A Work Project, presented as part of the requirements for the Award of a Masters Degree in Finance from the NOVA – School of Business and Economics.

Credit Default Swap (CDS) Prediction Model & Trading Strategy

Tiago Mota Dutra
Finance #642

A Project carried out on the Finance course, under the supervision of:

Prof. João Pedro Pereira

Lisbon, January 2015

Credit Default Swap (CDS) Prediction Model & Trading Strategy

Tiago Mota Dutra

Corresponding Author – Email: *tiago.mota.dutra@gmail.com*

Supervisor: João Pedro Pereira

Universidade Nova de Lisboa - *NOVA*

Faculdade de Economia – *School of Business and Economics*

Lisbon, January 2015

Abstract:

This project focuses on the study of different explanatory models for the behavior of CDS security, such as Fixed-Effect Model, GLS Random-Effect Model, Pooled OLS and Quantile Regression Model. After determining the best fitness model, trading strategies with long and short positions in CDS have been developed. Due to some specifications of CDS, I conclude that the quantile regression is the most efficient model to estimate the data. The P&L and Sharpe Ratio of the strategy are analyzed using a backtesting analogy, where I conclude that, mainly for non-financial companies, the model allows traders to take advantage of and profit from arbitrages.

Keywords: *Credit Default Swap; Econometric Prediction Model; Quantile Regression; Trading Strategy*

Index

I. Purpose of Project – General Overview

The present project is based on Credit Default Swaps (CDS). The CDS was invented by Blythe Masters from JP Morgan in 1994, and it is designed to transfer the credit exposure of fixed income products, such as loans and bonds, between two different parties. It is a financial swap agreement (see Appendix 1) in which the seller of the CDS will compensate the buyer (the creditor) in the event of a loan/bond default (by the debtor) or other credit event, usually by the face value of the loan/bond. In this scenario, the seller of the CDS takes possession of the defaulted loan/bond. In the case of non-default, the purchaser of the swap makes payments (to the seller of the swap) up until the maturity date of a contract. Consequently, a CDS is considered an insurance against non-payment by the debtor. The buyer of a credit default swap receives credit protection, whereas the seller of the swap guarantees the credit worthiness of the debt security. In doing so, the risk of default is transferred from the holder of the fixed income security to the seller of the swap. However, anyone can purchase a CDS, even buyers who do not hold the loan/bond instrument and who have no direct insurable interest in the loan/bond (these are called "naked" CDSs). If there are more CDS contracts outstanding than bonds in existence, a protocol exists to hold a credit event auction: the payment received is usually substantially less than the face value of the loan/bond.

After familiarizing with the topic of Credit Default Swaps, the identification of the empirical determinants that may explain the behavior of CDS will be presented. For the purpose of estimating a predictable model of CDS, the econometric software STATA is used. The study starts with fourteen different possible explanatory variables with daily data from January 1, 2009 to August 8, 2014. These variables are used to estimate several models in four different approaches, such as Fixed-Effect model, GLS Random-Effect model, Pooled OLS model and Quantile Regression model. In each of them the non-significant variables are different, which leads to a final model with different variables in each approach. In order to decide which model is the most appropriate, several tests and parameters are taken into consideration. By correcting the estimation errors, eliminating the non-significant variables and analyzing the R-Square, it is possible to select the best model that fits the CDS data. Implementing an explanatory model with an R-square higher than more or less 30%, there are good chances to profit by developing a trading strategy based on this prediction model.

Therefore, this project is not only an explanatory model of CDS. Considering that the econometric model is reasonable good, any financial trader will want to benefit from it. Consequently, a trading strategy of CDS is developed. Since the model gives the expected value of CDS at a specific point in time, the initial logic for a trading strategy is to be long in CDS if the expected value is higher than the actual CDS spread, or to be short if the opposite happens. The objective is to profit from the pricing inefficiency of the CDS security, with the expectation that this pricing disparity and its temporary divergence will cancel out by converging to the fair value predicted by the model.

II. A brief literature review

Although the existing academic work and literature on CDS spreads is recent, it provides interesting and important conclusions that contributed significantly to the development of the current work project. The first study of bond credit spreads, instead of the usual yields, was conducted by Collin-Dufresne *et al.* (2001). Before that, Merton (1974) predicted a relation between credit spreads and leverage, volatility, and interest rates. In 2003, another explanatory variable of credit spreads was discovered by

Campbell and Taksler (2003), where the rising idiosyncratic equity risk/volatility leads to increasing yields on corporate bonds relative to treasury bonds. Five years later, Cremers *et al.* explored the explanatory power of equity-options, such as at-the-money implied volatility and put skew, and concluded that these determinants explain one third of the credit spreads.

Recently, regarding the CDS spreads, Ericsson et al. and Zhang et al., both in 2009, proved that firm volatility, leverage and jump probability are important determinants of CDS spreads. Zhang et al (2009) showed that the volatility and jump effects are strongest for high-yield entities and financially stressed firms. In 2013, Pires, P. *et al*, when analyzing U.S. and European CDS company names, concluded that CDS premiums significantly increase with absolute bid-ask spreads across all conditional quantiles of the CDS distribution. Moreover, they also explain why a Quantile Regression is more appropriate to estimate CDS data than the usual OLS method.

III. Discussion of the topic

III.I. Variables

As explained in the previous section, some literature found correlations between several explanatory variables and credit spreads. The data in use in this project was imported from either a Bloomberg or an Eikon Thomson Reuters terminal on a daily basis over the last 6 years, from January 1, 2009 to August 4, 2014, in order to develop a daily trading strategy that presented a good performance over a considerable period of time (5 years). Considering that the data is from different companies (125 firms from the iTraxx Europe Index) and, for each company, there is six years' worth of daily data, all variables' values are organized in a panel data. The theoretical and logical relationship between the independent variables and the CDS spread is explained below.

Firstly, the dependent variable is the Credit Default Swap (CDS) spread, also known as CDS premium, price or quote. The sample consists of daily observations of the iTraxx Europe 125 corporate CDS names, where all quotes correspond to 5-year CDS contracts. The sample was imported as composite quotes from Bloomberg. The overall sample consists of a set of 125 companies over 1458 days, amounting to 182,250 CDS quotes. It is important to note that, since there are some missing data in the explanatory variables of certain companies, the total data analyzed is unbalanced.

Now, regarding the explanatory variables, there are fourteen taken into consideration. In each one of them, the economic sense and relationship with the dependent variable is explained below.

The “Daily Equity Return” variable is the percentage change of the company’s stock price from one day to another. Based on the structural model of Merton (1974), a higher drift in the firm’s asset value process increases the probability of the market value of the firm staying far from the default threshold, decreasing the probability of default and hence decreasing the CDS spread.

The “Total Debt-to-Total Equity” variable consists of a leverage ratio. This financial leverage ratio gives an idea of the company's methods of financing or its ability to meet its financial obligations. In the economic sense, the higher this ratio is, the greater the probability of not being able to meet its financial obligations, leading to a higher CDS.

The “Profit Margin”, a profitability ratio, is calculated by dividing the net income by revenues. It is an important measure since this financial metric is used to assess a business's ability to generate earnings as compared to its expenses and other relevant costs incurred. In that sense, as the profit margin increases, its ability to satisfy its obligations also increases, decreasing the risk of default. This will represent a lower CDS spread.

The "Equity Volatility" is defined as the Standard Deviation of the last thirty days equity return. Considering again the Merton (1974) theory, if the equity presents high volatility, this means that the value of the firm changes easily. This change means a high probability of crossing the default threshold and consequently a higher CDS premium.

