Distance to default as a measure of default risk

Author: Fernando Luna Goris

Degree: Bachelor of Economics

Tutor: Lidija Lovreta

June 9th, 2015
Companies are facing increasing uncertain times, key investment decisions and challenging financial needs. Now more than ever it is important to understand both the financial decisions of the business corporations as well as how they connect.

In this paper I would like to show the basics to understand complex pricing of assets such as derivatives. Since central, commercial and investment banks use lots of models to price securities and forecast financial markets I wanted to research on the basic theorems and ideas needed to do so.

The subject that I chose to study is the combination of the representative companies included in the *EURO STOXX 50* index, but excluding the financial institutions due to high leverage ratios that may lead to wrong interpretations. In order to estimate the credit (default) risk on every company's assets, I used two different approaches and then I compared and chose the one that was more precise with the results obtained.

Moody's KMV approach to measure distance to default is the most commonly used, but in the paper I show that it is not always the best option for estimating this measure.

In order to provide evidence about the functionality of the distance to default, I compare the estimations gathered from the empirical research with credit default swaps (CDS) that are linked to the sample companies of the project.

Two different sets of linear regressions were done. The first set of regressions is the cross-sectional one, while the second set is done with a sequence of time-series data points. By the comparison of these regressions, it’s expected to conclude whether the traditional and most widespread method is the best possible way to perform such estimations or if there are better ways to proceed and therefore, that may give us even better results.
INDEX

1. INTRODUCTION .......................................................................................................................... 5
   1.1. Motivation and sources of information ................................................................................. 5
   1.2. Literature review ..................................................................................................................... 6
   1.3. Definitions, objective and hypothesis .................................................................................... 8
2. METHODOLOGY ............................................................................................................................ 10
   2.1. KMV Overview ....................................................................................................................... 10
   2.2. Data collection ....................................................................................................................... 10
   2.3. Calculating the Default Point ............................................................................................... 12
       2.3.1. KMV’s Default Point ....................................................................................................... 12
       2.3.2. Nominal Value Default Point ......................................................................................... 12
   2.4. Calculating Market Value of Assets ..................................................................................... 13
   2.5. Annualizing Assets Volatility & Standard Deviation (Assets Risk) ................................. 13
   2.6. Calculation of the Distance to Default ................................................................................ 14
       2.6.1. Moody’s-KMV Approach .............................................................................................. 14
       2.6.2. Nominal Debt Approach .............................................................................................. 14
   2.7. Getting the averages ............................................................................................................. 15
       2.7.1. Performing correlations .................................................................................................. 15
   2.8. Regressions ............................................................................................................................ 15
       2.8.1. Cross-sectional regression ............................................................................................. 16
       2.8.2. Time-series regression .................................................................................................. 16
3. DATA ............................................................................................................................................ 18
   3.1. Data selection ........................................................................................................................ 18
   3.2. Statistics and relevant information to subject matter ......................................................... 19
4. EMPIRICAL RESEARCH ............................................................................................................... 23
   4.1. Getting the averages ............................................................................................................. 23
       4.1.1. Average per year (Cross-sectional regression) ............................................................... 23
       4.1.2. Average per day (Time-series regression) ..................................................................... 24
   4.2. Leverage ratio ....................................................................................................................... 25
   4.3. Comparing DTD P with DTD KMV averages ..................................................................... 25
       4.3.1. Supersector comparison and analysis ............................................................................ 26
4.3.2. Country comparison and analysis ................................................................. 28
4.3.3. Comparing DTD P and DTD KMV with EURO STOXX 50 index .......... 30
4.4. Correlations ........................................................................................................ 31

5. RESULTS OBTAINED .......................................................................................... 32
5.1.1. Results: Cross-sectional regression ............................................................ 32
5.1.2. Results: Time-series regression .................................................................... 34

6. CONCLUSIONS .................................................................................................... 37
6.1. Evidence ............................................................................................................. 37
6.2. Difficulties and further research ...................................................................... 39

7. BIBLIOGRAPHY .................................................................................................. 41
7.1. Academic papers ............................................................................................... 41
7.2. Electronic texts ................................................................................................ 41
7.3. Online databases .............................................................................................. 42
7.4. Software ............................................................................................................ 42

APPENDIX ................................................................................................................. 43
1. Transforming the Default Point: Linear interpolation ....................................... 43
1. INTRODUCTION

1.1. Motivation and sources of information

In today’s world, with an economical environment that seems to be almost purely defined by the ups and downs of the financial markets, I wanted to work specifically on something that really matters for any company that faces risk when states of nature are dynamic and influential on company’s decisions.

Any company that wants to grow in a highly competitive market has to face many crucial challenges from the very first moment they enter the market. One of these challenges is dealing with risk. Taking care of risky economical situations has to be well done and strategically planned.

In this paper I wanted to research on how companies establish their models and calculations to estimate those decisions that carry risk. I was particularly interested in the understanding of why companies fail so frequently and what weapons do they have to defend themselves from failing. Being more explicit, I wanted to investigate why some companies are able to face greater amounts of debt than others and what are those consequences of facing risky financial decisions. When a company decides to grow by the use of leverage it should be riskier than growing with just your own assets, but how can we determine it?

It is possible to do it with financial models that allow us to see how far a company is from failing. So this is apparently a very good tool since the company is able to act with decisions by knowing if their actions are leading the company to a good direction or if they are making the company go broke.

But are these models well defined? Are there better methods to estimate the effect of being subject to financial risk? This is what I wanted to project on this paper since I thought that if companies tend to fail it is because the methods they use are possibly redundant for its purpose and that maybe other approaches can lead them to take better decisions in the future.
Prior to the undertaking of the project, I had to research on different aspects and concepts that I considered relevant to the aim of the paper and in order to do so, I red chapters of books, academic articles and financial encyclopaedias such as Investopedia. My tutor, who is an expert in the field, has also been a great source of information and has helped me enormously on the methodology of the project.

The compilation of data required for the purpose of this report was done with Thomson Reuters Datastream, which made possible to gather big amounts of data in simple steps that allowed me to save a lot of time throughout the pursuit of the paper. Other software used for the subject matter has been MS Office, EViews, Mozilla Firefox and Adobe Acrobat Reader.

1.2. Literature review

One of the academic papers that origin the main goal of this paper is the one written by Crosbie and Bohn (2003) in which they describe the procedure needed to estimate “how far” are companies from default i.e. how to make probabilistic assessment of the likelihood of default. In essence, Crosbie and Bohn (2003) explain the distance-to-default measure as defined by Moody’s KMV. Jessen and Lando (2013) further support this sort of measure as a prosperous way to rank companies by how distant they are from actual default. They show that despite the fact the distance-to-default measure is based on simplifying assumptions, it turns out to be a good predictor of default and robust to model misspecifications.

The paper written by Alonso, Forte and Marqués (2006) describe the importance and relationship between CDSs and the distance-to-default measurement. Their research conducted points out the fact that CDSs are highly related to the financial situation of a company and thus, I’ll assume in the forthcoming estimations that they are a good point of view to compare with the distance to default.

Wang (2012) outlines that financial companies have a very unique way to structure their capital and that this fact has an important influence on how their leverage ratios differentiate from other industries. This idea would force me to avoid using financial
companies such as banks for the analysis of default risk from the EURO STOXX 50 index.

Some quantitative adjustments had to be done throughout the fulfilment of this paper such as mathematical modifications, volatility estimations, econometrical procedures, etc. One of these measures taken is the linear interpolation that according to the College of Engineering & Technology from the Youngstown State University, it is a way for constructing new data points that aren’t specified, but they actually fit with the sample used. Given a curve, the function will include estimated points that would form a sequence of both known points and unknown points that together would structure a workable set of data points. This measure will be necessary to be performed in order to link balance sheets since we want to extract non-dynamic information from them, such as liabilities.

