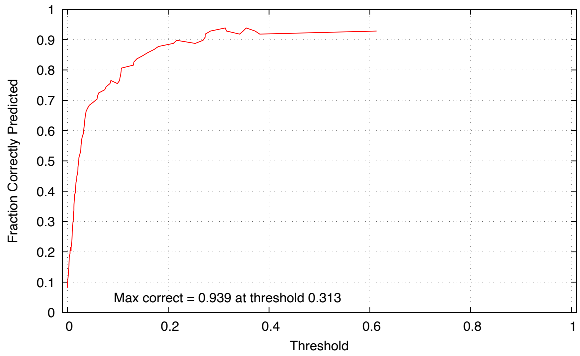
1Y – 6M spread, two lagged quarters



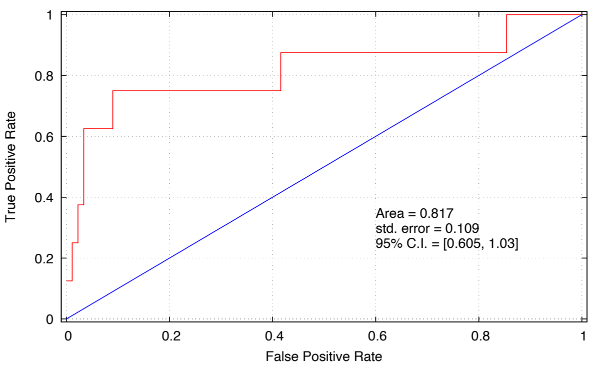


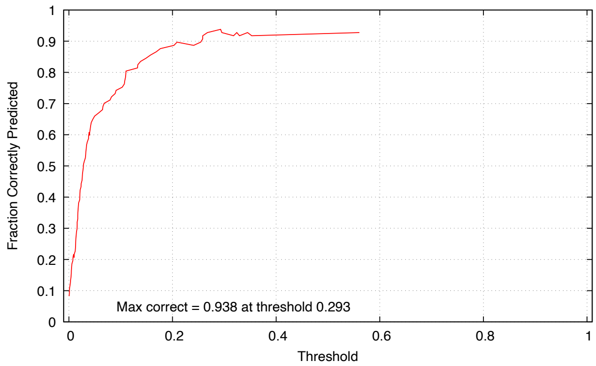
2Y – 6M spread, one lagged quarter



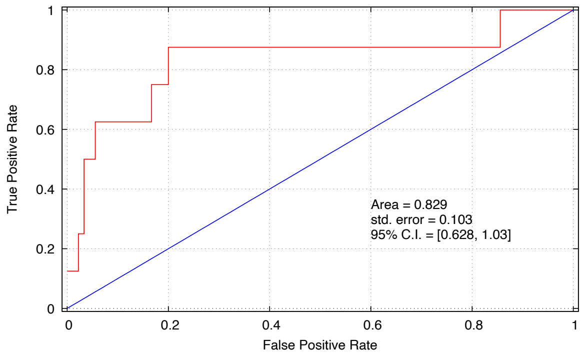


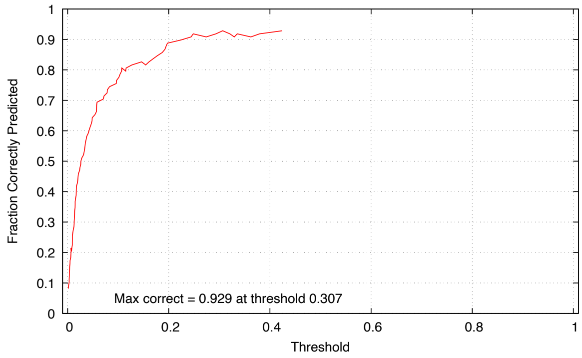
2Y – 6M spread, two lagged quarters



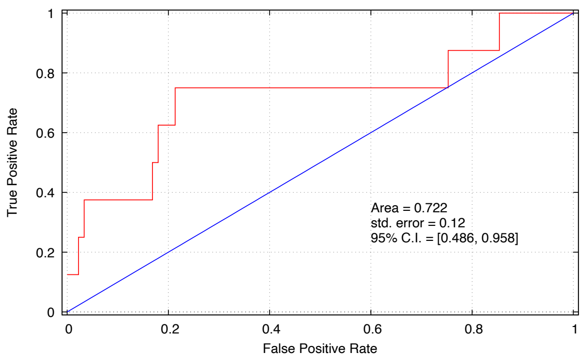


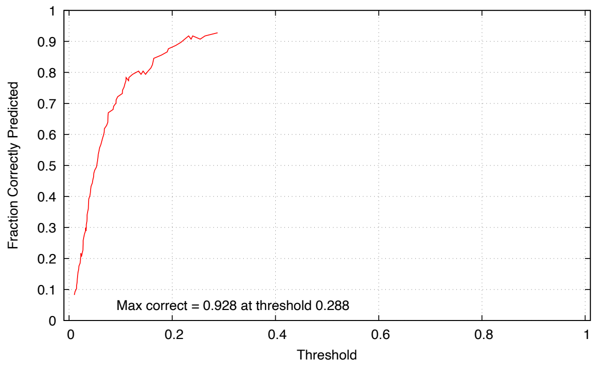
5Y – 6M spread, one lagged quarter



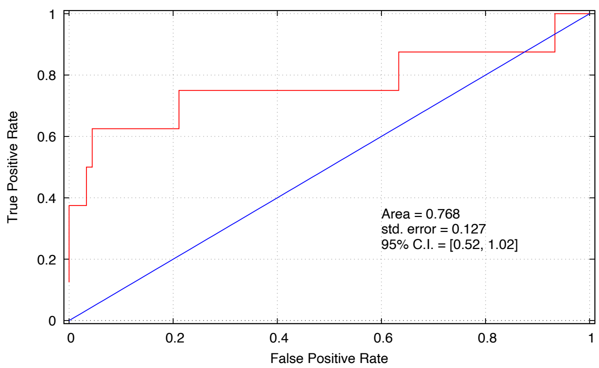


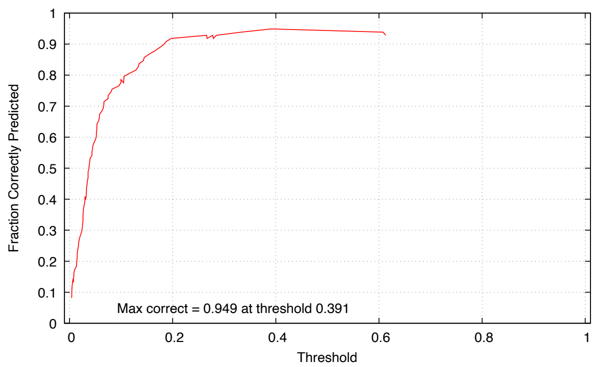
5Y – 6M spread, two lagged quarters



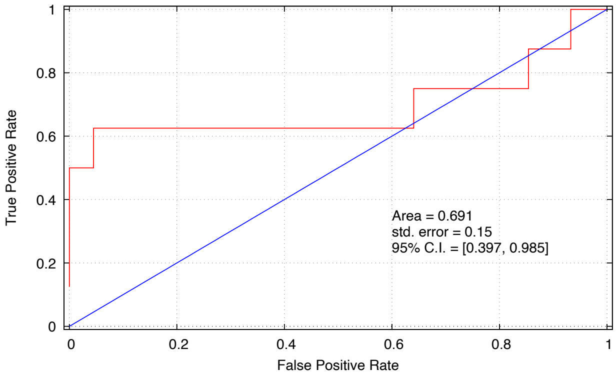


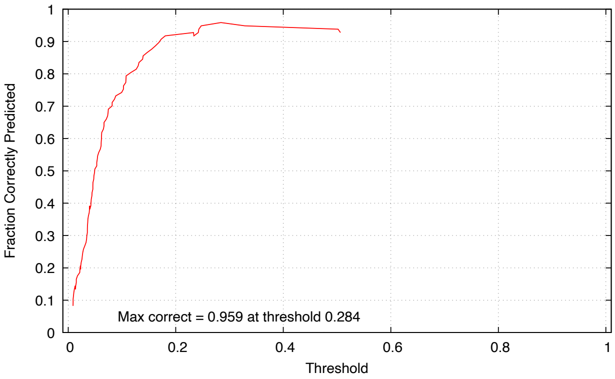
5Y – 3M spread, one lagged quarter



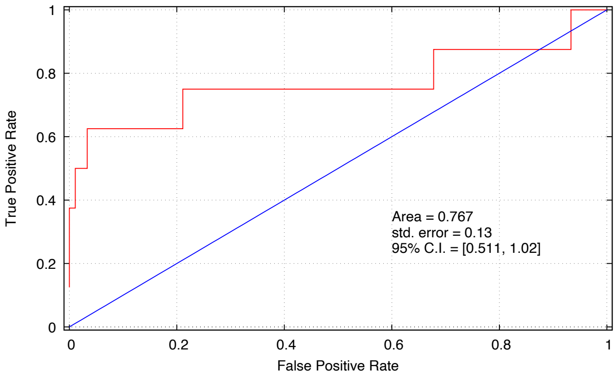


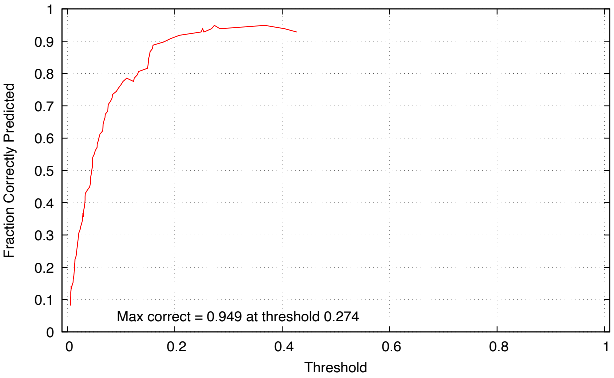