The "Put Implied Volatility" is the thirty day implied volatility calculated as a weighted average of the two put options closest to the at-the-money strike (source: Bloomberg). The buyers of put options are afraid of a reduction in the equity price. Therefore, in order to hedge against this scenario (reduction in equity price), buying a put option is a good solution. As it is known, the equity option implied volatility is a forward looking measure of volatility, thus providing timely warnings of credit deterioration. Hence, according to Cremers *et al.* (2008), as the volatility of the put option becomes higher, the hedgers are predicting huge variations on the equity price, which means higher probability of crossing the threshold of default. This leads to higher CDS spreads.

The "Put Skew" variable is the difference between the implied volatilities of deep-out-of-the-money and at-the-money put options. Buying deep-out-of-the-money puts on the firm's equity provides protection against very large losses, especially in the case of a default where the equity price may approach zero. Hence, both these puts and CDS can be used to trade credit risk and their price must thus be closely related. The higher the put skew, the more protection is being sought in the options market, thus indicating a higher probability of a downside jump on the firm's value and hence a higher CDS spread.

"Equity Liquidity" is determined as the relative bid-ask spread in the stock market. The logic behind this variable is the same as CDS liquidity, where the bid-ask spread is a measure of liquidity. The more liquid the security is (a lower bid-ask spread defined in this project as the variable Equity Liquidity), the less asymmetry information it will represent, leading to a lower CDS spread.

"CDS Liquidity" is determined as the absolute bid-ask difference of the CDS quotes. According to the empirical study "The Empirical Determinants of Credit Default Swap Spreads", by Pires, P. *et al* (2013), absolute bid-ask spread becomes more significant to predict CDS quotes than relative spread. This variable is associated with the level of information asymmetry in the market. The more liquid a company's security is, the less information asymmetry it will represent. Therefore, since high information asymmetry is seen as a risk, high liquidity represents lower CDS premium.

The "Structural Default Probability[1]" variable, from "The StarMine Structural Credit Risk Model (SCR)" (source: Eikon Thomson Reuters), estimates the probability that a company will go bankrupt or default on its debt obligations over the next 1-yr period by assessing the equity market's view of credit risk. The equity volatility, market value of equity and liability structure are used to infer a market value and volatility of assets. The final default probability is equivalent to the probability that the market value of assets will fall below a default point (which is a function of the company's liabilities) within 1 year. It is important to note that "The StarMine Default Probabilities" are not

[1] The StarMine Default Probabilities are similar to Moody's KMV EDF. Both StarMine SCR Model and Moody's KMV EDF are built from the structural default prediction framework introduced by Merton that models a company's equity as a call option on its assets. However, the StarMine SCR Model differs by the estimation method of the drift rate, by employing a different treatment of balance sheet liabilities for banks and insurance companies, and by optimizing the formulation of volatility and default point.

inferred from CDS spreads or Recovery Rates. They are extrapolated from variables that are exogenous to the credit markets.

The “Credit Rating” variable represents an evaluation of the credit worthiness of a debtor. In this metric the historical S&P Long Term Issuer Credit Rating for each firm is used. Using dummy variables, it is possible to estimate the relationship between the Credit Ratings and the CDS spread. Therefore, the dummy variable has fifteen possible values. Value 1 is for the AAA rating and value 15 stands for the B rating. The credit rating has an inverse relationship with the possibility of debt default. As a result, a high credit rating indicates a high debtor’s ability to pay back the debt and consequently a low probability of defaulting on debt; conversely, a low credit rating suggests a high probability of default.

Some macroeconomic variables are also included to predict the CDS behavior. Firstly, the “Risk Free Interest Rate” represents the interest rate that someone can get taking no risk. For this variable, the Germany Generic Government 10Y Yield is used. A higher risk-free rate may lead to lower credit spreads. For example, Longstaff and Schwartz (1995) suggest that a higher risk-free rate increases the risk-neutral drift of the firm value process, thus reducing the probability of default and the CDS premiums.

The “Slope of the Treasury Yield Curve” is calculated as the Germany Generic Government 10Y Yield minus Germany Generic Government 2Y Yield. In Longstaff and Schwartz (1995) a rising slope lowers credit spreads. The term structure slope is a well-known leading indicator of the business cycle, with a positively sloped structure usually signaling ‘good times’ and a consequent reduction in the number of defaulted companies and its CDS spreads.

The "5-year Swap Spread" is defined as the EUR Swap Annual 5Y rate minus the Germany Generic Government 5Y Yield. This spread is commonly seen as a credit spread reflecting overall credit conditions. A high spread (high interest rate swap) is a consequence of low confidence among investors and financial institutions regarding the probability of satisfying their financial obligations, which lead to higher CDS spreads.

The last macroeconomic variable in use is the "Market Implied Volatility". It is defined as the implied volatility of at-the-money put options on the Eurostoxx 50. For this variable, the VSTOXX Index is used. VSTOXX Index is based on a new methodology jointly developed by Deutsche Borse and Goldman Sachs to measure volatility in the Eurozone. VSTOXX is based on the EURO STOXX 50 Index options traded on Eurex. It measures implied volatility on options with a rolling thirty day expiry. Collin-Dufresne et al. (2001) show that credit spreads are related to overall market volatility, i.e., as the market implied volatility goes up, the probability of crossing the default threshold increases, leading to a higher CDS spread.

III.II. Econometric Model

In order to establish a predictive model for CDS prices, different regressions are tested. In the first stage, the Fixed Effect and GLS Random Effect model are estimated.

The Fixed Effect model is determined by the following multiple linear regression for individual $i = 1; \dots; N$, which is observed at several time periods $t = 1; \dots; T$:

$$y_{it} = \alpha + x'_{it}\beta + c_i + u_{it}, \qquad i = 1, \dots, N \; and \; t = 1, \dots, T, \tag{1}$$

where $\mathrm{y_{it}}$ is the dependent variable, $\mathrm{x'_{it}}$ is a *K*-dimensional row vector of time-varying explanatory variables, α is the intercept, β is a *K*-dimensional column vector of parameters, c_i is an *individual-specific effect* represented as a dummy variable and u_{it} is an *idiosyncratic* error term.

In this model it is assumed that each company has a different and specific intercept ($\alpha + c_i$). Fixed Effect model should be used if the companies are different, but the difference in the companies correspond to the other explanatory variables. The term fixed effect is due to the fact that, although the intercept may differ across companies, each company's intercept does not vary over time; that is, it is time invariant.

The Fixed Effect model estimates the data assuming that each company is different from each other. Therefore, the Fixed Effect Model should be used if there is a correlation between something individual about that company and the other explanatory variables.

If there is indeed a relationship between the c_i (FE dummy variable) and the other explanatory variables, when this dummy variable is left out, it will cause bias in the explanatory variables' coefficients. This is the reason for using the FE model. In other words, there is a need for a control for the individual specific effects because, if not, and if an omitted variable is correlated with other explanatory variables, it causes bias in the coefficients.

The Random Effect model is given by the following multiple linear regression model for individual $i = 1; \dots; N$, which is observed at several time periods $t = 1; \dots; T$:

$$y_{it} = \alpha + x'_{it}\beta + (c_i + u_{it}), \qquad i = 1, \dots, N \; and \; t = 1, \dots, T, \tag{2}$$

In the Random Effect model it is assumed that a company specific effect does indeed exist, but it is uncorrelated with the other explanatory variables ($Corr(u_i, x'_i) = 0$). Therefore, all the companies have a common mean value for the intercept (α). If the specific effect is actually uncorrelated with the explanatory variables, the dummy variable can be left out and the model still has unbiased coefficients.