Considering that companies normally supply one audited balance sheet per year -at the end of the year-, by the help of this mathematical concept described I will have to adjust each balance sheet in such a way that I create some sort of “fictitious” balance sheets per each day since the prices of stocks change in a daily basis.

Other relevant adjustments were conducted in some phases of the paper because of the limitations of the sample. Granger and Newbold (1973) specified the problematic of persistence in time-series regressions that may end up being spurious regressions when Durbin-Watson’s statistic tends to be around 0 (presence of unit roots). Bohn (2005) also researched about this complication and explained a way to test the presence of unit roots by the use of the Augmented Dickey-Fuller test on non-stationarity processes such as the one described previously. This test would allow us to either reject the possibility encompassing unit roots or fail to reject it.

In order to overcome this obstacle, Granger and Newbold (1973) illustrated some ways to fix the problem that could head us to wrong interpretations and in order to do so they recommend measures such as the inclusion of lagged variables or differences in the regressions.
1.3. Definitions, objective and hypothesis

An essential concept that will be constantly taken into account throughout this project is “default risk”. It is defined as the ambiguity encompassing a company’s ability to face debts and other obligations. When a firm reaches the point where it is unable to pay its debt and must stop its economical activity we say that the firm has gone into bankruptcy.

Thus, supposed the fact that a given company faces trouble to pay its debt when the total number of the company’s assets reaches a certain minimum level, according to Anastasija (2012) it is said that the company reached its default point (value of the debt).

Moody’s KMV model has probably been historically one of the most widely used methods to estimate how distant any given company is from its default (point). This concept is known as distance-to-default (DTD) and by contrasting empirically the values obtained, it is expected to be able to draw conclusions for the sample I worked for. The study from Crosbie and Bohn (2003) tells us that this model is an extended version from the Black-Scholes-Merton (1973) framework, produced by Stephen Kealhofer, John McQuown and Oldrich Vasicek.

The second step in the project is to compare the two approaches (KMV and Nominal Debt) with Credit Default Swaps (CDS). We should be able to observe which of the two better explains the market’s reality.

CDSs are considered insurance in the possible event of default because by purchasing a CDS, the buyer transfers the risk that the security will default. Therefore, they are a measure for credit risk of a given company and that is why it is going to be used in this project. The price of this “insurance” is known as credit spread and it is expressed in annual basis points as a percentage of the notional amount (face value). This credit spread is usually paid quarterly up until the end of the contract.

1 They founded the company KMV in 1989, which was sold to Moody’s in 2002 and now receives the name of Moody’s Analytics.
Since this is an empirical paper, I want and expect to show that there is one method that may lead me to attain better results than by the use of the traditional KMV model. The other method to compare is going to be the Nominal Debt approach, which relies fundamentally on the fact that default point of a given company is equal to its total debt.

The overall purpose of this paper is then to analyze the determination of distances to default when we only have access to information from equity capital markets and see if there are better methods than KMV’s to estimate distances to default. Studies like Jessen and Lando (2013) support the idea that distances to default are a robust measure for ranking firms according to their respective default risk and so I presumed this assumption. The comparison between the two methods and CDSs, should allow me to determine which of the two is better from a credit risk perspective. The main hypothesis of the paper is defined as follows:

**Hypothesis:** “*The nominal debt approach should give us more accurate estimates than the Moody’s KMV approach to measure the distance to default of a given company*”. 

This hypothesis will be finally verified by the use of regressions that would give me sufficient information to either reject the hypothesis or fail to reject the hypothesis. By the comparison of their respective coefficients of determination, which is a value that gives us clue on how well fitted is the data with model’s variables, we will see evidence committed to an outcome relevant for the subject matter.
2. METHODOLOGY

2.1. KMV Overview

Our starting point of this entire project is how KMV does its calculations to estimate the distance to default of a company. According to Crosbie and Bohn (2003), the formula that compounds all the terms needed to estimate the distance to default is:

\[
[\text{Distance to Default}] = \frac{[\text{Market Value of Assets}] - [\text{Default Point}]}{[\text{Market Value of Assets}] [\text{Asset Volatility}]}
\]

The three parameters needed to estimate Distance-to-Default (DTD) are: the default point, the market value of firm’s assets and the firm’s asset volatility. Once the three elements needed to obtain the distance to default value are estimated, it is then possible to get distance to default for every company, for every day. The first thing is to gather all the data needed to composite these three values required and then I will compare afterwards the KMV distances to default with the nominal debt estimations.

2.2. Data collection

In order to gather all the times-series data needed to proceed with the analysis, I used the Thomson Reuters Datastream, which is an application with lots of historical data related to the financial markets and the macroeconomic environment that permits to get very useful qualitative and quantitative information. I specially used this tool to collect the following financial data: closing adjusted price of equities, market capitalization, total liabilities, and long-term liabilities.

Before proceeding with the data collection, I had to decide the number of years that I wanted to analyze in the paper. I wanted to get data previous to the Lehman Brothers investment bank collapse on September 15, 2008, to be able to draw conclusions from the pre-crisis – post-crisis scenes. Therefore, I chose to collect financial data from 01/01/2008 to 31/12/2013, which means 1.566 business days (stock exchanges opened).
After that I knew how many days I was going to work on, I started to get the information necessary for the analysis. For all the companies to analyze, I got the share price, market capitalization, total liabilities, long-term debt and current liabilities for the 1.566 days.

I chose to study the distance to default on the Euro Stoxx 50’s components because I wanted to obtain relevant conclusions from this research about the European economy after having faced the financial crisis. Thus, during the interval period, there have been many changes in the index and a high volatility in some stages of the timeline driven by a speculation tendency.

This index includes the top 50 blue-chip companies from the Eurozone that are leaders in their sectors and respective countries. It was created the 26th of February, 1998 and it is managed by Stoxx Limited, a joint venture between Deutsche Börse AG and the SIX Group AG (the German and Swiss stock exchanges respectively).

I had to discard 8 of the 50 components due to high leverage ratios that can cause misunderstandings on the results. Therefore I did not include the banks such as BBVA, BNP Paribas, Santander, Deutsche Bank, Société Générale, ING, Intesa Sanpaolo and Unicredit. Crosbie and Bohn (2003) have also arguments for excluding these kind of companies and they explained it as “The credit risk of financial institutions is notoriously difficult to assess. Financial institutions are typically very opaque and thus judging the quality of their assets and determining the extent of their liabilities is almost always very difficult.” Thus, I decided not to include them for the analysis and work with 42 non-financial companies from the index.

Since I wanted to compare the behaviour of the two approaches for measuring DTDs with Credit Default Swaps, I also downloaded CDSs from the database for all the companies. I used 5-year senior unsecured CDSs since according to Coudert and Gex (2010), five years tend to be the most typical maturity and thus, the ones that provide the highest possible liquidity in the market.
Once these first steps were done, I was ready to start to calculate and interpret the fundamentals of the study which I would use later on to perform on the one hand, the MKV’s approach on measuring the distance to default, and on the other hand the Nominal Debt approach.

2.3. Calculating the Default Point

Since I am about to compare the two methods, it is necessary to first define the two different ways in which the default point is calculated. After that the two different sets of values are computed, it is required to perform some adjustments on the data like linear interpolations\(^2\).

2.3.1. KMV’s Default Point

Default Point is defined as the minimum value of company’s assets in which default happens and according to Moody’s-KMV it is expressed as:

\[
DP_t = STL_t + 0.5 \cdot LTL_t
\]

Where (DP) stands for “default point”, (STL) stands for “short term liabilities” and (LTL) represents “long term liabilities”.