5Y – 3M spread, two lagged quarters



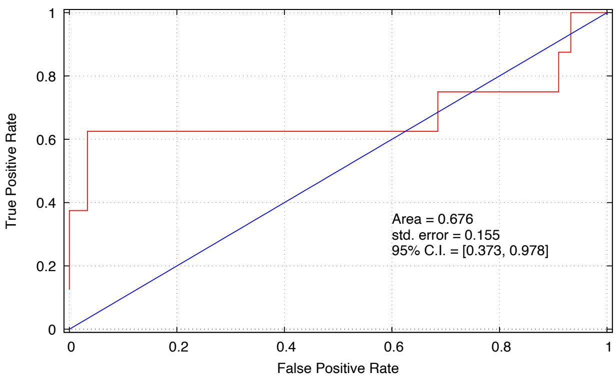


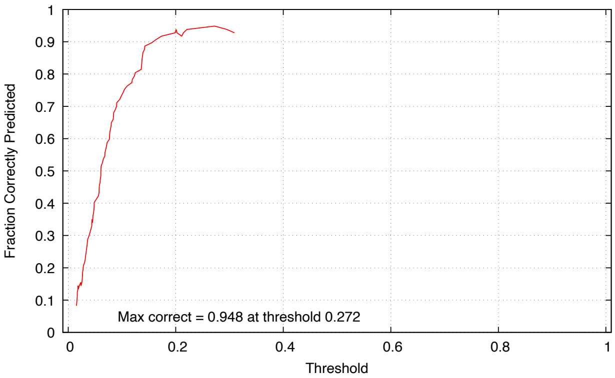
10Y – 3M spread, one lagged quarter





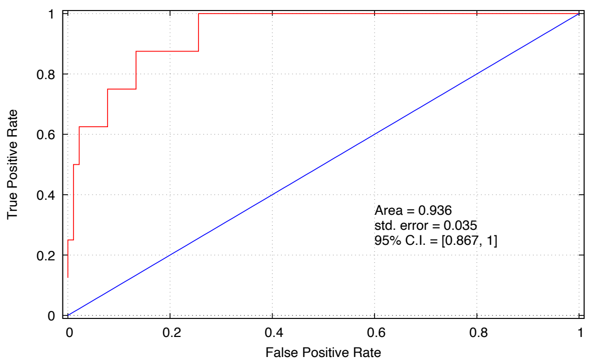
10Y – 3M spread, two lagged quarters

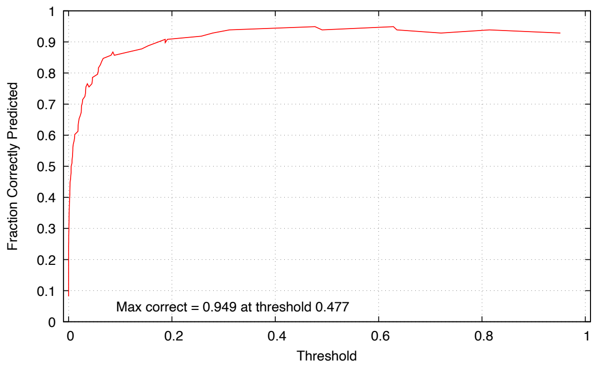




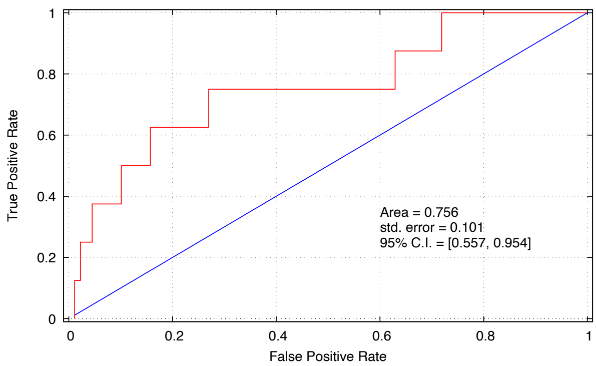
Third analysis. Predicting a recession (with alternative indicators)

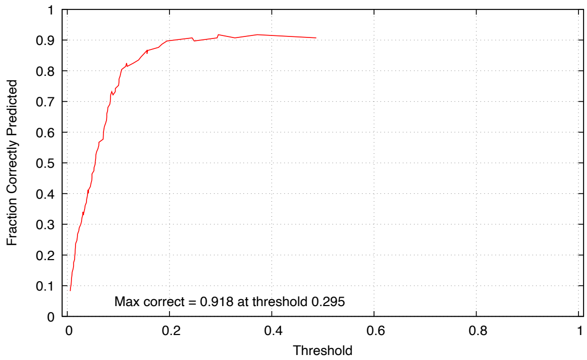
Composite Leading Indicator, one lagged quarter



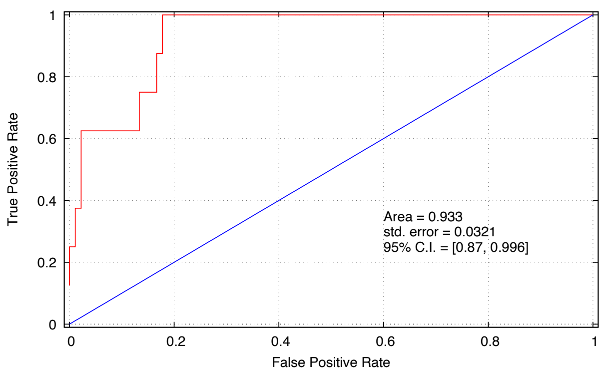


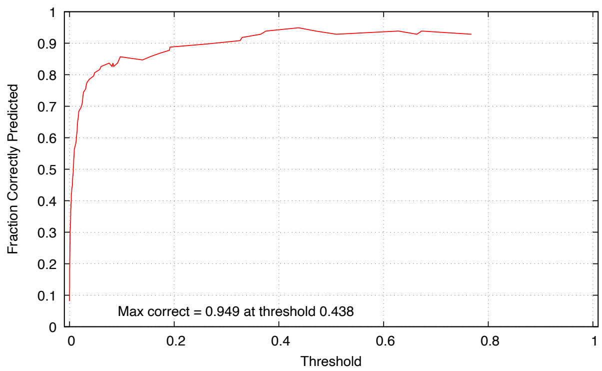
Composite Leading Indicator, two lagged quarters



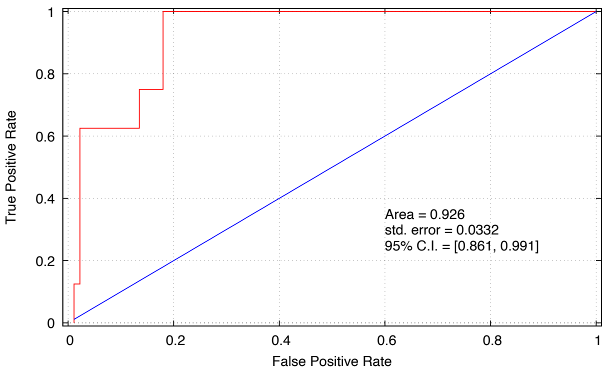


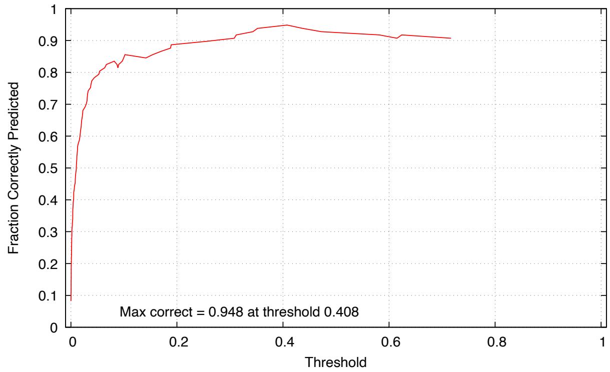
Consumer Confidence Index, one lagged quarter