In sum, the random effects assumption is that the individual specific effects are uncorrelated with the independent variables. The fixed effect assumption is that the individual specific effect is correlated with the independent variables. If the random effects assumption holds, the random effects model is more efficient than the fixed effects model. However, if this assumption does not hold (i.e., if the Specification Hausman Test fails), the random effects model is not consistent.

The outputs of both FE and RE models are in Extra-Appendix 1. In order to decide which model is the most appropriate and the one that fits the data better, the Specification Hausman Test is used (see Extra-Appendix 2). Since the p-value is lower than 5%, the null hypothesis that the GLS Random Effect (RE) model is appropriate to fit the data is rejected, i.e., the assumption that the individual specific effect is uncorrelated with the independent variables is rejected. Hence, the alternative hypothesis that the Fixed Effect (FE) model is appropriate and more efficient than RE model can be accepted, which leads to choosing the FE model over the GLS RE model.

In a second stage, the Pooled OLS is also tested.

$$y_{it} = \alpha + x'_{it}\beta + u_{it}, \qquad i = 1, \dots, n \text{ and } t = 1, \dots, T, \tag{3}$$

This model is the most restrictive one. Under this approach, it is assumed that unobserved individual heterogeneity does not exist, i.e., the model is estimated without considering any individual specific effect (see Extra-Appendix 1). Therefore, the dummy variable for specific effect does not exist. Hence, if there is no individual heterogeneity, pooled OLS will be consistent and efficient.

On the other hand, the fixed and random effects models assume the existence of unobserved individual heterogeneity. If there is heterogeneity, a tradeoff exists between bias and precision. If the unobserved heterogeneity is uncorrelated with the explanatory variables, RE will be more efficient. If the unobserved heterogeneity is correlated with the explanatory variables, RE will be biased and the FE will be consistent and efficient.

Regarding the decision between the FE model and the Pooled OLS model, and considering the assumption of the strict exogeneity by the Pooled OLS, i.e., the errors in the regression should have conditional mean zero, it is enough to check the veracity of the assumption to decide. Therefore, for the FE model, the F test that all $u_i=0$ is rejected, which means that FE should be preferred over Pooled OLS model. Consequently, among the three models presented above, the FE model is the preferable choice.

In a third stage, the quantile regression (QR) is tested (see Extra-Appendix 3):

$$y_{it} = \alpha^{(q)} + x'_{it}\beta^{(q)} + u_{it}{}^{(q)}, \qquad i = 1, \dots, n \; and \; t = 1, \dots, T, \tag{4}$$

Where y_{it} is the dependent variable, x'_{it} it is a K-dimensional row vector of time-varying explanatory variables, $\alpha^{(q)}$ is the intercept associated with the q^{th} quantile, $\beta^{(q)}$ is a K-dimensional column vector of parameters associated with the q^{th} quantile and $u_{it}{}^{(q)}$ is an idiosyncratic error term associated with the q^{th} quantile. In quantile regression we have $\beta^{(q)}$ instead of β to make clear that different choices of quantile estimate different values for the parameter.

The main difference between this model and the others already tested in this project is the fact that in the QR model, instead of regressing on the mean of the dependent variable, it regresses on the median or any other quantiles. Considering that no distribution assumptions are made in QR, this difference makes the estimation model efficient even when the data is not normally distributed.

Therefore and according to the conclusions of Pires, P., Pereira, J. and Martins, L. (2013), the main advantages of QR are the flexibility of modeling data with heterogeneous conditional distributors; the fact that the median regression is more robust to outliers than the pooled OLS regression; the efficiency when the distribution of the data is skewed; and the richer characterization and description of the data: can show different effects of the independent variables across the spectrum of the dependent variable.

The QR model is estimated for five different quantiles: 0.10; 0.25; 0.50 (median); 0.75 and 0.90 (see Appendix 2). For each quantile regression, the non-significant variables are excluded, leading to final estimation models with different explanatory variables. As concluded by Amato, J. and Remolona, E. (2003), the goodness-of-fit of the model increases with CDS premiums, in accordance with the credit spread puzzle.

For the economic interpretation purpose of the signal of each coefficient output, the median (0.50 quantile) regression is used. Considering this, the output signal for each coefficient should be in accordance with the economic interpretation of each explanatory variable. Note that for the case of CDS and Equity Liquidity, since these variables represent the bid-ask spread, the higher the bid-ask spread is, the higher the CDS and Equity liquidity measure value is, leading to a lower CDS quote. The same happens with the Credit Ratings variable, where a higher value for this measure (dummy variable) stands for a lower credit rating, which leads to a higher CDS spread.

The QR allows for the identification and analysis of the different impacts that the same explanatory variable has in the high quantile CDS spreads in contrast to the same impact in the low quantile CDS spreads. Contrary to the standard OLS models, in the QR models the relationship is between the independent variables and the conditional quantiles of the dependent variable, rather than just the conditional mean of the CDS quote, which means that QR gives a more comprehensive view of the effect of the independent variables on the dependent variable.

According to the graph in Extra-Appendix 4, the CDS spread daily data is sorted by quantiles from the minimum value (0.0138%) to the maximum (16.23%). From the graph it is possible to realize that at the 0.90 quantile there is high CDS spreads compared to the rest of quantiles. In addition, through this graph it is possible to verify that the median is around 1% while the mean is around 7.5%. This difference is justified by the impact that outliers have on the mean, i.e., the mean is being affected by skewness. Therefore, since the typical OLS assumes normal distribution of data and the estimation is based on the dependent's variable mean, it will be less efficient due to the skewness. On the other hand, since the QR output is based on the dependent variable's median (or any other quantiles), the estimation is less affected by skewness, remaining efficient.

Using the graphs of Extra-Appendix 5, it is possible to analyze the different impact of each explanatory variable in the CDS spread for each quantile, as well as understand the need for using Quantile Regression over the OLS regression. The coefficient magnitudes are in the vertical axis while the CDS quantiles are in the horizontal axis. The OLS coefficient is plotted as a horizontal line with the confidence interval as two horizontal lines around the coefficient line. This OLS coefficient does not vary by quantiles. The quantile regression coefficients are plotted as lines varying across the quantiles with confidence intervals around them. If the quantile coefficient is outside the OLS confidence interval, then there are significant differences between the coefficients of the QR and the coefficients of the Pooled OLS regression. The same happens with the zero value. If the confidence intervals of the quantile regression coefficients do not include the zero value, this means that they are statistically different than zero (statistical significant). For the purpose of interpretation, the graph of the "Equity Volatility" variable's coefficient is used. Note that QR is more efficient when the effect of the independent variable varies across the level of the dependent variable. In a QR model, it is assumed that the effect is not constant across spectrum. In this graph, only for the 0.80 quantile the coefficient is not statistically different from the OLS coefficient. The same happens with all the other coefficients, where only in a specific and tight quantile the coefficient is not statistically different from the OLS coefficient, meaning that the use of a Quantile Regression is more suitable than the typical OLS regression. Another interpretation can be taken from the equity volatility coefficient's graph: it is possible to conclude that the effect of the Equity Volatility explanatory variable increases for CDS with higher spreads (higher quantiles). The same analogy can be made for the rest of the explanatory variables.

With the purpose of comparing the output of the Pooled OLS regression with the output of the Quantile Regression, Appendix 2 is used. Through the table in this appendix, it is possible to analyze the different impacts of the explanatory variables in the CDS spread according to the several quantiles. For example, in the case of Equity Liquidity, if this explanatory variable increases 1%, the CDS spread increases 1.58 basis points (0.0158%) at the 0.10 quantile. For the 0.90 quantile, if the Equity Liquidity increases 1%, the CDS spread increases 37.63 basis points (0.3763%). In other words, the effect of Equity Liquidity increases for the CDS with higher spreads (higher quantiles), which justifies the use of a QR over the Pooled OLS that has a non-varying coefficient.