2.3.2. Nominal Value Default Point

While according to the nominal debt approach it is simplified as the inclusion of the whole debt as a measure to account the default point.

\[
DP_t = TL_t = P_t
\]

Where (TL) stands for “total liabilities”.

\(^2\) Clarified in the Appendix.
2.4. Calculating Market Value of Assets

According to Crosbie and Bohn (2003), it is assumed by the model that the market value of assets cannot be extracted with exact precision and for this reason I use a proxy to estimate it. I added the market capitalization plus the interpolated data collected from the previous step for every day of the data set.

\[ V_t = MC_t + P_t \]

Where (V) stands for the proxy of “market value of assets”, (MC) represents “market capitalization” and (P) is the value calculated in 2.1.2.

2.5. Annualizing Assets Volatility & Standard Deviation (Assets Risk)

To proceed with the calculation of the distance to default, I had to estimate the daily assets return in such a way that I can after transform it into an annualized value (which will be fixed):

\[ r_t = \ln \left( \frac{V_t}{V_{t-1}} \right) \]

Where (r) is the daily return and (V) represent the values estimated in 2.3. Once I had the 1.565 possible values, I calculated the standard deviation of the data set.

\[ \sigma_{daily} = \sqrt{\frac{1}{T-1} \sum_{i=1}^{t} (V_t - \bar{V})^2} \]

Then (\(\sigma\)) stands for the involved assets volatility per day, (V) is firm’s asset value estimated in 2.3. and (\(\bar{V}\)) is the mean of the total values estimated (one each different for each company). Then I was able to transform the daily value into a unique annualized assets volatility value.

\[ \sigma_{annual} = \sigma_{daily} \cdot \sqrt{250} \]
I multiplied the daily standard deviation by the square root of 250 since the stock exchanges normally are opened 250 days throughout the year.

2.6. Calculation of the Distance to Default

Now I was ready to proceed with the proxy estimation of the DTD for all the companies. This is a continuous value, not fixed. So for every day of the data set I had one different DTD value for the two approaches.

2.6.1. Moody’s-KMV Approach

These are the components used on the Moody’s KMV model:

\[
DTD (KMV)_t = \frac{(V_t - DP_t)}{(V_t \cdot \sigma_{annual})}
\]

Where (V) stands for the values estimated in 2.3., (DP) stands for the values calculated in 2.2.1., and (\(\sigma\)) represents the value obtained in 2.4. for a given company.

2.6.2. Nominal Debt Approach

In this approach, instead of using (DP) from 2.2.1., I use the interpolated nominal debt values (total liabilities) to estimate the distance to default:

\[
DTD (P)_t = \frac{(V_t - P_t)}{(V_t \cdot \sigma_{annual})}
\]

Where (V) corresponds for the values obtained in 2.3., (P) is the default point calculated in the nominal debt way (2.2.2.) and (\(\sigma\)) is the value calculated in 2.4.

From the two equations, it could be concluded that the only difference between the two approaches for estimating distances to default is the way they include their respective default point.
2.7. Getting the averages

To proceed with the analysis I calculated the averages of the two different values I looked for on the previous steps, the DTD P and the DTD KMV because there are very large amounts of data. In order to simplify the calculations, I considered the average a good measure for central tendency since I assume a normal distribution.

Once I calculate all the averages for all the companies, I will then have one “DTD P” value and one “DTD KMV” value for every company in the analysis (42). The same process for calculating CDSs averages, getting 42 values.

2.7.1. Performing correlations

For both types of methods estimated I performed a table of correlations that are described in greater detail in section 4.4. The basis is to execute a correlation coefficient between the two different tables obtained and then to draw conclusions on how each variable is able to explain the result of another.

This table was performed with the CORREL function from Excel.

2.8. Regressions

Finally, when all the data transformation was done, I was ready to go ahead with the last step on the paper that would allow me to either reject the hypothesis or fail to reject the hypothesis set at the very start of this project.

In order to do so, with the help of the software EViews, I completed two different kinds of linear regressions that were estimated by the use of the ordinary least squares (OLS) method.
2.8.1. Cross-sectional regression

Since the objective is to see which of the two approaches gives us a better estimation of credit risk (CDS), I will use the averages of CDS that were calculated in the previous step as dependent variables while the DTD P and DTD KMV would be interpreted as independent variables (each, respectively).

Seven regressions were done for the Nominal Debt approach and seven more for KMV since there are data for six years while the last term stands for the full average mean containing the 1,566 days between 01/01/2008 and 31/12/2013. The two linear equations estimated are as follows:

\[
\text{MEAN CDS}_{c,\tau} = \alpha + \beta \cdot \text{MEAN DTD P}_{c,\tau}
\]

\[
\text{MEAN CDS}_{c,\tau} = \alpha + \beta \cdot \text{MEAN DTD KMV}_{c,\tau}
\]


The software would allow me to illustrate relevant details from the regression like the \(R^2\), Adjusted – \(R^2\), Durbin-Watson statistic, and others, that I would use later on to examine the results obtained. The number of observations included in these regressions range between 30 and 37 because of the lack of some data that should contain CDSs spreads that were unavailable.

2.8.2. Time-series regression

Like on the other type of regression, I use again the CDS averages as dependent variables while the averages of P and KMV correspond to independent variables in the analysis. In this case, the number of observations taken into account for the regressions was much higher (1,566).
To verify the presence of unit roots (non-stationarity), I used the unit roots test and in parallel I performed another regression in which I used differences between variables to compare the results. These are the estimated equations:

\[ d(MEAN CDS_t) = \alpha + \beta \cdot d(MEAN DTD P_t) \]

\[ d(MEAN CDS_t) = \alpha + \beta \cdot d(MEAN DTD KMV_t) \]

Because of the possibility of a presence of persistence in the data, I decided to include a lag (1) of the dependent variable as an independent variable for both types of regressions from the above combined with their respective differences as shown below:

\[ d(MEAN CDS_t) = \alpha + \beta \cdot d(MEAN DTD P_t) + \gamma \cdot d(MEAN DTD P_{t-1}) + \delta \cdot d(MEAN CDS_{t-1}) \]

\[ d(MEAN CDS_t) = \alpha + \beta \cdot d(MEAN DTD KMV_t) + \gamma \cdot d(MEAN DTD KMV_{t-1}) + \delta \cdot d(MEAN CDS_{t-1}) \]
3. DATA

3.1. Data selection

In the table 3.1. we can see all the companies that were included in the analysis. There are a total of 42 companies including data about their respective supersector and country.