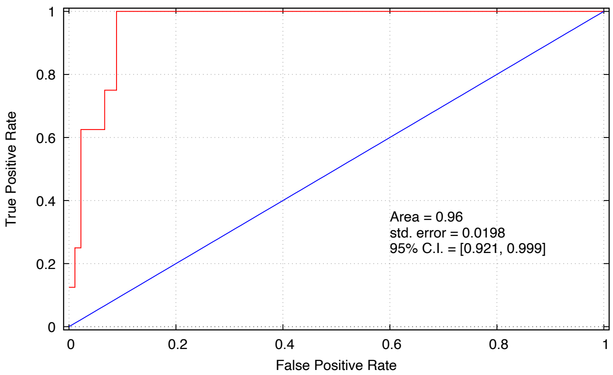


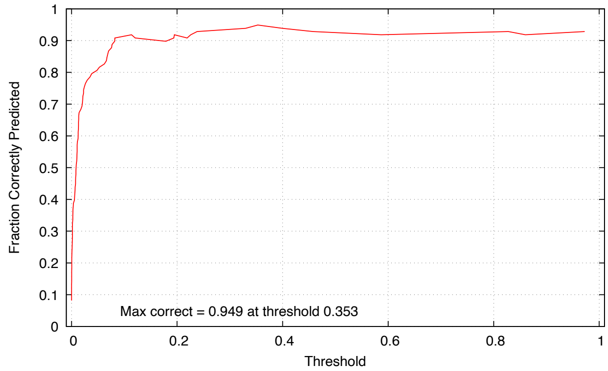
Consumer Confidence Index, two lagged quarters



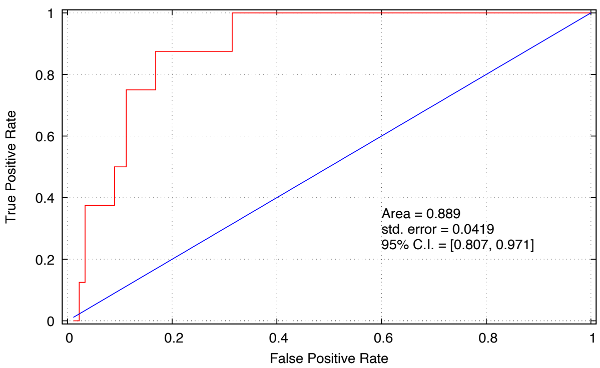


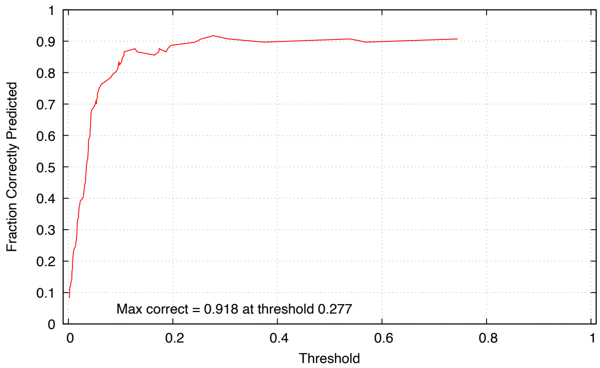
Business Confidence Index, one lagged quarter



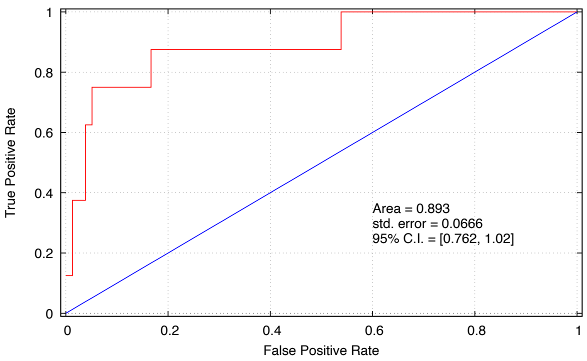


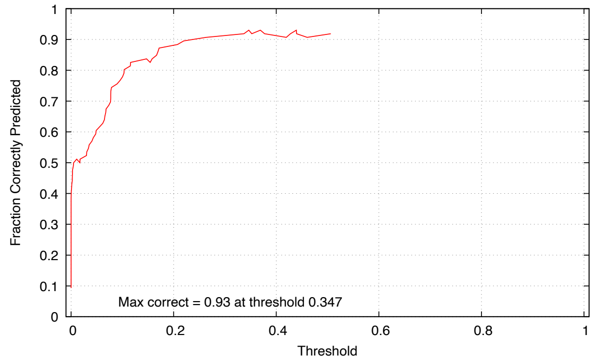
Business Confidence Index, two lagged quarters



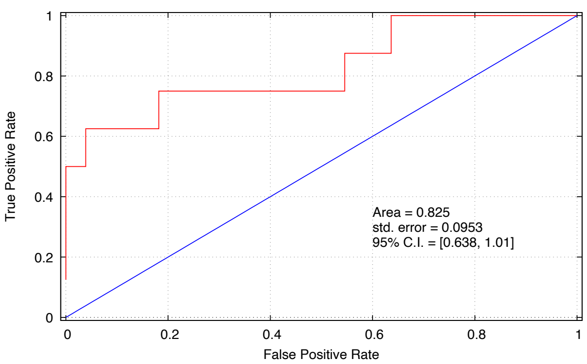


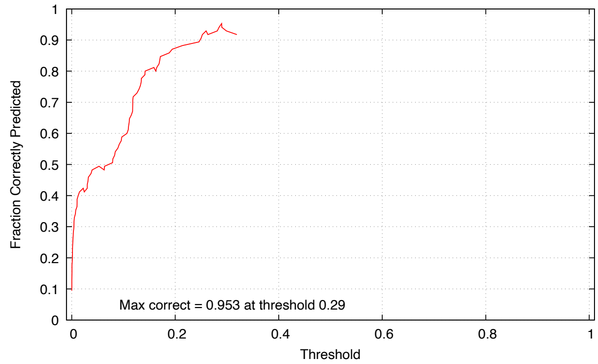
FTSE MIB, one lagged quarter





FTSE MIB, two lagged quarters





9.4 – Next table provides the full results of the estimations of the probit models for the four alternative indicators.

		1 lagged quarter	2 lagged quarters	3 lagged quarters	4 lagged quarters
BCI	Pseudo r-squared	0.468833	0.222375	0.06083	0.00382
	P-value	0.0001***	0.001***	0.066*	0.6437
CCI	Pseudo r-squared	0.420935	0.379262	0.333743	0.276358
	P-value	0.0002***	0.0003***	0.0004***	0.0008***
CLI	Pseudo r-squared	0.472719	0.129654	0.000853	0.063153
	P-value	0.0002***	0.0116**	0.8282	0.0794*
FTSE MIB	Pseudo r-squared	0.321516	0.180924	0.085284	0.024728
	P-value	0.0043***	0.0211	0.0631*	0.2718

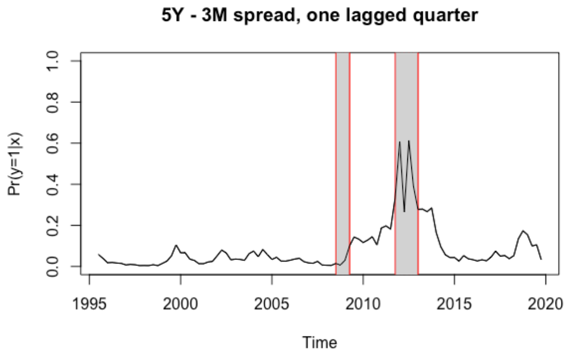
		5 lagged quarters	6 lagged quarters	7 lagged quarters	8 lagged quarters
BCI	Pseudo r-squared	0.006222	0.038034	0.072963	0.059441
	P-value	0.5684	0.1807	0.0739*	0.1004

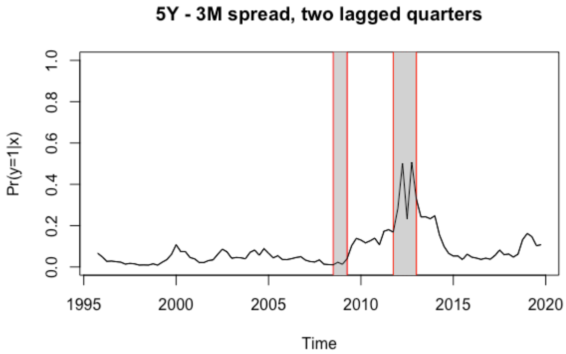
CCI	Pseudo r-squared	0.220973	0.15656	0.083726	0.033508
	P-value	0.0018***	0.0062***	0.0372**	0.1785
CLI	Pseudo r-squared	0.218447	0.38861	0.456519	0.36398
	P-value	0.0035***	0.0011***	0.0012***	0.001***
FTSE MIB	Pseudo r-squared	0.005264	0.0002	0.002947	0.00564
	P-value	0.6031	0.9185	0.6949	0.5882

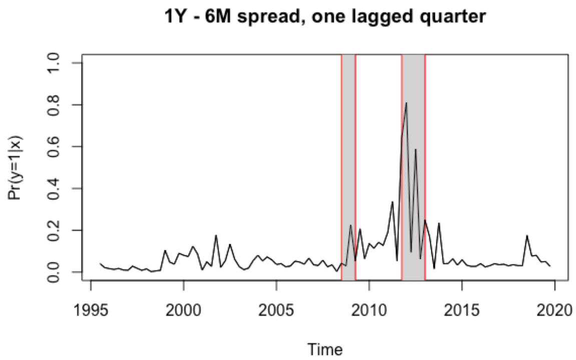
9.5 – Next table provides the ROC, AUC, maximum percentage of predicted values, threshold, false positive rates and false negative rates for the four alternative indicators.