Considering that the Fixed Effect, Random Effect and Pooled OLS regressions assume that the data follows a normal distribution and that they are not efficient when the data presents heteroskedasticity, the following tests have been conducted:

The Breusch and Pagan test (see Extra-Appendix 6) is significantly different than zero, which means that there is heteroskedasticity. Since the Quantile Regression is still efficient even the data presents heteroskedasticity, the use of quantile regression over OLS is justified.

The Jarque Bera Normality test (see Extra-Appendix 7) is significantly different than zero, which leads to the rejection of the normality of the data. Since the FE, RE and Pooled OLS regressions assume normal data, the use of QR is preferred.

After opting for the use of QR over the other regressions tested in this project, a test to verify if the QR explanatory model becomes better when the CDS spreads of the financial companies are separated from the CDS spreads of the non-financial companies was conducted. To do so, the R-square of the QR with all the companies is compared with the R-square of the QR with only financial companies and the R-square of the QR with only non-financial companies. Through Appendix 3, it is possible to conclude that the R-square of both at 0.50 quantile (42.48% for non-financial companies and 44.91% for financial companies) is higher than the single R-square of the QR with all the companies included (41.48%). Given this conclusion, instead of proceeding with a unique model for both, two distinct models for the financial and non-financial companies are established in order to develop the trading strategy.

III.III. Trading Strategy

In this section a trading strategy is developed in order to give practical utility to the econometric model presented above. To do so, a trading strategy is developed. A trading strategy is a fixed plan designed to achieve a profitable return by going long and/or short positions in markets. For the trading strategy purpose, each QR estimated has different explanatory variables for each quantile, conforming if they are statistically significant or not in the respective quantile, in order to guarantee a more accurate and efficient strategy.

Considering that the econometric model developed in this project is a predictable model, the main idea of the trading strategy is to take advantage and profit from differences between the actual CDS quote and the one predicted by the model. As an analogy, see the Euromoney report by Currie and Morris (2002). Therefore, if the model predicts a price for the next day higher than the today's price, the CDS must be bought and vice-versa. This logic is known as "Convergence Price Theory" and the strategy as "Capital Structure Arbitrage". Before achieving the final strategy, different approaches were tested in order to produce the one with the best performance (highest Sharpe Ratio and Accumulated Profit).

Although the model presents a good explanatory power (0.50 quantile's R-square of 44,91% for financial companies and 42,48% for non-financial companies), it is essential to backtest in order to understand which strategy would have performed better in the past. Assuming that the strategy will have approximately the same performance as in the backtesting, the one that gives the highest Sharpe Ratio and Accumulated Profit should be chosen.

Before giving an in-depth explanation of the strategy, it is important to understand how the CDS trading works in real life. As already explained, the trader that buys a CDS has to pay a coupon (which can be quarterly, semiannually or annually) pre-defined (usually 100 bp). As it is known, the CDS spread that is quoted in the market varies on a daily basis according to, among other factors, the demand and supply for this security. Consequently, if a CDS that is quoted in the market with a spread of 84 basis points is bought, the buyer will still pay a coupon of 100 bp (quarterly, semiannually or annually). To compensate, he/she must receive the difference in advance, i.e., he/she must receive the following initial cash amount:

$$P = \sum_{t=1}^{T} \frac{-(CDS\ spread_t - 100bp)}{(1+y)^t} \times N\,, \tag{5}$$

$$with\ t = 1; \ldots; T\ quarters/semesters/years$$

Where P stands for the cash amount that he/she must receive/pay in front, N is the Notional, i.e., the amount that the buyer is investing and has in exposition, and y is the discount rate from the SWAP curve. Considering that the buyer is long in a CDS with 84 bp of spread and assuming that in the next day the CDS spread goes up to 90 bp and he/she decides to sell his/her position, the following profit should be made:

$$\Delta P = \Delta CDS\ spread \times D^* \times N \tag{6}$$

Where D^* represents the modified duration of the respective 5-year CDS. To better understand this explanation, please see Appendix 4.

In all strategies' approaches, a CDS position is entered into whenever the actual and the prediction CDS spread diverged from each other by a proportional threshold amount. After several approaches had been tested, the final version of the strategy was achieved. The decision rule of this strategy is to go long when the predicted CDS spread is higher than the real CDS spread of the day before (t-1) plus a certain proportion of the Standard Deviation (SD) of the CDS spreads, and to go short when the inverse happens. This SD corresponds to the first one sixth of the data, i.e., in this part of the data no long/short decision is made for the backtesting purpose: it is only used to get entry indicators, such as the SD. In this approach a stop loss limit is defined: when the daily return of the previous day is lower than a certain value (stop loss limit), there will not be a position for the current day. On the next day it is possible to go long or short again according to the criteria previously explained, as usual. If those criteria are not respected, no position is kept until the inverse happens (see Extra-Appendix 8).

The predicted CDS spread is calculated by using the output from the QR estimation, where the coefficients from the output are multiplied by the respective known variable's value:

$$\hat{y}_{it} = \widehat{\alpha}^{(\mathrm{q})} + x'_{it}\hat{\beta}^{(q)}, \quad i = 1, \dots, n \; and \; t = 1, \dots, T, \tag{7}$$

This is done for the 0.10, 0.25, 0.50, 0.75 and 0.90 quantiles of the Non-Financial companies and Financial companies. In other words, if the real CDS of the day before is in the 0.10 quantile of data, the today's predicted CDS is calculated using the coefficients estimated in the 0.10 QR. The same logic is applied for the remain quantiles.

Regarding the daily return (in percentage), it is calculated by multiplying the daily 5-year CDS spread variation (in percentage) by a general modified duration, which, by assumption, is considered to be 4.5 years for all CDS securities:

$$Daily\ return\ \%_t = (CDS\ spread\ \%_t - CDS\ spread\ \%_{t-1}) \times D^* \tag{8}$$

Regarding the daily stop loss limit value and the constant S referred above, the values that lead to the highest Accumulated Profit for those parameters are obtained by using the function “solver” in excel. Therefore, the combination that results from the solver is a daily stop loss of -3.25% and a proportion of SD of 2.27 times for the Non-Financial companies, and a daily stop loss of -11.82% and a proportion of SD of 6.27 times for Financial Companies.

In all approaches tested, the same procedures are made in order to decide which one has the best performance. For all of them, the solver excel function is used to get the indicators’ values that provide the greater Accumulated Profit. Then, the final choice is based on the Sharpe Ratio. The approach that presents the highest Sharpe Ratio is the one chosen to proceed with the trading strategy. For SR calculation purpose, a risk free rate of 0.75% (Germany Generic Government 10Y Yield in Oct 14) was assumed to determine the excess return.

In Extra-Appendix 9, some statistics and graphs of the trading strategy for the non-financial companies, financial companies and both together are presented.

Some conclusions can be drawn from these outputs. As it is possible to verify in Extra-Appendix 9, the performance in the financial companies is not as good as expected, presenting a SR of only 0.59. Conversely, the trading strategy for non-financial companies has a good performance with a SR of 1.10. By including both in one unique trading strategy, the SR obtained is 0.97. Although in the three cases the accumulated profit is considerably low, this is not a huge concern since, as already explained, the initial cash amount paid is not the notional/exposure amount, which means that the position can be easily leveraged, keeping the same SR. In other words, it is possible to have an investment in CDS with a notional of €10 million where it is just necessary to pay (or receive) a much lower initial cash amount:

$$P = \sum_{t=1}^{T} \frac{-(CDS\ spread_t - 100bp)}{(1+y)^t} \times N \quad (9)$$

Therefore, through this trading strategy it is possible to achieve an annual return of 2.31% (average annual return of both financial and non-financial companies' trading strategy) over a notional amount of €10 Million[2], without paying any initial cash amount (assuming a CDS spread equal to 100 bp). In this scenario, the only expenditure that the CDS buyer has is the 100 bp times the notional amount (€10 MIO), resulting in an annual payment of €100.000, as illustrated in Appendix 5.