<table>
<thead>
<tr>
<th>Components</th>
<th>Supersector</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 AIR LIQUIDE</td>
<td>Chemicals</td>
<td>FR</td>
</tr>
<tr>
<td>2 AIRBUS GROUP NV</td>
<td>Industrial Goods &amp; Services</td>
<td>FR</td>
</tr>
<tr>
<td>3 ALLIANZ</td>
<td>Insurance</td>
<td>DE</td>
</tr>
<tr>
<td>4 ANHEUSER-BUSCH INBEV</td>
<td>Food &amp; Beverages</td>
<td>BE</td>
</tr>
<tr>
<td>5 ASML HLDG</td>
<td>Technology</td>
<td>NL</td>
</tr>
<tr>
<td>6 ASSICURAZIONI GENERALI</td>
<td>Insurance</td>
<td>IT</td>
</tr>
<tr>
<td>7 AXA</td>
<td>Insurance</td>
<td>FR</td>
</tr>
<tr>
<td>8 BASF</td>
<td>Chemicals</td>
<td>DE</td>
</tr>
<tr>
<td>9 BAYER</td>
<td>Chemicals</td>
<td>DE</td>
</tr>
<tr>
<td>10 BMW</td>
<td>Automobiles &amp; Parts</td>
<td>DE</td>
</tr>
<tr>
<td>11 CARREFOUR</td>
<td>Retail</td>
<td>FR</td>
</tr>
<tr>
<td>12 DAIMLER</td>
<td>Automobiles &amp; Parts</td>
<td>DE</td>
</tr>
<tr>
<td>13 DANONE</td>
<td>Food &amp; Beverages</td>
<td>FR</td>
</tr>
<tr>
<td>14 DEUTSCHE POST</td>
<td>Industrial Goods &amp; Services</td>
<td>DE</td>
</tr>
<tr>
<td>15 DEUTSCHE TELEKOM</td>
<td>Telecommunications</td>
<td>DE</td>
</tr>
<tr>
<td>16 E.ON</td>
<td>Utilities</td>
<td>DE</td>
</tr>
<tr>
<td>17 ENEL</td>
<td>Utilities</td>
<td>IT</td>
</tr>
<tr>
<td>18 ENI</td>
<td>Oil &amp; Gas</td>
<td>IT</td>
</tr>
<tr>
<td>19 ESSION INTERNATIONAL</td>
<td>Healthcare</td>
<td>FR</td>
</tr>
<tr>
<td>20 GDF SUEZ</td>
<td>Utilities</td>
<td>FR</td>
</tr>
<tr>
<td>21 IBERDROLA</td>
<td>Utilities</td>
<td>ES</td>
</tr>
<tr>
<td>22 INDITEX</td>
<td>Retail</td>
<td>ES</td>
</tr>
<tr>
<td>23 L’OREAL</td>
<td>Personal &amp; Household Goods</td>
<td>FR</td>
</tr>
<tr>
<td>24 LVMH MOET HENNESSY</td>
<td>Personal &amp; Household Goods</td>
<td>FR</td>
</tr>
<tr>
<td>25 MUENCHENER RUECK</td>
<td>Insurance</td>
<td>DE</td>
</tr>
<tr>
<td>26 NOKIA</td>
<td>Technology</td>
<td>FI</td>
</tr>
<tr>
<td>27 ORANGE</td>
<td>Telecommunications</td>
<td>FR</td>
</tr>
<tr>
<td>28 PHILIPS</td>
<td>Industrial Goods &amp; Services</td>
<td>NL</td>
</tr>
<tr>
<td>29 REPSOL</td>
<td>Oil &amp; Gas</td>
<td>ES</td>
</tr>
<tr>
<td>30 RWE</td>
<td>Utilities</td>
<td>DE</td>
</tr>
<tr>
<td>31 SAINT GOBAIN</td>
<td>Construction &amp; Materials</td>
<td>FR</td>
</tr>
<tr>
<td>32 SANOFI</td>
<td>Healthcare</td>
<td>FR</td>
</tr>
<tr>
<td>33 SAP</td>
<td>Technology</td>
<td>DE</td>
</tr>
<tr>
<td>34 SCHNEIDER ELECTRIC</td>
<td>Industrial Goods &amp; Services</td>
<td>FR</td>
</tr>
<tr>
<td>35 SIEMENS</td>
<td>Industrial Goods &amp; Services</td>
<td>DE</td>
</tr>
</tbody>
</table>
3.2. Statistics and relevant information to subject matter

In order to get know a bit more on the subject in which I worked for during all the process of this project, I decided to include some relevant statistics that would allow me to get better insights from the data and thus, know more useful stats about the subject matter. All these information could be relevant when the time to conclude arrives.

Table 3.2. Top 10 Companies in the EURO STOXX 50 by Market Capitalization (29/04/2015)

<table>
<thead>
<tr>
<th>Company</th>
<th>Market Capitalization Million Euros</th>
<th>Capitalization gained/lost in 2015 Million Euros</th>
<th>Year's variation %</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANHEUSER-BUSCH INBEV</td>
<td>182,696</td>
<td>31,746</td>
<td>21.03</td>
</tr>
<tr>
<td>SANOFI</td>
<td>126,969</td>
<td>26,569</td>
<td>20.7</td>
</tr>
<tr>
<td>UNILEVER</td>
<td>124,156</td>
<td>24,756</td>
<td>26.53</td>
</tr>
<tr>
<td>TOTAL</td>
<td>114,504</td>
<td>13,114</td>
<td>12.9</td>
</tr>
<tr>
<td>BAYER</td>
<td>111,886</td>
<td>18,436</td>
<td>16.04</td>
</tr>
<tr>
<td>VOLKSWAGEN</td>
<td>110,320</td>
<td>23,82</td>
<td>26.21</td>
</tr>
<tr>
<td>L'OREAL</td>
<td>98,636</td>
<td>19,457</td>
<td>26.17</td>
</tr>
<tr>
<td>SANTANDER</td>
<td>93,657</td>
<td>-5,617</td>
<td>-4.79</td>
</tr>
<tr>
<td>DAIMLER</td>
<td>93,247</td>
<td>20,616</td>
<td>26.37</td>
</tr>
<tr>
<td>INDITEX</td>
<td>91,255</td>
<td>17,375</td>
<td>23.52</td>
</tr>
</tbody>
</table>

Source: Bloomberg

According to the table 3.2., it can be appreciated that there have been some big positive changes for the top ten companies, with the exception of Santander. The biggest company -in market capitalization terms- is the Belgian multinational that produces beverages and beer. It is followed by the French healthcare giant, which is the one that enjoyed the highest growth this year, and in third place we find Unilever, which also is having a good year. Santander is the only bank in the top ten and therefore the biggest bank in the Eurozone.
As of 7 May 2015, the index (SX5E) is valued at 3.556,21 points and the total market capitalization consists of 2.738.046 million Euros. The index valuation is calculated with the Laspeyres formula\(^3\), and it is represented with the following variables:

\[
Index_t = \sum_{i=1}^{n} (p_{it} \cdot s_{it} \cdot ff_{it} \cdot cf_{it} \cdot x_{it}) = \frac{M_t}{D_t}
\]

Where:

\(p_{it}\) = Price of company (i) at time (t)

\(s_{it}\) = Number of shares of company (i) at time (t)

\(ff_{it}\) = Free float factor of company (i) at time (t)

\(cf_{it}\) = Weighted cap factor of company (i) at time (t)

\(x_{it}\) = Exchange rate from local currency into index currency for company (i) at time (t)

\(M_t\) = Free float market capitalization of the index at time (t)

\(D_t\) = Divisor of the index at time (t)

**Figure 3.1. EURO STOXX 50 All-Time Historical Chart**

![EURO STOXX 50 All-Time Historical Chart](image)

**Source: The Wall Street Journal**

---

\(^3\) Eurostat defines the Laspeyres price index as “an index formula used in price statistics for measuring the price development of the basket of goods and services consumed in the base period. The question it answers is how much a basket that consumers bought in the base period would cost in the current period. It is defined as a fixed-weight, or fixed-basket, index that uses the basket of goods and services and their weights from the base period”.
The historical index value since 1999 faced two big recessions. The first happened right after the dot-com bubble in which it reached the highest historical value, while the second was after the Lehman Brothers collapse due to the subprime crisis that led the index to the lowest value in its history.

The period that I chose to study includes big ups and downs, but unfortunately, it was not possible to include the last bullish cycle from 2013 to today because of a lack of financial data needed for the study.

Given the big set of data, it is possible to extract valuable information and thus, conclusions from them by organizing it by “supersectors” and nationalities. There are 15 different sectors with representative companies in the index chosen, and in terms of geographical distribution, there are up to 7 different countries that include companies in the EURO STOXX 50.