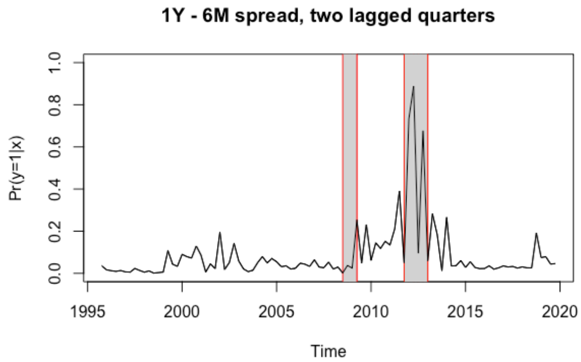
		1 lagged quarter	**2 lagged quarters**
BCI	AUROC	0.96 (0.0198)	0.889 (0.0419)
	Max correctly predicted	94.9%	91.8%
	Threshold	0.353	0.277
	False Positive rate	0.0219	0.0329
	False Negative rate	0.5000	0.6250
CCI	AUROC	0.933 (0.0321)	0.926 (0.0332)
	Max correctly predicted	94.9%	94.8%
	Threshold	0.438	0.408
	False Positive rate	0.0219	0.0219
	False Negative rate	0.5000	0.5000
CLI	AUROC	0.936 (0.035)	0.756 (0.101)
	Max correctly predicted	94.9%	91.8%
	Threshold	0.477	0.295
	False Positive rate	0.0219	0.0219
	False Negative rate	0.5000	0.75
FTSEMIB	AUROC	0.893 (0.0666)	0.825 (0.0953)
	Max correctly predicted	93%	95.3%
	Threshold	0.347	0.29
	False Positive rate	0.0659	0.0219
	False Negative rate	0.6250	0.8750

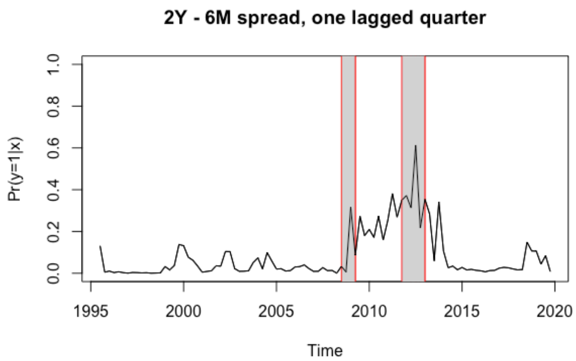
9.6 - The following graphs show the probabilities distributions for the best spread predictors (5Y – 3M, 1Y- 6M, and 2Y – 6M) and the alternative four indicators. Results are significant only considering a maximum of two lagged quarters.

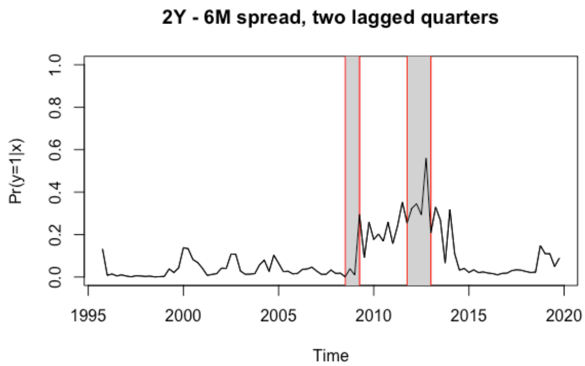


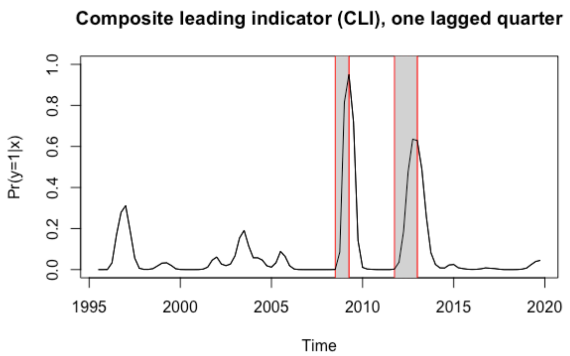


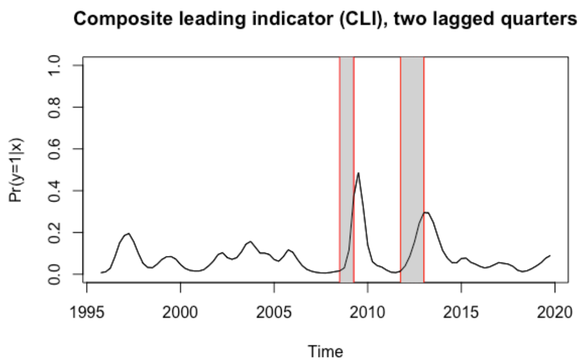




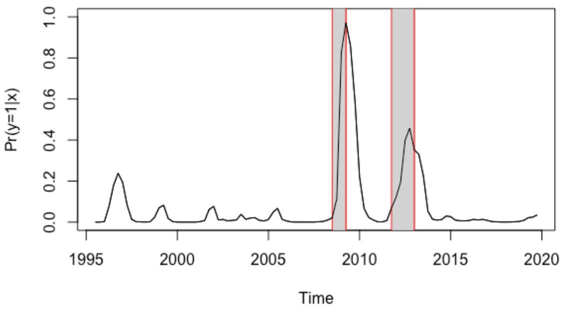




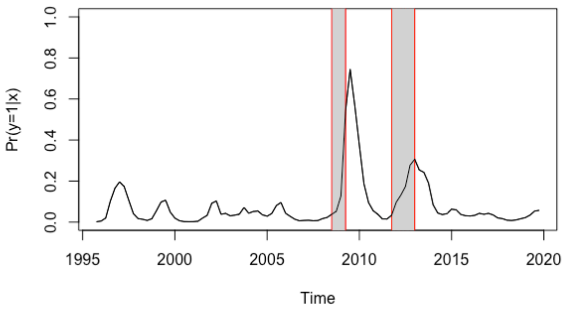




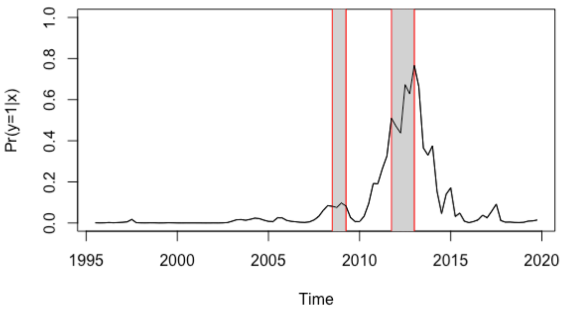
Business Confidence index (BCI), one lagged quarters



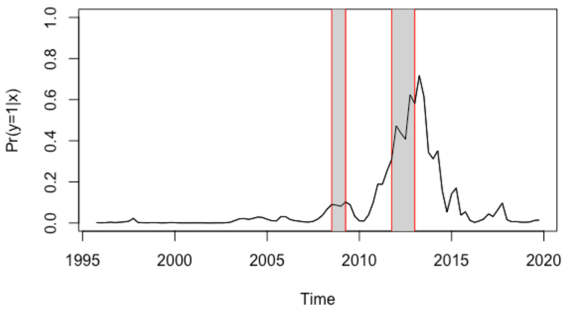
Business Confidence index (BCI), two lagged quarters



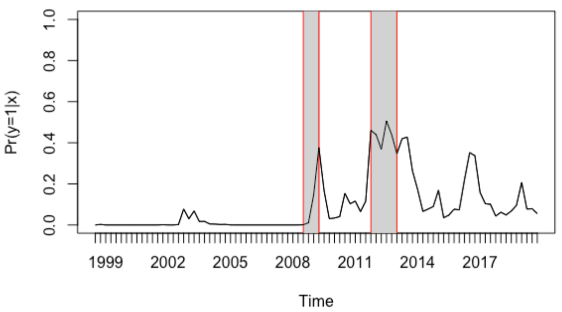
Consumer Confidence index (CCI), one lagged quarter