Comparing the individual performance of the non-financial companies with the financial companies, the trading strategy of the non-financial companies presents much better results (Sharpe Ratio of 0.56 vs. Sharpe Ratio of 1.10). The main reason for this difference is the fact that financial companies, by definition, are much more volatile (see Zhang *et al*. (2009)), increasing the Standard Deviation and consequently the risk incurred. In that sense, a fair decision is to opt for trading only non-financial companies due to its stabilized character, getting a higher Sharpe Ratio than trading both financial and non-financial companies.

It is commonly known that the CDS has a high correlation with bonds (debt market). Through the graphs in Extra-Appendix 9, it is possible to verify that the trading strategy's daily returns present more volatility between 2011 and 2012 and consequently higher Standard Deviation. The reason behind this volatility is connected to the government debt crisis that happened in those years. The fragile economic

[2] 2.31% x €10 MIO = €231.000 annually

situation of Portugal, Greece, Italy, Ireland, Spain and other European countries and the request for external help and debt restructuration of some of these countries led to a reduce on the investor's confidence and an increase in the market's instability. This justifies an increase in the traders' speculation and the huge variations of the CDS spreads, which are directly linked to the government debt movements. This situation only stabilized when the European Central Bank (ECB) announced the introduction of the LTRO (Long Term Refinancing Operations[3]) program at the end of 2011, to help ease the Eurozone crisis.

Considering this high volatility period, a new approach to the trading strategy was tested, where some hedge positions were included in order to avoid the instability caused by the market movements (systematic risk), such as the ones verified in 2011. In other words, with the hedge positions, it is supposed to only be exposed to the specific risk and not to the idiosyncratic risk. The most obvious way found to hedge was to have the inverse position that existed in the CDS, in the own iTraxx Europe 125 Index. Therefore, if the position is long/short in a CDS, a short/long in the iTraxx Europe 125 Index should be taken by the same notional value. Using this approach, the strategy's return represents only the return specific to the CDS, excluding the return that may be due to other reasons unrelated to that company. This strategy did not present the performance that was expected, not being acceptable to proceed with the trading activity. The reasons that may explain this performance is owing to the fact that the CDS prediction model already includes four macroeconomic variables, which means that the whole model already takes the market movements into consideration. This means that it does not make sense to try to hedge against the idiosyncratic risk, something that is already predicting the market movements.

[3] LTRO program is a cheap loan scheme for European banks.

It is important to note that, since a CDS is an over-the-counter (OTC) security, i.e., they are traded via a dealer network as opposed to on a centralized exchange, the bid-ask spread is wider due to the lower liquidity of these kind of markets (OTC markets). Considering that the mid-spread was used when developing the trading strategy, this assumption may distort the conclusions from the trading strategy. Another assumption made when developing the trading strategy was the non-existence of accrued interests.

IV. Conclusion

In this project, in addition to the development of a CDS prediction model based on both specific and macroeconomic explanatory variables, a trading strategy based on the Convergence Price Theory is also developed in order to guarantee the practicality of the model in real trading life.

Regarding the CDS prediction model, it is proved that, according to the specifications of CDS, a quantile regression (which when estimating the coefficients considers the dependent variable's median instead of the mean) is more efficient and accurate to fit the data than the traditional OLS. This is justified by the heterogeneity of the dependent variable, where the impact of the explanatory variables on CDS spreads varies according to whether firms have conditionally high or low risk. Moreover, the median regression is more robust to outliers than the pooled OLS regression which, considering that the distribution of the data is positive skewed, justifies the choice of a QR model. Moreover, with the Quantile regression, it is possible to estimate different coefficients for as many quantiles as desired, leading to different final CDS prediction models, thereby leading to a more accurate and better performed trading strategy.

Hence, regarding the trading strategy, although the average annual profit is relative low (around 2% a year), a worthy Sharpe Ratio and the CDS trading's method, where the initial cash amount invested is much lower than the notional amount (allowing for leverage), make me confident in the trading strategy's quality and the respective opportunity to take advantage and profit from the CDS prediction model. Considering the specific characteristics of financial companies, where the returns are much more volatile, reducing the Sharpe Ratio to 0.59, the use of a trading strategy only for the non-financial companies' CDS is suggested, since it presents much less risk and a lower standard deviation of returns.

V. References

Amato, J. and Remolona, E. 2003. "*The Credit Spread Puzzle*", BIS Quarterly Review, 51–63.

Campbell, J. and Taksler, G., 2003. "*Equity Volatility and Corporate Bond Yields*", Journal of Finance, Vol. 58, No. 6.

Collin-Dufresne, P., Goldstein, R. and Martin, J. 2001. "*The Determinants of Credit Spread Changes.*" The Journal of Finance, Vol. 56, No. 6

Cremers, M., Driessen, J., Maenhout, P. and Weinbaum, D., "*Individual Stock‑Option Prices and Credit Spreads.*" Journal of Banking and Finance, Vol. 32, 2008, pp. 2706‑15.

Currie, A., and J. Morris. 2002. "*And Now for Capital Structure Arbitrage*." Euromoney (December):38-43

Ericsson, J., Jacobs, K. and Oviedo, R., 2009. "*The Determinants of Credit Default Swap Premia*", Journal of Financial and Quantitative Analysis, Vol. 44.

Koenker, R. and Bassett, J. 1978. "*Regression Quantile.s.*" Econometrica, Vol. 46, No. 1 (January, 1978)

Longstaff, Francis A. and Schwartz, Eduardo S. 1995. *"A Simple Approach to Valuing Risky Fixed and Floating Rate Debt."* The Journal of Finance, Vol. 50, No. 3, Papers and Proceedings Fifty-Fifth Annual Meeting, American Finance, Association, Washington, D.C., January 6-8, 1995. (Jul., 1995), pp. 789-819.

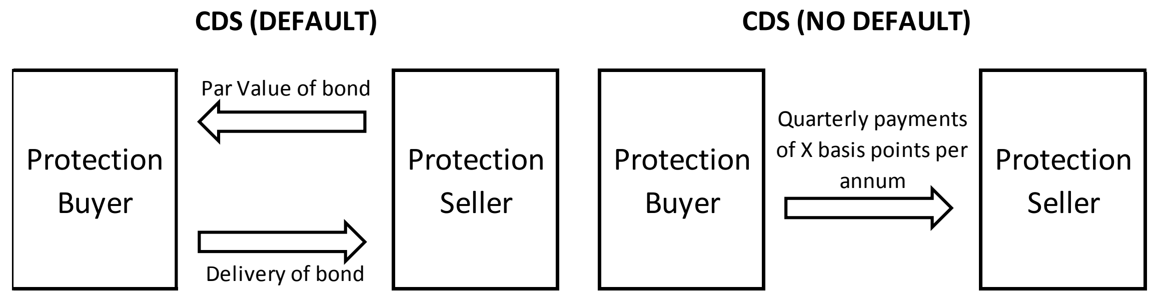
Merton, R.C. 1974. *"On the Pricing of Corporate Debt: The Risk structure of Interest Rates."* Journal of Finance, Vol. 29, No. 2

Pires, Pedro., Pereira, João. and Martins, Luís. 2013. "*The Empirical Determinants of Credit Default Swap Spreads: a Quantile Regression Approach.*" European Financial Management, Vol. 9999, No. 9999, 2013, 1-34

Zhang, B., Zhou, H. and Zhu, H., 2009. "*Explaining credit default swap spreads with the equity volatility and jump risks of individual firms"*, Review of Financial Studies, Vol. 22, 2009, pp. 5099‑5131.