The proportion of countries represented is low since there are 19 countries constituting the Eurozone, so the companies are quite concentrated in the big economies of Western Europe. The vast majority of the companies in the index are from France and Germany, which are also the top two industrial economies in the Eurozone. The two are able to include more than half of the companies.

Figure 3.2. EURO STOXX 50 Components distributed by country
The top three industries in the index are the Industrial Goods & Services (12%), Utilities (12%) and Insurance (10%). Overall the components together form a quite balanced index. In the empirical research section, it is going to be seen that there are big differences on how every sector behaves depending on different variables.

Figure 3.3. EURO STOXX 50 Components distributed by supersector

Although the banking sector has not been included in the analysis, it is the one with the highest representation in the index with up to 8 banks.
4. EMPIRICAL RESEARCH

4.1. Getting the averages

It was necessary to calculate the averages of the two different values I looked for on the previous steps (DTD P and the DTD KMV) because there are very large amounts of data and in order to simplify the calculations, I considered the average a good measure for central tendency since I assume a normal distribution.

Once I calculate all the averages for all the companies, I will then have one “DTD P” value and one “DTD KMV” value for every company in the analysis (42 for each method). The same process for calculating CDSs averages, getting 42 values.

4.1.1. Average per year (Cross-sectional regression)

Again I repeated the same process, but only considering the values that are included on the range of a year. For instance, for the year 2008 I have 262 values per company.

I got the averages from those values and I repeated this step 7 times since I considered 6 full years to analyse in the paper plus a row that contemplates the full average of the 6 years (2008-2013). Here is an example of the performance of what is been explained for all the years on the company Daimler:

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean DTD P</th>
<th>Mean DTD KMV</th>
<th>CDS Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>1.77</td>
<td>2.89</td>
<td>165.38</td>
</tr>
<tr>
<td>2009</td>
<td>1.02</td>
<td>2.88</td>
<td>194.52</td>
</tr>
<tr>
<td>2010</td>
<td>1.78</td>
<td>2.88</td>
<td>103.83</td>
</tr>
</tbody>
</table>

Table 4.1. P, KMV and CDS averages per year example (Daimler)
<table>
<thead>
<tr>
<th>Year</th>
<th>Mean DTD P</th>
<th>Mean DTD KMV</th>
<th>CDS Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>1.84</td>
<td>2.87</td>
<td>126.05</td>
</tr>
<tr>
<td>2012</td>
<td>1.54</td>
<td>2.86</td>
<td>128.41</td>
</tr>
<tr>
<td>2013</td>
<td>1.99</td>
<td>2.86</td>
<td>81.31</td>
</tr>
<tr>
<td>2008 – 2013</td>
<td>1.66</td>
<td>2.79</td>
<td>133.33</td>
</tr>
</tbody>
</table>

Given that I want to compare two approaches, this operation was done twice, one for the P values and the second for the KMV values as seen on the table 4.1. With this whole table I will obtain the data necessary to compute the cross-sectional regression (2.8.1).

As we can see, even with the presence of a very intense financial crisis, the KMV values continued to be too stable for what I expected. This is a clue that tells me that maybe nominal debt approach may explain distances to default in a better way.

### 4.1.2. Average per day (Time-series regression)

In this step, instead of having one value per company for the three different variables, I have one value for each variable, per every single day of the data set in which I take into account all the companies.

Thus, I classified a table separating the two approaches to measure distances to default and another for the CDSs averages obtained for the 1,566 days. In this table it is then shown the combined averages of all the companies per day that I will use later on to do the time-series regressions.
4.2. Leverage ratio

In this project, I used the leverage ratio to extract qualitative information. This ratio is a good tool to determine the likelihood of a company to pay its debt considering that some industries operate with higher leverage ratios than others.

*Investopedia* explains this concept as “Uncontrolled debt levels can lead to credit downgrades or worse. On the other hand, too few debts can also raise questions. If a company's operations can generate a higher rate of return than the interest rate on its loans, then the debt is helping to fuel growth in profits. A reluctance or inability to borrow may be a sign that operating margins are simply too tight.”

Therefore, not always a low leverage ratio is good or a high leverage ratio is bad. It really depends on many variables like the company’s ability to generate cash-flows, the characteristics of the industry, the market’s behaviour, regulatory laws, and others.

4.3. Comparing DTD P with DTD KMV averages

After filling all the required data for the purpose of this paper, it is possible to transform it into qualitative information by sorting the results into the three components by which I divided the table below.

| Table 4.2. DTD mean comparison (P vs. KMV) and leverage mean by company |
|---------------------------------|-----------------|-----------------|-----------------|
| Company                        | MEAN DTD P      | MEAN DTD KMV    | MEAN LEVERAGE   |
| AIR LIQUIDE                    | 3.598           | 4.017           | 0.359           |
| AIRBUS GROUP NV                | 2.542           | 2.737           | 0.801           |
| ALLIANZ                        | 2.466           | 3.112           | 0.948           |
| ANHEUSER-BUSCH INBEV           | 2.503           | 3.385           | 0.503           |
| ASML HLDG                      | 2.754           | 2.730           | 0.210           |
| ASSICURAZIONI GENERALI        | 1.418           | 3.379           | 0.941           |
| AXA                            | 1.742           | 2.015           | 0.951           |
| BASF                           | 3.008           | 3.330           | 0.427           |
| BAYER                          | 3.233           | 3.495           | 0.420           |
| BMW                            | 1.250           | 2.997           | 0.760           |
| CARREFOUR                      | 2.388           | 2.969           | 0.660           |
| DAIMLER                        | 1.657           | 2.791           | 0.722           |
| DANONE                         | 3.341           | 3.898           | 0.392           |
| DEUTSCHE POST                  | 3.329           | 3.627           | 0.850           |
| DEUTSCHE TELEKOM               | 2.032           | 3.634           | 0.645           |
| E.ON                           | 2.245           | 2.936           | 0.695           |
| ENEL                           | 0.695           | 3.033           | 0.760           |
According to the table 4.2., the arithmetic means of the two approaches differ by 0.631 points, being the KMV’s a higher mean value.

The company that shows the lowest distance to default at both levels is the automaker Volkswagen, which also has a leverage value above the total average. On the other side, the company with the highest distances to default combined with a relatively low leverage ratio is Essilor International, a French company that produces optical goods for the medical equipment sector (Healthcare).

4.3.1. Supersector comparison and analysis

While the leverage mean is 0.556, we can see that there is a high variation of this value depending on the company. For instance, INDITEX is the company with the lowest leverage ratio (0.092), while AXA, Allianz and Assicurazioni Generali (the three in the insurance sector) have a value of above 0.9. This may be explained because of the industry where they operate, the regulatory laws, the country’s culture, and others. For this, I decided to perform an analysis by industry so that I could compare both the distance to default and leverage ratios and draw conclusions.
Table 4.3. DTD mean comparison (P vs. KMV) and leverage mean by supersector