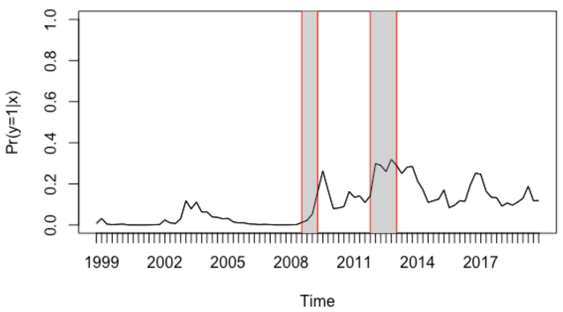
Consumer Confidence index (CCI), two lagged quarters



FTSE MIB, one lagged quarter



FTSE MIB, one lagged quarter



- Page intentionally left blank -

LUISS Guido Carli
LIBERA UNIVERSITÀ INTERNAZIONALE DEGLI STUDI SOCIALI

Department Business and Management

Chair Quantitative Methods for Management

The term spread as a predictor for recessions. Empirical evidence on the Italian case - Summary -

Professor Marco Pirra
SUPERVISOR

Professor Paolo Porchia
CO-SUPERVISOR

Nicolò Bongiovanni
CANDIDATE

Academic Year: **2019/2020**

1. INTRODUCTION

The inversion of the yield curve is an extremely rare phenomenon. When the long-term yields fall below the short-term rates, the fear of an impending recession starts to spread among the markets, motivated by a general lack of confidence in the near future. Nowadays policymakers have the possibility to rely on a vast range of methods and an enormous amount of data to forecast future market conditions, but an analysis of the yield curve movements, which proved to be valid in the past, could still represent a further and useful study. It is important to mention that the occurrence of the inversion of the term structure in a specific country does not necessarily act as predecessor for the same event in other nations. Nevertheless, for the western world, the curve inversion in the United States is usually denoted as a strong alert on the global economy’s health, and whenever this phenomenon occurs, financial markets start to tremble. The inversion of the yield curve represents a sort of “break” of the traditional theory regarding the value of time. In a traditional context, where the slope of the yield curve is positive or slightly upward sloped, investors who choose to invest in long-term bonds will earn a higher compensation. However, this reasoning is overturned when we face an inverted yield curve. In this situation, investors do not feel safe in buying short-term securities despite their higher returns. Indeed, they prefer to invest in long-term debt, because of the shared expectation of an economic downturn in the long run. Acquiring bonds with long-term rates will ensure the investors a higher return whether the policy makers will cut the interest rates, a measure that represents the traditional reaction that central banks carry out to fight a recession time.

The main purpose of this paper is to study the historical inversions of the Italian yield curve and to assess the curve predictive power concerning future recessions, since most of the literature related to this topic have a United States or European outlook. To reach our objective, we will conduct an econometric analysis using the so-called “probit model”. This method allows to assign a probability level to the recession event. More precisely, this model allows to convert the steepness of the yield curve into a probability of a recession with a specific time forecast. The model is based on the definition of two important factors, whose misinterpretation could lead to different results: recession and spread. Throughout the model development, both recession and spread definitions will be clearly addressed. Moreover, in order to gain a better understanding of the accuracy level of the term spreads’ performance, we will compare the obtained results with a bunch of alternative indicators. The largest part of the existing literature mostly refers to the past decade and beyond, providing an attractive space for more recent studies such as this one. This document contributes to the body of literature that looks at the relationship that exists between movements of term spreads (inverted yield curve) and the rise of a recession, providing an empirical evidence on the Italian case.

2. THE YIELD CURVE

2.1 - Fundamentals and Italian scenario

"The yield curve is a graph that plots the yields of various bonds against their term to maturity"[54]. This definition summarizes in a few words the main concept that underlies this subject, whose characters are bonds, debt instruments that investors are willing to buy and hold in exchange for a proper compensation defined by the interest rates. Taking into consideration the nation's debt, the line represented by the combination of the various levels of interest rates and corresponding maturities defines the yield curve. The general expectation combines longer maturity with higher yields, because of the risks faced by long-term debt holders. Among the uses and interpretations of the yield curve, one main and distinctive use stands out: it sets the benchmark for the listing of all debt market instruments (such as bank lending rates and mortgage rates). Secondly, the curve is used to anticipate future economic growth, providing market expectations about the course of future interest rates. Short-term rates are directly affected by the actions of the Central Banks; indeed, the curve is fundamental for the transmission of the monetary policy. These institutions cut or raise the short-term interest rates according to the current economic situations: usually, they cut rates to stimulate consumptions and investments, while they raise rates to prevent potential high levels of inflation. The role of the yield curve is therefore dual[55]: on one side it is essential for monetary policy purposes, and on the other side it provides relevant insights regarding investors' expectations about the future outlook of the overall economy.

Evidence on historical graphs shows that four basic representations of the yield curve have been identified: positive sloped, flat, negative sloped and humped[56]. Moreover, three studies have been developed to justify these shapes: the theory of pure expectations, the liquidity premium theory and the preferred habitat theory. In the theory of pure expectations, market expectations are perfectly reflected in the form of the curve. This theory relies on the basic assumption that investors do not have preferences for different maturities, as long as they maximize their returns. They invest according to their expectations about future rates. This process leads to the changes in the slope of the curve, and therefore, an inverted yield curve would mean that market players fear an impending recession. The same assumption (market expectations are perfectly reflected in the form of the curve) does not hold for the liquidity premium theory. In this theory, the curve reflects the future short-term

[54] Moorad Choudhry, 2019. "Analysing and Interpreting the Yield Curve". Wiley Finance Series, second edition

[55] The yield curve may be used for other reasons as well. However, the main purposes are the two described above (example of other uses: pricing of interest rate derivatives).

[56] Humped: short-term rates rise until they reach a peak in the medium term, from which they decrease.

rates plus a liquidity premium that depends on the maturity, which represents a sort of compensation for investors who are willing to buy securities expiring later in time. The liquidity premium theory is therefore consistent with an upward sloped yield curve. The third and last theory relies on the assumption that different market players have different priorities and, therefore, they have preferences for specific maturities. This theory can explain all shapes of the yield curve, and the spread will rise because of the different levels of supply and demand for specific kinds of instruments.

Since this paper seeks to prove a direct link between the inversion of the Italian yield curve and a subsequent recession, it is necessary to provide a glimpse about the current Italian scene concerning yields levels. Figure 1 shows the graphical representation of three Italian yield curves at three different dates. The light blue line is the curve on the 30th April 2020, the red line is the curve on the 30th April 2016, while the green line is the curve on 30th November 2014. We can notice that the two most recent yield curves both have the short-term rates below the zero threshold, which is a feature of the current financial markets. The green line, corresponding to the oldest date, shows a higher slope and higher yields for most of the maturities. Indeed, between 2014 and 2020 interest rates decreased significantly. As mentioned in the previous paragraph, the yield curve is used as a benchmark for different purposes. For example, it is used to set the interest rates of bank loans and mortgages. Figure 2 provides the monthly interest rates imposed on loans for two different targets, privates and companies. In the graph, both Italian values (red lines), and European values (blue lines) are showed. The continuous lines represent the interest rates payed by companies to access to bank capitals, while the dotted lines represent the interest rates payed by privates when they request a mortgage for house purchase. The first impression is that approaching recent years, an overall and sharp decrease in the rates occurred, both in Italy and in Europe. From the chart, it also stands out that nowadays borrowing money became cheaper compared to the past, both for families and for companies. Indeed, in the last ten years, central banks often decreased short-term interest rates to stimulate the economies, increase consumptions and push companies to make more investments.

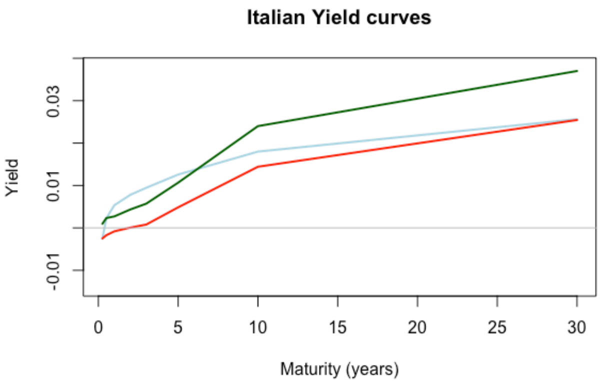


Figure 1: Author representation (data provided by ISTAT)

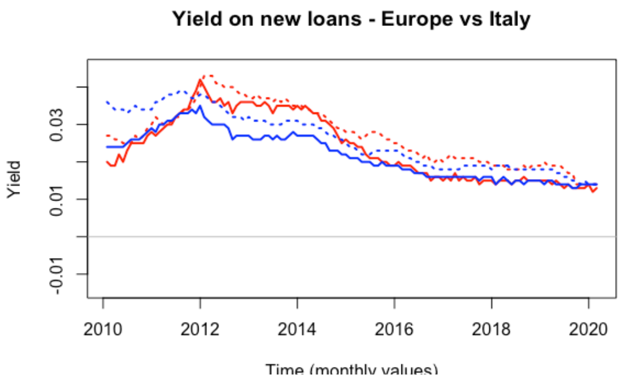


Figure 2: Author representation (data provided by ECB)

3. LITERATURE REVIEW

In previous studies, researchers initially focused on the relationship between the curve and real activity, a topic analyzed by several authors, including Harvey (1988), Laurent (1988, 1989), Chen (1991), and Estrella and Hardouvelis (1991). To have a more recent outlook, it is possible to refer to reports published in the last few years. In the next paragraph, researches conducted by David Miller (2019), Johansson and Meldrum (2018), Rudebusch and Williams (2008), O. Emre Ergungor (2016), Bernard and Gerlach (1996) and Fabio Moneta (2003) are briefly described.