VI. Appendices

Appendix 1 – CDS contract illustration



Appendix 2 – Pooled OLS vs. Quantile Regression Outputs by quantiles

	Pooled OLS	.10 QR	.25 QR	.50 QR	.75 QR	.90 QR
Equity Return	-.1085991*	.1116075	*	*	*	-.4990048
	(-1.69)	(2.38)	*	*	*	(-4.91)
Debt-to-Equity	7.54E-06	7.71e-06	8.12e-06	.0000137	6.98e-06	*
	-7.96	(6.13)	(8.46)	(16.71)	(9.54)	*
Profit Margin	-0.0004002	-.0001102	-.0005992	-.0009843	-.0007838	*
	(-6.57)	(-2.81)	(-15.73)	(-18.67)	(-12.12)	*
Equity Volatility	18.79513	11.66134	12.94759	14.53608	17.52392	23.29411
	-57.47	(46.45)	(55.28)	(51.49)	(53.33)	(44.87)
Put Implied Volatility	2.129287	1.038721	1.241379	1.461252	1.992459	2.634595
	-92.83	(55.14)	-74.58	(73.71)	(84.14)	(66.70)
Put Skew	2.053833	1.072702	1.417518	1.741376	2.506614	3.265822
	-28.3	(12.37)	(23.10)	(27.73)	(34.14)	(27.36)
Equity Liquidity	23.27108	1.583463	6.19434	16.77271	27.39624	37.62943
	-60.2	(5.06)	(23.18)	(50.13)	(59.94)	(42.68)
CDS Liquidity	5.867065	3.540653	4.80847	6.351532	8.12867	9.311201
	-179.75	(109.10)	(191.08)	(224.91)	(237.44)	(155.03)
Structural DP	0.061307	.0757791	.1059037	.1825096	.1813807	.1868445
	-21.58	(26.03)	(44.44)	(74.29)	(71.78)	(49.30)
Rating	0.0819932	.0554131	.0618934	.0667756	.0746055	.0803061
	-135.42	(145.36)	(164.82)	(127.41)	(102.77)	(64.95)
Rf Interest rate	-2.618036	-4.763832	-5.807943	-5.760369	-4.04223	-1.424513
	(-10.39)	(-29.05)	(-34.94)	(-26.42)	(-15.04)	(-3.31)
Treasury yield curve slope	-13.13478	-13.06628	-13.10441	-13.85594	-10.58565	-9.24833
	(-20.71)	(-30.79)	(-30.79)	(-25.26)	(-15.79)	(-8.64)
5-year swap spread	69.51416	41.67326	46.80322	48.2411	49.15824	63.12581
	-67.11	(49.60)	(61.62)	(53.84)	(47.14)	(39.24)
Market Implied Volatility	1.830041	.7966775	1.020732	1.083088	1.478467	1.954649
	-63.84	(37.68)	(50.92)	(43.80)	(49.23)	(39.95)
Constant	-0.6313262	-.3691274	-.4490565	-.4941998	-.5753137	-.6379091
	(-65.70)	(-53.71)	(-68.84)	(-59.43)	(-55.47)	(-38.64)
R-squared	0.6667	.2840	.3410	.4148	.4850	.5407

t-stats in brackets; one star for non statistical significant at 0.05 level

Appendix 3 –Quantile Regression of non-financial and financial companies

	Non-Financial					Financial				
	.10 QR	.25 QR	.50 QR	.75 QR	.90 QR	.10 QR	.25 QR	.50 QR	.75 QR	.90 QR
Equity Return	.1275111	*	*	*	*	*	*	*	*	*
	(2.83)	*	*	*	*	*	*	*	*	*
Debt-to-Equity	7.19e-06	8.03e-06	8.55e-06	6.92e-06	4.97e-06	.0000561	-.0000243	-.0001008	-.0002125	-.000379
	(9.11)	(11.03)	(12.33)	(9.17)	(3.86)	(10.63)	(-4.79)	(-13.93)	(-20.85)	(-28.34)
Profit Margin	-.0018517	-.0019977	-.0013522	-.0011457	-.0014206	*	-.0008339	-.0007593	-.0007334	*
	(-48.84)	(-38.58)	(-17.48)	(-10.78)	(-8.98)	*	(-13.79)	(-8.79)	(-7.41)	*
Equity Volatility	9.009949	11.67051	15.17604	18.43201	24.79477	*	8.540058	14.05026	13.47688	20.28915
	(33.76)	(42.11)	(48.82)	(47.47)	(38.88)	*	(19.68)	(23.91)	(17.28)	(21.49)
Put Implied Volatility	.710102	.842247	.9477413	1.056415	1.044004	.8952088	2.037596	2.550943	3.444967	4.627032
	(35.20)	(42.16)	(43.56)	(39.03)	(22.31)	(35.59)	(61.48)	(55.37)	(54.49)	(54.99)
Put Skew	.9496305	1.251704	1.490625	2.20533	2.90434	*	-.0347689	-1.039228	*	*
	(11.71)	(19.53)	(24.27)	(29.21)	(21.13)	*	(-5.64)	(-6.28)	*	*
Equity Liquidity	1.007157	3.814043	14.11664	24.13719	31.06385	*	9.555332	16.15757	31.43578	27.34068
	(3.22)	(12.57)	(41.64)	(49.81)	(33.49)	*	(18.76)	(20.33)	(23.96)	(13.09)
CDS Liquidity	3.291738	4.291013	5.782102	7.818921	9.944256	2.362793	3.840092	4.89616	6.370376	5.622795
	(131.76)	(163.54)	(191.79)	(195.20)	(140.37)	(43.45)	(72.11)	(80.42)	(76.97)	(46.22)
Structural DP	.0682761	.1069194	.1717153	.2897211	.3275975	*	-.0347689	-.0438158	-.0413056	-.0965333
	(13.02)	(22.92)	(42.54)	(84.02)	(77.98)	*	(-10.77)	(-9.69)	(-6.26)	(-10.55)
Rating	.0727969	.08302	.0905158	.0997789	.1078643	.07212	.059754	.0694772	.0941905	.1317963
	(203.97)	(198.11)	(161.80)	(120.94)	(73.82)	(52.01)	(47.45)	(35.47)	(33.47)	(35.68)
Rf Interest rate	-5.703468	-6.582318	-5.926824	-4.578487	-2.194612	-11.94887	*	-1.235713	-2.59567	-7.701674
	(-37.88)	(-37.53)	(-28.41)	(-17.47)	(-5.30)	(-33.47)	*	(-2.02)	(-2.91)	(-7.10)
Treasury yield curve slope	-12.91307	-11.03426	-9.505849	-8.644842	-7.239738	29.89479	-10.07108	-16.07828	-21.46499	-30.11585
	(-33.30)	(-24.99)	(-18.31)	(-13.34)	(-7.08)	(29.91)	(-11.34)	(-10.56)	(-9.62)	(-10.42)
5-year swap spread	32.30332	40.01891	36.74004	32.03653	34.44785	130.8127	143.9229	151.0545	157.5989	191.0455
	(44.47)	(52.96)	(43.59)	(30.97)	(21.71)	(75.54)	(75.84)	(59.06)	(45.47)	(42.51)
Market Implied Volatility	.3656809	.5188488	.4181153	.4515775	.6773351	1.250591	2.417912	3.021578	3.538939	4.076833
	(18.91)	(25.05)	(17.79)	(15.69)	(14.75)	(23.77)	(46.31)	(40.39)	(32.88)	(27.99)
Constant	-.4360138	-.594658	-.7149224	-.7798079	-.7991586	-.5806224	-.2911829	-.2455709	-.3110188	-.5337047
	(-70.68)	(-89.08)	(-88.95)	(-73.00)	(-45.53)	(-28.84)	(-15.60)	(-9.44)	(-8.74)	(-11.77)
R-squared	.3008	.3554	.4248	.4953	.5465	.2577	.3853	.4491	.5134	.5529

t-stats in brackets; one star for non statistical significant at 0.05 level

Appendix 4 – Illustration of CDS payments

Go long in a CDS with spread = 84 bp
Notional = 10 MIO
Discount rate from SWAP curve