<table>
<thead>
<tr>
<th>SECTOR</th>
<th>MEAN DTD P</th>
<th>MEAN DTD KMV</th>
<th>MEAN LEVERAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemicals</td>
<td>3,280</td>
<td>3,614</td>
<td>0,402</td>
</tr>
<tr>
<td>Industrial Goods &amp; Services</td>
<td>2,778</td>
<td>3,042</td>
<td>0,599</td>
</tr>
<tr>
<td>Insurance</td>
<td>2,178</td>
<td>2,994</td>
<td>0,935</td>
</tr>
<tr>
<td>Technology</td>
<td>2,808</td>
<td>2,828</td>
<td>0,271</td>
</tr>
<tr>
<td>Automobiles &amp; Parts</td>
<td>1,082</td>
<td>2,178</td>
<td>0,762</td>
</tr>
<tr>
<td>Retail</td>
<td>2,843</td>
<td>3,109</td>
<td>0,376</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>2,199</td>
<td>3,675</td>
<td>0,598</td>
</tr>
<tr>
<td>Oil &amp; Gas</td>
<td>2,740</td>
<td>3,200</td>
<td>0,503</td>
</tr>
<tr>
<td>Utilities</td>
<td>1,717</td>
<td>2,827</td>
<td>0,702</td>
</tr>
<tr>
<td>Personal &amp; Household Goods</td>
<td>3,609</td>
<td>3,820</td>
<td>0,333</td>
</tr>
<tr>
<td>Construction &amp; Materials</td>
<td>1,780</td>
<td>2,691</td>
<td>0,655</td>
</tr>
<tr>
<td>Healthcare</td>
<td>3,975</td>
<td>4,001</td>
<td>0,238</td>
</tr>
<tr>
<td>Real Estate</td>
<td>2,344</td>
<td>3,463</td>
<td>0,505</td>
</tr>
<tr>
<td>Media</td>
<td>2,769</td>
<td>3,378</td>
<td>0,553</td>
</tr>
<tr>
<td>Food &amp; Beverages</td>
<td>2,922</td>
<td>3,642</td>
<td>0,447</td>
</tr>
</tbody>
</table>

Figure 4.1. P & KMV average DTD by supersector

On the one hand, it is noticeable that the Healthcare sector is the one that gives the highest distance to default for both approaches while on the other hand, is the Automobiles & Parts industry the one that its distance to default is closest to zero in both approaches, being the nominal debt approach significantly lower than the KMV’s.
According to the figure 4.2. and, as expected, the insurance sector is the one with the highest leverage ratios in the data set given its highest arithmetic mean. At the other extreme, I find that healthcare followed by technology, are the industries that operate with the lowest leverage ratios (0.238 and 0.271 respectively).

So there may probably be a significant correlation between low leverage ratios and high distances to default.

4.3.2. *Country comparison and analysis*

<table>
<thead>
<tr>
<th>COUNTRY</th>
<th>MEAN DTD P</th>
<th>MEAN DTD KMV</th>
<th>MEAN LEVERAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td>2,725</td>
<td>3,219</td>
<td>0,508</td>
</tr>
<tr>
<td>Germany</td>
<td>2,423</td>
<td>3,099</td>
<td>0,659</td>
</tr>
<tr>
<td>Belgium</td>
<td>2,503</td>
<td>3,385</td>
<td>0,503</td>
</tr>
<tr>
<td>Netherlands</td>
<td>3,169</td>
<td>3,398</td>
<td>0,389</td>
</tr>
<tr>
<td>Italy</td>
<td>1,646</td>
<td>3,210</td>
<td>0,735</td>
</tr>
<tr>
<td>Spain</td>
<td>2,433</td>
<td>3,182</td>
<td>0,455</td>
</tr>
<tr>
<td>Finland</td>
<td>1,991</td>
<td>2,050</td>
<td>0,454</td>
</tr>
</tbody>
</table>
Although the data set is not equally divided by the same number of companies by country, I wanted to reflect in the graph above an evaluation on how the companies behave depending on the country where they have their headquarters. Apparently, the Finnish, Spanish and Italian companies are the ones that are closer to zero. This may have sense according to the economical difficulties that Spain and Italy have faced during the recession.
Again, here I can see evidence that the country closest to zero on the distance to default’s mean is also the one with the highest leverage ratio. This time, Italy (0.735) is followed closely by Germany (0.659). The Netherlands is the country with the lowest leverage ratio (0.389).

4.3.3. Comparing DTD P and DTD KMV with EURO STOXX 50 index

Figure 4.5. All companies’ average DTD: P & KMV throughout the chosen timeline

Figure 4.6. EURO STOXX 50 value throughout the chosen timeline

Source: The Wall Street Journal
It is possible to see that there are some big similarities between figures 4.5. and 4.6. due to a consistent positive correlation between the index value and the estimated distances to default. This is normal and expected since the stock prices are included in the calculation for both approaches.

4.4. Correlations

I chose to compare the average distances to default from both approaches (calculated in 2.3.) against the CDSs averages of every year.

<table>
<thead>
<tr>
<th></th>
<th>CORRELATION</th>
<th></th>
<th>CORRELATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTDKMV08</td>
<td>-0.515</td>
<td>DTDP08</td>
<td>-0.724</td>
</tr>
<tr>
<td>DTDKMV09</td>
<td>-0.507</td>
<td>DTDP09</td>
<td>-0.714</td>
</tr>
<tr>
<td>DTDKMV10</td>
<td>-0.130</td>
<td>DTDP10</td>
<td>-0.518</td>
</tr>
<tr>
<td>DTDKMV11</td>
<td>0.096</td>
<td>DTDP11</td>
<td>-0.570</td>
</tr>
<tr>
<td>DTDKMV12</td>
<td>0.162</td>
<td>DTDP12</td>
<td>-0.647</td>
</tr>
<tr>
<td>DTDKMV13</td>
<td>0.250</td>
<td>DTDP13</td>
<td>-0.664</td>
</tr>
<tr>
<td>DTDKMV0813</td>
<td>-0.291</td>
<td>DTDP0813</td>
<td>-0.616</td>
</tr>
</tbody>
</table>

It is visible that better results have been obtained on the nominal debt approach in any of the years compared. The difference is quite significant. Even for the 6-years average the difference is big (0,325).

Other relevant aspects that can be read from the table above is that I found negative correlations for every year on the nominal debt side, while on the KMV side, the results are quite disperse. It was expected to find negative correlations, which means that if distances to default increase, credit default swaps decrease since the price for “insurance” should go down.

I consider that these results are very relevant for the direction of the paper and thus, I continue following the set target with forthcoming data interpretation via econometric modelling.
5. RESULTS OBTAINED

5.1. Results: Regressions

With the completion of the cross-sectional regressions, I found enough evidence to support the initial hypothesis given the fact that the obtained coefficients of determination are sensibly higher in the regressions that included the nominal debt approach instead of those who included the KMV approach, which is the most widespread method used in the risk management industry.

5.1.1. Results: Cross-sectional regression

\[
MEAN\ CDS_{c,t} = \alpha + \beta \cdot MEAN\ DTD\ P_{c,t}
\]

\[
MEAN\ CDS_{c,t} = \alpha + \beta \cdot MEAN\ DTD\ KMV_{c,t}
\]

Table 5.1. R-squared results (Cross-sectional regressions)

<table>
<thead>
<tr>
<th></th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>08 - 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>0.523</td>
<td>0.510</td>
<td>0.268</td>
<td>0.324</td>
<td>0.419</td>
<td>0.440</td>
<td>0.379</td>
</tr>
<tr>
<td>KMV</td>
<td>0.265</td>
<td>0.257</td>
<td>0.016</td>
<td>0.00</td>
<td>0.026</td>
<td>0.062</td>
<td>0.084</td>
</tr>
</tbody>
</table>

From the table above, we can see classified the coefficients of determination by year and by approach. Overall, the results from regressions showed that CDSs seem to be better explained than by the nominal debt measure than by KMV’s since at all levels the values obtained are higher.

Table 5.2. Overall results from cross-sectional regressions
In 2008, the year in which Lehman Brothers collapsed, the distances of default obtained were the highest for all the timeline and so were the coefficients of determination. In 2009 these coefficients mostly maintained the previous values, but in 2010 I observed that it is the year in which distances to default plummeted. This could be explained by the fact that it was a clear change of tendency combined with a tremendous uncertainty in the financial markets.

From 2010 to 2013, the coefficient of determination values grew slowly up until a close level to 2008 in the case of the nominal debt approach, but not in the KMV’s since I notice that the growth was even slower and still very far away from the initial level in the analysis.