David Miller (2019), analyzing the predictive ability of several term spreads, demonstrated that it is hard to find a unique best predictor, and that the time horizon of the forecast is determinant to assess the accuracy of the spread. Johansson and Meldrum (2018) tested the predictive power of the spread applying three different variants of the probit model: the first considering the term spread, the second considering the first three principal components of yields, and the third adjusting the spread for the term premium. These different models proved to be best fit for different forecasts, but as general result, the term structure performed as a great predictor. Two other economists, Rudebusch and Williams (2008), compared the predictive power of the term spread with the results obtained by the Survey of Professional Forecasters (SPF)[57]. The main result showed that, in several occasions, the term spread proved to outperform the SPF forecasts. Following the idea that current economic conditions identify specific and unique scenarios, Emre Ergungor (2016) added two new variables to the regression: the credit spread and the growth of corporate profits[58]. The results combining these new indicators provided positive relations between the variables and the ability to predict recessions in the short future. Each of these researches had a US background. Focusing on Europe, it is a bit more difficult to find studies that analyze the same topics. However, authors including Bernard and Gerlach (1996) and Estrella and Mishkin (1996) focused on different sets of European countries, finding evidence that the term spread revealed to be a good indicator in this region as well. Moreover, Fabio Moneta (2003) also tested the predictive power of the term spread across Europe, but with an innovative technique: he retrieved national data and he aggregated them to obtain a Euro area series[59]. The predictive ability of the spread was then compared with other economic variables, and the results showed that the term spread was once again the best predictor among the sample.

---

[57] This survey is a quarterly analysis conducted by professionals of the economic sector, who are asked to provide their thoughts concerning several subjects, such as expected inflation, GDP growth and unemployment rate.

[58] The credit spread defines the difference in yields between the rates of securities that corporate borrowers pay, for example, the difference between rates of a high-quality borrower, rated AAA, and a bad-quality borrower, rated BBB. The change in growth of corporate profits refers to inflation-adjusted quarterly profits data between 1970 and 2016.

[59] Based on: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Spain and Portugal.

4. METHODOLOGY

4.1 - Research question and data collection

As far as we are aware, studies or researches which focus exclusively on the Italian landscape and Italian yield curve are not present in the literature. To study the effects that different yield curve spreads have on the observed GDP, we propose a model whose objective is to translate the steepness of the yield curve into a probability of a recession in the future. Our first goal is to find the spread that proves to have the highest explicative power regarding a future recession. Therefore, we will analyse the term spreads' impacts on GDP in different time horizons. The model that we use is the "probit model", and will help us in testing the hypothesis of our work, which is: 'The term spread, based on the Italian yield curve, may be considered as a reliable predictor for impending national recessions or economic downturns'.

The first step that we need to take for the implementation of the model is to define the yield spread. For the purpose of our work, the issuer considered is the Italian Government. Although there is a general acceptance that the spread between the 10 years and 3 months interest rates may represent the most accurate tool to predict recessions, we have decided not to rely on this hypothesis. Therefore, the instruments used in our work are: the 3-months interbank rate (EURIBOR[60]), the 6-months and 12-months BOT ("Buoni ordinario del Tesoro"[61]), the 2-years CTZ ("Certificati zero-coupon"), the 5-years and 10-years BTP ("Buoni del Tesoro Poliennali"). In this work, we decided to consider the 3-months interbank rate as the shortest instrument, because both the overnight and 1-month rate could not represent proper indicators for market expectations, due to their direct control by the European Central Bank. The fifteen combinations of spreads that we tested are the following: 6M - 3M, 1Y - 3M, 2Y - 3M, 5Y - 3M, 10Y - 3M, 1Y - 6M, 2Y - 6M, 5Y - 6M, 10Y - 6M, 2Y - 1Y, 5Y - 1Y, 10Y - 1Y, 5Y - 2Y, 10Y - 2Y, 10Y - 5Y (Y = years, M = months). The first date in which there is availability for all these instruments is January 1995; hence our study covers the time frame between January 1995 and December 2019, providing a focus on more recent data and economic events.

The second element that we need to define for our analysis is the recession. Recessions' boundaries are always difficult to identify. Movements of national Gross Domestic Product are based on data that are often revised according to different calculations, thus finding exact dates that mark

[60] Euro interbank offered rate (EURIBOR) is a benchmark rate, daily calculated, which points the medium rate of the financial transactions between European banks.
[61] Banca d'Italia also issues 3-months BOT, however, very spare data were available about this instrument.

the beginning and the end of a recession is a challenging task. Moreover, the European system does not provide a unique and unmistakable definition of these phenomena. Given the Italian shade of our study, the definitions provided by Borsa Italiana[62] were taken as reference. This institution defines the so-called technical recession and the economic crisis. A technical recession occurs when the gross domestic product shows a negative variation for at least two consecutive quarters[63], while an economic crisis occurs whether the negative variation of the economic output reports less severe values. We chose to analyze both definitions according to two reasons. Firstly, because the goal of this paper is to understand whether the term spreads have some predictive power in predicting an economic downturn, may it be very severe or slightly severe. Secondly, because considering exclusively the definition that identifies the technical recession, only few data of the sample would be used, since in the last twenty-five years few large recessions occurred. The GDP measures provided by the database show quarterly values of Italian gross national product at current prices, and data have been seasonally adjusted[64].

In order to understand the accuracy of the term spreads' predictive power, we compared the results with the analysis of other alternative indicators. These alternative measures are the Composite Leading Indicator (CLI), the Consumer Confidence Index (CCI), the Business Confidence Index (BCI) and the FTSEMIB[65]. The Composite Leading Indicator is a composite index, which monitors the movements of different trends occurring in a specific nation. The trends considered are time series which exhibits strong relationships with movements of GDP, and that cover a large range of key statistics. The Consumer Confidence Index focus on consumers' and families' behaviors. These categories are asked to provide their opinion regarding their expected financial situation, their sentiment about the overall economic health, unemployment and savings trends[66]. The Business Confidence Index is built with a similar approach, but the target of the analysis is composed by enterprises and players of the industrial sector. These subjects provide their opinion regarding developments in production, order and stocks of finished goods, selling prices, export conditions, and competitive positions. This index is useful to understand the current trends concerning the production side of the economy, that usually represents a good benchmark to monitor output growth. Lastly, the

[62] Borsa Italiana S.p.A. is the national institution managing Italian stock exchange and financial markets.
[63] This definition has often been considered as the benchmark to identify recessions. The main reason is that temporary economic slowdowns may not turn into recessions and may be country specific.
[64] Seasonally adjusted data are net of seasonal fluctuations. It means they are corrected for calendar effects.
[65]CLI, CCI and BCI data are provided by the Organization for Economic Co-operation and Development (OECD). This organizations monthly studies these indicators for a large set of countries (40 countries all over the world).
[66] For more information about the surveys conducted for both the CCI and BCI, please refer to the document "The Joint Harmonised EU Programme of Business and Consumer Surveys", published by the European Commission.

FTSE MIB is the most significant Italian share price index. This index refers to the basket made up by the 40 most capitalized companies, whose shares are daily traded on the stock exchange markets.