Years	t=0	t=1	t=2	t=3	t=4	t=5	
Coupon Pre-defined (paid)		-100	-100	-100	-100	-100	b.p.
Coupon quoted ("should pay")		84	84	84	84	84	b.p.
I'm paying more than what is quoted		-16	-16	-16	-16	-16	b.p.
Initial received cash amount	**96,791 €**						

If the CDS spread goes up to 90 bp and I sell my position
Notional = 10 MIO
Discount rate from SWAP curve

Years	t=0	t=1	t=2	t=3	t=4	t=5	
Coupon Pre-defined (received)		100	100	100	100	100	b.p.
Coupon quoted ("should receive")		-90	-90	-90	-90	-90	b.p.
I'm receiving more than what is quoted		10	10	10	10	10	b.p.
Compensation (I pay ahead)	**-67,141 €**						

Initial received cash amount	96,791 €
Compensation (I pay ahead)	-67,141 €
Profit & Loss	**29,650 €**

Δ % Spread	0.06%
Mod. Duration (years)	4.94
Notional	10 MIO
ΔP = Δ % Spread × Mod. Duration × Notional	**29,650 €**

Appendix 5 – CDS annual payments illustration

Go long in a CDS with spread = 100 bp
Notional = 10 MIO
Discount rate from SWAP curve

Years	t=0	t=1	t=2	t=3	t=4	t=5	
Coupon Pre-defined (paid)		-100	-100	-100	-100	-100	b.p.
Coupon quoted ("should pay")		100	100	100	100	100	b.p.
I'm paying more than what is quoted		0	0	0	0	0	b.p.
Initial paid/received cash amount	**0 €**						

Trading Strategy's annual return = 2.31%
Notional = 10 MIO

Years	t=0	t=1	t=2	t=3	t=4	t=5	
Coupon Pre-defined (paid)		-100	-100	-100	-100	-100	b.p.
Trading Strategy's annual return = 2.31%		2.31%	2.31%	2.31%	2.31%	2.31%	
Notional = 10 MIO		10 MIO	10 MIO	10 MIO	10 MIO	10 MIO	€
P&L		**131,000 €**	**131,000 €**	**131,000 €**	**131,000 €**	**131,000 €**	

EXTRA

APPENDICES

Extra-Appendix 1 – Fixed Effect, Random Effect and Pooled OLS Models Estimation

	Fixed Effect (FE)	Random Effect (RE)	Pooled OLS
Equity Return	0.1309558	0.1305669	-.1085991*
	-2.72	-2.71	(-1.69)
Debt-to-Equity	0.00000334	0.00000337	7.54E-06
	-4.35	-4.4	-7.96
Profit Margin	-0.0001016*	-0.0001062	-0.0004002
	(-1.92)	(-2.01)	(-6.57)
Equity Volatility	15.69452	15.71789	18.79513
	-60.54	-60.6	-57.47
Put Implied Volatility	1.597726	1.600275	2.129287
	-82.84	-82.96	-92.83
Put Skew	-0.1934741	-0.1873557	2.053833
	(-3.35)	(-3.24)	-28.3
Equity Liquidity	13.2545	13.29275	23.27108
	-39.32	-39.42	-60.2
CDS Liquidity	3.990147	3.993594	5.867065
	-146.18	-146.26	-179.75
Structural DP	0.0822216	0.0822199	0.061307
	-31.77	-31.77	-21.58
Rating	0.1422101	0.1412768	0.0819932
	-112.35	-112.38	-135.42
Rf Interest rate	-0.8414393	-0.8349338	-2.618036
	(-4.41)	(-4.37)	(-10.39)
Treasury yield curve slope	-9.336355	-9.351161	-13.13478
	(-19.57)	(-19.59)	(-20.71)
5-year swap spread	86.08983	86.03488	69.51416
	-109.93	-109.8	-67.11
Market Implied Volatility	0.9272666	0.9312598	1.830041
	-40.69	-40.85	-63.84
Constant	-0.9901436	-0.9831941	-0.6313262
	(-84.96)	(-44.24)	(-65.70)
R-squared	0.6173	0.6183	0.6667

t-stats in brackets; one star for non statistical significant at 0.05 level

Extra-Appendix 2 – Specification Hausman Test

```
. hausman Fixed .

Note: the rank of the differenced variance matrix (11) does not equal the number of coefficients being
        tested (13); be sure this is what you expect, or there may be problems computing the test.
        Examine the output of your estimators for anything unexpected and possibly consider scaling your
        variables so that the coefficients are on a similar scale.

                  ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     Fixed        Random       Difference          S.E.
-------------+----------------------------------------------------------------
NetDebtShare |  -.0003586    -.0003539        -4.68e-06        1.85e-06
ProfitMargin |  -.0020172    -.0020071        -.0000102               .
 EquityVolat |   15.78387     15.80902        -.0251546               .
 PutImpVolat |   1.767914     1.771783        -.0038692               .
     PutSkew |  -.2903816    -.2878094        -.0025721               .
EquityLiquid |   21.29959      21.3878        -.0882161               .
  CDSLiquid  |    3.47173      3.47625        -.0045193               .
StructuralDP |    .028975     .0290225        -.0000475        .0001026
      Rating |   .1153448     .1149683         .0003765        .0001635
   RfIntRate |   -.960407     -.950722         -.009685               .
Treasyield~e |  -5.398952    -5.418983          .020031               .
yrswapspread |   102.8442     102.7791         .0651259               .
MarketImpl~y |   1.189676     1.195703        -.0060271               .
------------------------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from xtreg
           B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(11) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                           =       29.84
                Prob>chi2 =      0.0017
                (V_b-V_B is not positive definite)

.
```

Extra-Appendix 3 – Quantile Regression Methodology

The quantile regression (QR), introduced by Koenker and Basset (1978), is an extension of the conditional mean to a collection of models for different conditional quantile functions. The quantile regression is describe by the following equation:

$$y_{it} = \alpha^{(q)} + x'_{it}\beta^{(q)} + u_{it}{}^{(q)}, \quad i = 1, \dots, n \; e \; t = 1, \dots, T, \qquad (1)$$

Where y_{it} is the dependent variable, x'_{it} it is a K-dimensional row vector of time-varying explanatory variables, $\alpha^{(q)}$ is the intercept associated with the q^{th} quantile, $\beta^{(q)}$ is a K-dimensional column vector of parameters associated with the q^{th} quantile and $u_{it}{}^{(q)}$ is an idiosyncratic error term associated with the q^{th} quantile.

While the Ordinary Least Square (OLS) minimizes $\sum u_{it}{}^2$, the QR minimizes

$$\sum q|u_{it}| + \sum (1-q) \times |u_{it}|, \tag{2}$$

a sum that gives the asymmetric penalties $q|u_i|$ for underprediction and $(1-q) \times |u_i|$ for overprediction.