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>KMV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>α</td>
<td>β</td>
</tr>
<tr>
<td>2008</td>
<td>201.99</td>
<td>-36.8</td>
</tr>
<tr>
<td>2009</td>
<td>197.15</td>
<td>-41.66</td>
</tr>
<tr>
<td>2010</td>
<td>134.62</td>
<td>-19.6</td>
</tr>
<tr>
<td>2011</td>
<td>206.03</td>
<td>-36.92</td>
</tr>
<tr>
<td>2012</td>
<td>264.05</td>
<td>-55.28</td>
</tr>
<tr>
<td>2013</td>
<td>170.62</td>
<td>-29.76</td>
</tr>
<tr>
<td>08 - 13</td>
<td>191.1</td>
<td>-33.74</td>
</tr>
</tbody>
</table>

Apart from R² values, in the table 5.3 we can see more information about the regressions done for each year, such as the alpha and beta coefficients as well as the Durbin-Watson statistics obtained. D-W statistic is a measure for detecting the possible presence of autocorrelation and when the value obtained is around 2, it means that there is evidence to discard the presence of autocorrelation and thus, the regressions are well fitted. As we can observe, in most cases the values obtained are relatively close to 2, so I assume that the estimations are good enough for the comprehensive purpose.

So, overall these results from cross-sectional regressions show enough evidence to be unable to reject the hypothesis.
5.1.2. Results: Time-series regression

\[ MEAN \, CDS_t = \alpha + \beta \cdot MEAN \, DTD \, P_t \]

\[ MEAN \, CDS_t = \alpha + \beta \cdot MEAN \, DTD \, KMV_t \]

In this case, the regressions were done with a total of 1,566 observations since it was not necessary to perform means for each year. Only one independent variable was included, which includes the averages obtained for every day of the dataset from the 42 companies.

Table 5.4. Time-series Regression #1

<table>
<thead>
<tr>
<th></th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( R^2 )</th>
<th>DW</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>389.24</td>
<td>-110.3</td>
<td>0.49</td>
<td>0.01</td>
</tr>
<tr>
<td>KMV</td>
<td>460.99</td>
<td>-110.99</td>
<td>0.42</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Again, it is visible in table 5.4., that I obtained better results for the nominal debt approach due to a higher R-squared value. The counterpoint in this case is the fact that the Durbin-Watson’s statistic values obtained are too far away from 2, which make me think about the possibility that I performed spurious regressions and thus, there is a presence of unit roots.

According to Granger and Newbold (1973), spurious regressions are a common circumstance in time-series regressions that have low values for the DW statistic although the R-squared values are high enough. When this situation happens, there is a need for doing some adjustments and in the paper they recommend to either include lagged variables or to take first differences of the variables involved in the regressions.

For this reason, I decided to first examine if there actually is a presence of unit roots. I performed the Augmented Dickey-Fuller\(^4\) test and these are the results:

\(^4\) Sets up an autoregressive model AR(1) where it tests whether there is a presence of unit roots or not.
Table 5.5.1. Augmented Dickey-Fuller Test (Nominal Debt)

Null Hypothesis: MEANDTDP has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=23)

<table>
<thead>
<tr>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller test statistic</td>
<td>-2.658602</td>
</tr>
</tbody>
</table>


Table 5.5.2. Augmented Dickey-Fuller Test (KMV)

Null Hypothesis: MEANDTDMV has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=23)

<table>
<thead>
<tr>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller test statistic</td>
<td>-2.241963</td>
</tr>
</tbody>
</table>


Table 5.5.3. Augmented Dickey-Fuller Test (CDS)

Null Hypothesis: MEANCDS has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on SIC, maxlag=23)

<table>
<thead>
<tr>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller test statistic</td>
<td>-2.522466</td>
</tr>
</tbody>
</table>


Since the p-values obtained for the three tests done are higher than 0.05, I cannot reject the null hypothesis and so, there may be a presence of unit roots. In order to solve this issue, I decided to perform two new regressions that include differences of the variables and compare the results as suggested by Granger and Newbold (1973).

\[ d(MEAN\ CDS_t) = \alpha + \beta \cdot d(MEAN\ DTD\ P_t) \]

\[ d(MEAN\ CDS_t) = \alpha + \beta \cdot d(MEAN\ DTD\ KMV_t) \]
Table 5.6. Time-series Regression #2

<table>
<thead>
<tr>
<th></th>
<th>α</th>
<th>β</th>
<th>R²</th>
<th>DW</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>0,01</td>
<td>-65,48</td>
<td>0,29</td>
<td>1,68</td>
</tr>
<tr>
<td>KMV</td>
<td>0,01</td>
<td>-74,9</td>
<td>0,28</td>
<td>1,68</td>
</tr>
</tbody>
</table>

This time, I achieved DW values closer to 2 and although the R-squared values turned down, the fit is still better for the nominal debt approach. I decided then to perform new version of regressions from the previous ones by including lagged variables together with differences as suggested by Granger and Newbold (1973) and then compare the results.

\[
d(\text{MEAN CDS}_t) = \alpha + \beta \cdot d(\text{MEAN DTD } P_t) + \gamma \cdot d(\text{MEAN DTD } P_{t-1}) + \delta \cdot d(\text{MEAN CDS}_{t-1})
\]

\[
d(\text{MEAN CDS}_t) = \alpha + \beta \cdot d(\text{MEAN DTD KMV}_t) + \gamma \cdot d(\text{MEAN DTD KMV}_{t-1}) + \delta \cdot d(\text{MEAN CDS}_{t-1})
\]

Table 5.7. Time-series Regression #3

<table>
<thead>
<tr>
<th></th>
<th>α</th>
<th>β</th>
<th>γ</th>
<th>δ</th>
<th>Adjusted R²</th>
<th>DW</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>0</td>
<td>-65,81</td>
<td>-24,79</td>
<td>0,16</td>
<td>0,386</td>
<td>2,03</td>
</tr>
<tr>
<td>KMV</td>
<td>0</td>
<td>-75,21</td>
<td>-28,42</td>
<td>0,16</td>
<td>0,385</td>
<td>2,03</td>
</tr>
</tbody>
</table>

From table 5.7., it can been observed that the DW’s statistic values now are very approximated to two, therefore this is an indication that the estimated model probably has no problems of autocorrelation and again, the observed Adjusted R² values show that the nominal debt method for measuring distance to default is the one who slightly better explains how CDSs behave. In this last regression it was used the *adjusted R-squared* to compare the two equations since I included up to four variables, although there are no big differences between the two types of R-squared versions estimated (0,387 for P and 0,386 for KMV on the “simple” R-squared estimation).

Overall, after having completed the three different sets of time-series regressions and after having observed that at any of them the coefficients of determination are higher for the nominal debt approach (like in the cross-sectional regressions), I could not find any argument that may reject the initial hypothesis.
6. CONCLUSIONS

6.1. Evidence

After having completed all the steps of this paper and after having interpreted the results obtained, I can say that the research done combined with the procedures taken to analyze the subject matter, **the data gives me enough evidence to rely on the successfulness of the alleged hypothesis and thus, being unable to reject it.** The fundamental estimations made me think this way because at any levels of the regressions prosecuted, I found that CDSs are better explained always by the use of the nominal debt approach over the KMV’s.

By comparing the R-squared values between the two methods in all of the regressions done, it is visible that not only the ones in the nominal debt approach are higher, but also they performed more accurately to what really happened in the market, even including periods of a very intense speculation (high volatility) in the financial markets.