4.2 - The analysis: the probit model

The probit model is a binary outcome model. It is classified as a non-linear regression model specifically designed for binary dependent variables[67]. In these models, dependent variables can only assume two specific values, usually represented by the values ‘yes’ and ‘no’ (that are generally turned into ‘one’ and ‘zero’ when it comes to statistical modeling). Equation [1] describes the probit model with one independent variable X:

$$\Pr(y_{t+k} = 1|X_t) = \Phi\,(\beta_0 + \beta_1\, x_t) \qquad [1]$$

In this model, Y represents the binary variable, X represents the independent variable and Φ represents the cumulative distribution function of a standard normal distribution. The distinctive characteristic of this regression is that it does not model the value of Y, but the probability that Y would take the value of 1, according to the observation of the independent variable X. When estimating probit models, one of the measures useful to evaluate the goodness of fit is represented by the percentage of correctly predicted values, which shows how many observations were correctly predicted according to the actual observations. The second measure that it is useful to estimate the statistical goodness of fit is the so-called Pseudo R-squared[68]. Values of this measure must be considered with a different approach compared to typical values of R-squared estimations. Indeed, the estimated model provides a good fit if this value falls into the range of [0.2 - 0.4][69]. The predicted values computed by the models are based on a classification rule, that assigns to the binary dependent variable the value of 1, if the estimated probability for the specific observation is higher than a defined threshold. To help us in understanding the thresholds that maximize the correctly predicted values for each spread, we will compare ROC (Receiver operating characteristic) and AUC (Area under the ROC curve, also called AUROC) estimations. Based on these estimations, a distinction between false positive and false negative rates will show the accuracy of the spreads predictions. Recalling equation [1], we can now consider the inputs of our analysis. The input to this calculation is the value of the

[67] Probit models are often used jointly with another similar methodology: the logit models. The difference between them lies in the distribution functions used to model probabilities: the standard normal distribution is used for the probit model while the logistic distribution is used for the logit model. However, results of the two models are usually very similar.
[68] Also called the McFadden R-squared.
[69] McFadden himself suggested this range of values as a measure of a good fit. In addition, please refer to Louvier et al, 2000: “Stated Choice Methods”.

term spread observed at time *t*, which represent the variable X. Variable Y, which is the binary variable, represents the recession. To be more precise, it assumes the value of 1 when a recession has occurred in time $t + k$[70], and 0 otherwise. Since our purpose is to focus on the predictive ability of the term spread, we took into considerations term spread impacts on eight different lagged quarters. Moreover, since our purpose is not to focus exclusively on recessions, but on economic crises as well, the two definitions of recessions mentioned in the previous chapter will be analyzed. Therefore, having at our disposal 15 combinations of spreads, 8 lagged quarters and 2 definition of recessions, our analysis will be conducted on 240 probit models.

5. RESULTS

A concise comparison of the most important results of our study is provided in Table 1. This table is divided in two sections: in the upper part, results calculated with term spreads as explanatory variables are listed, in the lower part the results shown were obtained considering the four alternative indicators as explanatory variables.

The three term spreads that proved to be the best predictors are the 5Y – 3M, 1Y – 6M, and 2Y – 6M spreads, and this result is consistent both when predicting an economic crisis, and when predicting a recession[71]. Comparing the results for the two definitions of recession, we can notice that these spreads provided a better forecast when considering the definition which describes the recession as a decline in GDP for two consecutive quarters. Each of these three spreads provided the highest accuracy according to three different measures: the 1Y – 6M spread with two lagged quarters showed the highest pseudo r-squared value (0.267), the 2Y – 6M spread showed the lowest false negative rate (0.5), and the 5Y – 3M spread showed the highest percentage of correctly predicted values. However, since the pseudo r-square value of the 2Y – 6M spread is almost the same compared to the 1Y – 6M spread's (0.265 vs 0.267 respectively), and that the percentage of correctly predicted values is almost the same compared to the 5Y – 3M term spread's (93.9% vs 94.9% respectively), we will refer to the 2Y – 6M as our overall best spread. The second section of the table provides the results obtained by the estimations of the probit models considering the four alternative indicators. These models show better goodness of fit compared to the previous results; indeed, pseudo r-squared values present very high numbers. However, comparing the percentages of correctly predicted values, the false positive and false negative rates, the results are very much similar, if not worse, to spreads' estimations.

[70] *K* refers to the lagged quarter. In our analysis, *k* will assume values from one to eight.
[71] Considering the second analysis (predicting a recession), the 5Y – 3M spread provided a pseudo r-squared value slightly lower compared to the 5Y – 6M (0.219 vs 0.22), however, we pointed the former as one of the best spread because the false positive rate and the percentage of correctly predicted values showed better estimations.

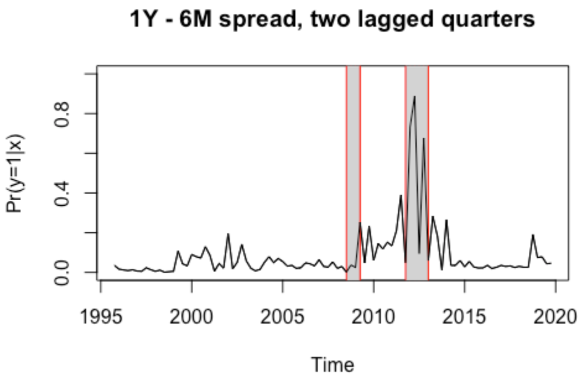
Moreover, the results of these indicators are very good when the time horizon considered is exclusively one lagged quarter; extending the horizon, pseudo r-squared values largely decrease for each of them, providing even worse estimations compared to the models based on the term spreads. Among the set of alternative indicators, the Composite Leading Indicator with one lagged quarter showed the best estimations, suggesting that a composite analysis of macro variables fluctuations had a better predictive ability compared to yield curve fluctuations.

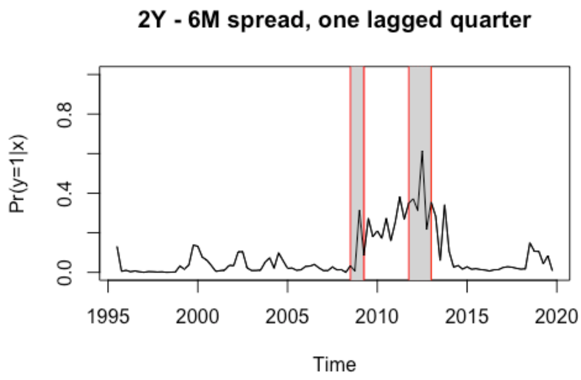
Table 1: Results

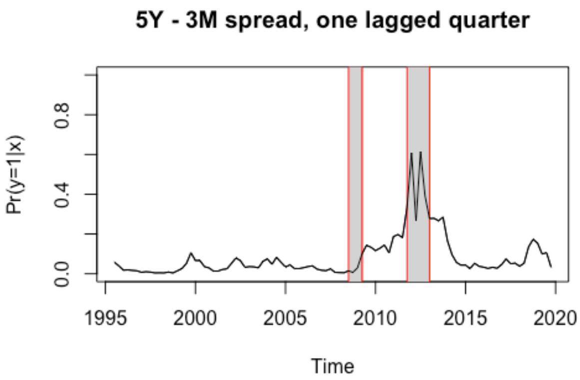
			Pseudo R2	p-value	Predicted values	False positive rate	False negative rate
Predicting an **economic crisis**	**1Y - 6M**	**1 lag**	0.227536	0.0004***	89.8%	0.0233	0.6154
		2 lags	0.082932	0.0161**	89.7%	0.0000	0.8462
	2Y - 6M	**1 lag**	0.219885	0.0002***	90.8%	0.0233	0.6154
		2 lags	0.090697	0.0107**	88.7%	0.0349	0.6923
	5Y - 6M	**1 lag**	0.168826	0.0008***	87.8%	0.0349	0.6923
		2 lags	0.065062	0.0286**	87.6%	0.0000	1.000
	5Y - 3M	**1 lag**	0.184896	0.0005***	90.8%	0.0000	0.7692
		2 lags	0.096502	0.0088***	90.7%	0.0000	0.6923
	10Y - 3M	**1 lag**	0.144501	0.0017***	90.8%	0.0233	0.6154
		2 lags	0.072544	0.0217**	89.7%	0.0000	0.7692
Predicting a **recession**	**1Y - 6M**	**1 lag**	0.212865	0.0016***	92.9%	0.0329	0.6250
		2 lags	0.267187	0.001***	94.8%	0.0000	0.7500
	2Y - 6M	**1 lag**	0.264991	0.0008***	93.9%	0.0329	0.5000
		2 lags	0.227596	0.0015***	93.8%	0.0329	0.5000
	5Y - 6M	**1 lag**	0.220905	0.0019***	92.9%	0.0329	0.6250
		2 lags	0.102251	0.023**	92.8%	0.0000	1.000
	5Y - 3M	**1 lag**	0.219154	0.0014***	94.9%	0.0000	0.6250
		2 lags	0.161591	0.0048***	95.9%	0.0000	0.6250
	10Y - 3M	**1 lag**	0.177166	0.0036***	94.9%	0.0109	0.6250
		2 lags	0.100225	0.0234**	94.8%	0.0000	0.6250

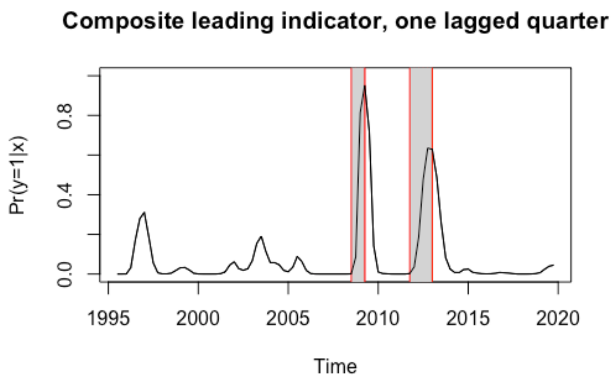
Predicting a **recession**	**BCI**	**1 lag**	0.468833	0.0001***	94.9%	0.0219	0.5000
		2 lags	0.222375	0.001***	91.8%	0.0329	0.6250
	CCI	**1 lag**	0.420935	0.0002***	94.9%	0.0219	0.5000
		2 lags	0.379262	0.0003***	94.8%	0.0219	0.5000
	CLI	**1 lag**	0.472719	0.0002***	94.9%	0.0219	0.5000
		2 lags	0.129654	0.0116**	91.8%	0.0219	0.75
	FTSEMIB	**1 lag**	0.321516	0.0043***	93%	0.0659	0.6250
		2 lags	0.180924	0.0211**	95.3%	0.0219	0.8750

Results provided above are helpful to conclude that there at least three different spreads that may be considered as good predictors, hence it is possible to affirm that these results support the paper's hypothesis. Since comparing statistical figures may result tricky, a graphical representation where the probabilities of recession are plotted may be helpful. The aim of these representations is to provide a concrete vision of the actual level of probabilities estimated according to the time horizon. The graphs below are quite intuitive, and they probably represent the most representative final output of the entire study. The black lines draw the trends of the probability scores, while the grey shaded areas define the two recessions occurred in our time horizon.









Some conclusions may be derived from the observation of these graphs. Each term spread correctly predicted with good accuracy the Sovereign Debt crisis occurred between the last quarter of 2011 and the first quarter of 2013. Indeed, in each graph we can observe that the probability levels heavily increase approaching this specific recession period. Another aspect that catches the attention is the difference between the probability distributions of the spreads' estimations. The 1Y – 6M and 2Y – 6M plots present more jagged lines, with more up and downs, while the 5Y – 3M shows a smoother distribution. This difference may be explained by the composition of the explanatory variables. The rate of returns of the instruments with shorter maturities are way more likely to be similar, and the difference among these returns may turn negative more often. On the other side, it is almost a rare event to assist to an inversion of the rate of returns of two securities which have very distant maturities (i.e. 3-months interbank rate and 5-years BTP). This difference represents the basic logic that allow us to state that the 5Y – 3M spread may be considered more suited to predict recessions. In fact, a negative spread considering these debt instruments is a rare event, and thus it may represent a signal for an unusual situation of the overall economy. Switching the focus on the comparison between different explanatory variables, the most notable difference between the results obtained with the spreads and those obtained with the Composite Leading Indicator, is the ability that the CLI proved to have in predicting the first recession, in 2008-2009 years. Each of the term spreads did not succeed in predicting this economic crash with good accuracy, even if probability's levels slightly increase for each of them. However, we also need to mention that observing the CLI's graph with a more rigorous eye, we may see that both recessions were predicted late. Indeed, the two peaks of the probability levels of the last graph are reached in correspondence of the end of both recessions. Nevertheless, this consideration should not surprise. If we think to the fundamental elements of the Composite Leading Indicator, such as industrial production, unemployment rate or import and export balance, it is reasonable to expect a good but late prediction; indeed, all these elements could be defined more as consequences of a recession than as predictors.

6. DISCUSSION

From an overall point of view, our analysis provided results that could be described significant and consistent with our initial expectations. Nevertheless, before being able to define the Italian yield curve a great predictor, it is important to understand the limits of our work and the areas of improvement. Although the statistical side of the work provided good results, there are a few topics and considerations that need to be mentioned. When referring to an inversion of the yield curve, the doctrine usually refers to the basis point difference between a security with a long maturity and another security with a very short maturity. Indeed, most of past literature is based on the analysis of the 10Y – 3M spread, which in the United States proved to forecast past recessions with great accuracy. In order not to limit our analysis to that specific element, we estimated the models for a large set of combinations. Among our results, the best estimations are provided by the 2Y – 6M spread, which is made up by the rates of return of two securities with, respectively, a short/medium and a very short maturity. Therefore, it would be quite pretentious to consider this spread as an inversion of the yield curve, since no long-term rates are part of the estimator. However, if we don't focus our attention simply on this spread, our consideration may change. Indeed, the 5Y – 3M spread with one lagged quarter is one of the three best estimators of our sample as well. A bond with a five years maturity, maybe could not be described as a long-term instrument such as the ten years government bond, but at the same time, it can't be described as a short-term debt obligation[72]. We could thus identify five years as a medium/long maturity. With this legit interpretation, it is now correct to state that the inversion of the yield curve based on the 5Y – 3M spread may be defined as a good predictor for recessions in Italy. Another topic worthy of attention is the low accuracy shown for the prediction of the first recession, which represents one issue concerning the robustness of our results. However, it is difficult to state whether this misinterpretation was due to the poor performance of the model in forecasting this specific recession, or whether it was due to issues related to data availability. We mention data availability because probit models of our analysis were estimated on a quite restricted amount of data. Indeed, referring to the statistical estimation, it has been difficult for our models to correctly predict the values of the dependent variable, considering that this variable was observed only eight times among the whole sample. This scarcity of data may have complicated the estimation of the recession probabilities, and therefore it probably represents the main limit of our study. Moreover, econometric models whose purpose is to forecast GDP fluctuations, usually rely on a vast set of variables. Using exclusively the term spread to predict declines of national accounts may therefore lead to correct, but improvable results. Finally, it is important to provide a comment regarding the predictive power of the curve in the current environment. As explained in the first chapter, the signals provided by the yield curve are very sensitive to financial markets conditions. Nowadays, financial markets present peculiar characteristics, and the activity that they support is massive. Therefore, it became more and more complicated to understand if changes in financial markets are due to technical factors or economic fundamentals. Another topic of discussion may be the current situation of interest rates. Financial markets are experiencing a period of extremely low interest rates, and this trend is not expected to end in the short term. Therefore, applying the study of the inversion of the yield curve, which proved to be very powerful in past macroeconomic conditions, may not be totally effective. Moreover, a recent study conducted by the European Central Bank, analyzed the recession probabilities estimated with adjusted probit models. In this analysis, the term spread was adjusted for the effect of different variables, such as the impact of the quantitative easing program undertaken by the Federal Reserve, or the spillover effect of US asset purchases by the European Central Bank. The result of this analysis showed that the probability estimations based on the yield curve term spread became less significant when several adjustments were made. This result should not surprise too much. Indeed, the quantitative easing programs and the measures adopted by

[72] The traditional classification of time horizons is the following: short-term rates are those with maturities lower or equal to twelve months (in our case, the interbank rate, the 3-months BOT and 6-months BOT), medium-term rates have maturities included between two and five years (in our case, the 2-years CTZ and the 5-years BTP) and long-term rates have maturities higher than five years (in our case, the 10-years BTP).

central banks in the last ten years have distorted interest rates. This analysis is a more advanced kind of study, but it could represent an interesting suggestion for future research related to our paper.

7. CONCLUSION

Our analysis was drove by the interesting relationship existing between the movements of the yield curve and the start of a recession. Assuming that this relationship was something certain and unavoidable, an inversion of the curve would be considered with more attention. Indeed, in the United States, where this relationship proved to be accurate several times, this event receives much more credit. The idea to test this hypothesis in Italy derives from the low attention that these events usually receive in our country. The statistical models that we used to translate the observations of real phenomena into specific values had a concrete objective: to test if these events could represent a useful signal to partially avoid the disastrous consequences that a recession necessarily brings. Our results proved that, although an inversion of the curve may not be considered as a universal principle to predict the start of a recession, it still represents an alarm for economic instability. Besides considering the term spread between the 2-years and 6-months securities, which showed the best estimations, but which cannot be described as a proper inversion of the curve, the results showed that an inversion between the 5-years and 3-months government bonds correctly predicted most of the observations of the sample. These results were then compared with those obtained by the estimations of other indicators, such as the CLI, the CCI, the BCI and the FTSEMIB. These variables, apart from providing better statistical goodness of fit compared to the spread's estimations, did not particularly improved the results previously obtained. Only the Composite Leading Indicator succeeded in predicting recessions with higher accuracy compared to the term spread, even though it provided slightly late predictions. Considering the weaknesses of our study, the main drawback was represented by the limited availability regarding the length of historical series and the quantity of recession periods of our sample, which may both have affected the robustness of our results. However, we believe that this analysis proved to be coherent with our initial expectations. Moreover, it may be considered as a starting point for future literature. Future research may focus for example on the impact that current macroeconomic conditions are having on the predictive ability of the Italian term spread. Finally, although our results were significant, we need to mention that in our opinion the United States' yield curve could still represent a good benchmark, because of the predominant position that the US still has among global economies, and because of the large interconnections that nowadays exist among countries on a global scale.