The results display that there is no real need to multiply by one half the long term debts in order to calculate the appropriate default points. This argument contrasts what Crosbie and Bohn (2003) suggested in their article. In this same article we can find that they also stated that “we have assumed that the default point is described by the firm’s liabilities and amortization schedule. Of course we know that this is not true. Unfortunately ex ante we are unable to specify the behaviour of the liabilities and thus the uncertainty in the adjustments in the liabilities must be captured elsewhere”.

Crosbie and Bohn (2003) say by their own that it is then hard to define the default point, but they do not really say a reason why they estimated their default points that way instead of doing it like I did on this paper (nominal debt approach), which seems more logical.

Supported with the idea described in the academic paper written by Jessen and Lando (2013) that tells us that “distance-to-default is a robust measure for ranking firms according to their default risk under most violations of the Merton model’s assumptions”, I state that by taking the nominal debt approach instead of KMV’s, even
more robust distances to default can be produced since they are better correlated with CDSs.

From the empirical results, I also deduct that effectively, distances to default behaved differently in the pre-crisis than in the post-crisis scenario. From table 6.1., it is noticeable that CDSs (42-companies average) reached their record high in 2012 (145,49 bp) while the Nominal Debt distances to default projected their minimum value in the timeline in 2009, right after Lehman Brothers investment bank collapsed, as expected.

<table>
<thead>
<tr>
<th>Table 6.1. CDSs vs. DTD (Nominal Debt) Averages Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CDS 08</strong></td>
</tr>
<tr>
<td>104,15</td>
</tr>
<tr>
<td><strong>P 08</strong></td>
</tr>
<tr>
<td>2,72</td>
</tr>
</tbody>
</table>

While CDSs spreads plummeted from 2012 to 2013 (from 145,49 bp to 97,90 bp), distances to default skyrocketed in that same timeframe (from 2,39 to 2,74). This argument fortifies indeed the robustness of the nominal debt approach on how these variables are negatively correlated since the values went to opposite directions in general. This is the moment in which the European economical instability started to slowly cease to be present and some growth from the biggest economies started to happen. The previous years in the timeline are moments of high unstableness and economies in the European continent struggled to grow.

It is a bit surprising and has to be mentioned that in 2013, distances to default surpassed the pre-crisis levels of 2008 in only six years of difference, but it can be explained by the fact that the Lehman Brothers’ default occurred in September and in the 2008’s average there are included also the distances to default from the days previous to the financial collapse (01/01/2008 – 15/09/2008) that were particularly high.

I also noticed that in 2009 the European CDSs spreads maintained quite constant although I expected to see a big downfall of this value since it actually happened for the distance to default estimated in 2009. It took up to two years for the CDSs to reach their minimum value in all the timeline projected (86,54 bp).
We can possibly think that since the values obtained in 2013 for both CDSs and distances to default are similar to 2008’s, we can expect to happen a new financial crisis in 2013, but it is not what actually happened because as I mentioned before, in 2008 it was included the effect of a massive financial shock that is not included in 2013.

Therefore, overall this table projects a bullish economical cycle and it is what happened for example in the U.S. up to today in which the S&P 500, the NASDAQ Composite (both from the U.S.) and the DAX 30 indexes (from Germany) reached their all-time record high in 2015.

Finally, I also encountered that some industries tend to be farther from their default point than others, being the healthcare companies the ones that showed a tendency to be more distant to default than the rest, while the automobiles & parts industry tend to be the closest to their respective distance to default.

6.2. Difficulties and further research

One of the difficulties that I found over time was the fact of having to deal with lots of data. It was necessary to be very organized with all the Excel sheets and to make a good use of the formulas available in order to save time.

Another obstacle that I found was that some data that I intended to download was not available like in the case of some CDSs. Fortunately that was just an exception that did not affect the results obtained at the end because it was possible to deal with the lack of data by doing some statistical adjustments.

For further research, I think that it would be interesting to compare the results obtained in this paper with components from another index, for example I propose the 50 top companies from the S&P 500 in order to be able to draw conclusions about how the American and European markets differ one from the other and observe which of the two suffered the most the financial crisis from a distance-to-default perspective. As we saw in table 6.1, the effect of the Lehman Brothers collapse needed some time to be
reflected on the financial markets in Europe. It would be interesting to compare if the same happened in the U.S. with their blue-chip companies.

I have also curiosity about making some sort of a comparison between public debt and private debt throughout the years 2008-2013. I guess that by comparing them, we should be able to explain if companies’ distances to default have any correlation with their respective countries risk to fail by implying the nominal debt approach.

Since Crosbie and Bohn (2003) specified in their article the fact that there are big difficulties to estimate distances to default on financial companies like banks, I would suggest a study on how to interpret the way these companies differentiate from the others and to put together a new specified model able to estimate their distances to default properly. This is a major cause since there are many random variables and causes that cannot be predicted due to regulatory laws on banks, state intervention that may lead to too big to fail causes and others.

I am also attracted by the idea on how companies decide to use financial leverage to grow faster than their competitors assuming credit risks, so I suggest a study that would be able to explain if companies really have in mind how distant they are from default when taking this kind of decisions.

---

5 Credit Suisse defines the concept as “if the bank incurs losses, shareholders' equity falls. It doesn't come to a crash until the losses incurred are so great that the shareholders' equity is all used up. Insolvency proceedings can take years, and they often have dramatic consequences. In many cases the bank must go out of business. In 2008, the bubble burst, and the banks were in trouble. Because there were fears for the entire financial system, the most important banks - UBS among them - were rescued. It was “too big to fail.”
7. BIBLIOGRAPHY

7.1. Academic papers

   http://www.macs.hw.ac.uk/~mcneil/F79CR/Crosbie_Bohn.pdf


   http://www.bde.es/t/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/06/Fic/dt0639.pdf

   http://www.willisresearchnetwork.com/assets/templates/wrn/files/Distance-to-default_June%202012%5B1%5D.pdf


7.2. Electronic texts


   http://www.investopedia.com/terms/c/creditdefaultswap.asp


7.3. Online databases


7.4. Software


c. Mozilla Firefox 37.0.2, Mozilla Corporation, United States of America, 2015, Available from: https://www.mozilla.org/firefox/new/

APPENDIX

1. Transforming the Default Point: Linear interpolation

I had to adapt the data obtained in order to establish the proper calculation of all the items in the models. To do so, I transformed the non-dynamic financial data such as short/long-term liabilities (1 balance sheet per year) into one balance sheet per day. The mathematical approach for this purpose is called linear interpolation, which allowed me to connect the data obtained from the balance sheets into a continuous growing/un-growing data set by filling the unknown gaps.

As we can see from the graph, the blue line connects with unknown data, the seven known red points.

The formula is expressed as the last value of the column, minus the first value of the column and then all divided by the number of values in the column.

Here is an example for interpolating the Total Liabilities column:

\[
\{2,2,2,3,3,3,5,5,5\} \quad N = 9 \rightarrow (5 - 2) / 9 = 1/3
\]

Once I get this value, I add it for all the numbers of the column consecutively and I get the following column:

\[
\{2, (2 + 1/3), (2 + 1/3 + 1/3), (2 + 1/3 + 1/3 + 1/3), (2 + 1/3 + 1/3 + 1/3 + 1/3), (2 + 1/3 + 1/3 + 1/3 + 1/3 + 1/3), (2 + 1/3 + 1/3 + 1/3 + 1/3 + 1/3 + 1/3), (2 + 1/3 + 1/3 + 1/3 + 1/3 + 1/3 + 1/3 + 1/3), (2 + 1/3 + 1/3 + 1/3 + 1/3 + 1/3 + 1/3 + 1/3 + 1/3) \}
\]

The process has to be repeated as many times as balance sheets are available for every company in the dataset (6) and now that all the required data is filled conveniently for the purpose of the models, I can proceed with further calculations.