In a standard regression which uses OLS, it will give a one single output of parameters/estimators/coefficients. In the QR, the output is several sets of outputs of parameters/estimators/coefficients for each quantile.

In QR we have $\beta^{(q)}$ instead of β to make clear that different choices of q^{th} estimate different values of β.

Quantile Regression Coefficients

The q^{th} conditional quantile of y_{it} given x_{it} is

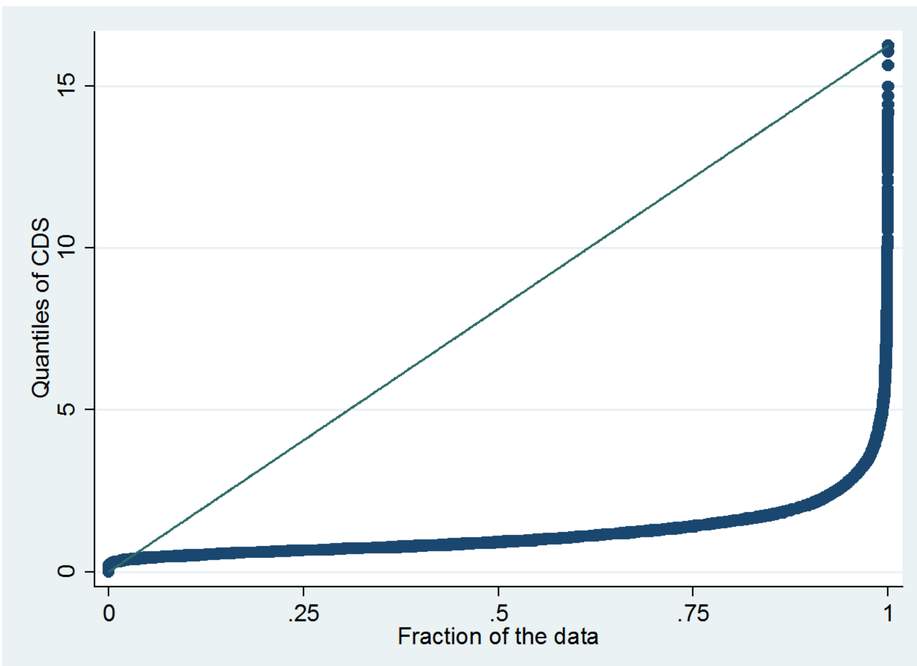
$$Q_q(y_{it}|x_{it}) = x'_{it}\beta^{(q)} \tag{3}$$

which follows from the necessary assumption concerning the error term, $u_{it}{}^{(q)}$, $Q_q\left(u_{it}{}^{(q)}\middle|x_{it}\right)$, i.e., the conditional q^{th} quantile of the error term is equal to zero. The quantile regression method allows the marginal effects to change at different points in the conditional distribution by estimating the partial derivatives of the conditional quantile function with respect to the set of explanatory variables,

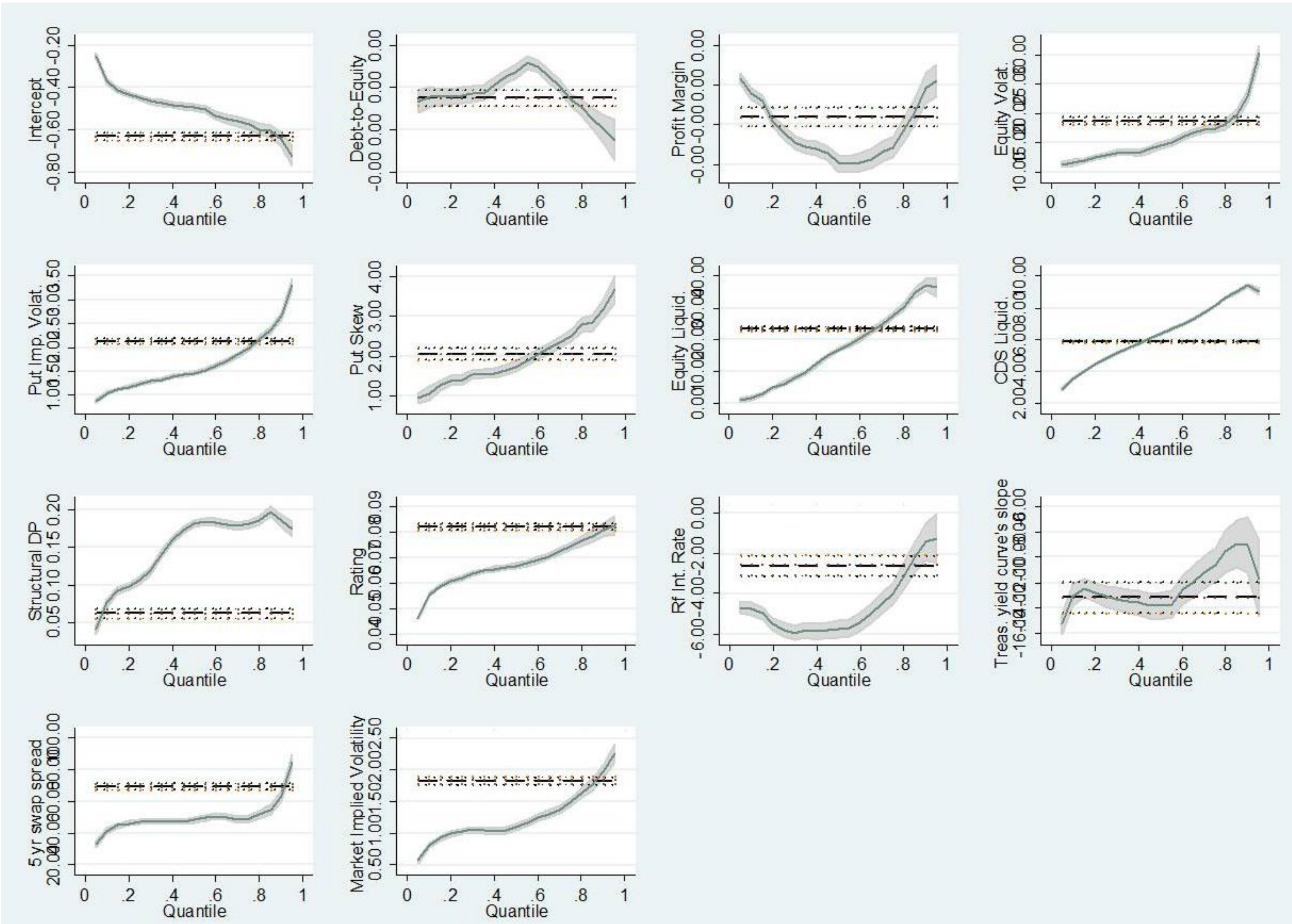
$$\frac{\delta Q_q(y_{it}|x_{it})}{\delta x} = \beta^{(q)} \tag{4}$$

Using different values for q^{th}, this way allowing for parameter heterogeneity. A quantile regression parameter $\beta^{(q)}$ estimates the change in a specified quantile q^{th} of the dependent variable y produced by a one unit change in the independent variable x. The marginal effects are for infinitesimal changes in the regressor, assuming that the dependent variable remains in the same quantile.

Extra-Appendix 4 – Dependent Variable by Quantiles



Extra-Appendix 5 – Coefficients plotted by quantiles



Extra-Appendix 6 – Breusch-Pagan test for heteroskedasticity

```
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: DebttoEquity ProfitMargin EquityVolat PutImpVolat PutSkew EquityLiquid
                    CDSLiquid StructuralDP Rating RfIntRate Treasyieldcurvesslope yrswapspread
                    MarketImpliedVolatility

         chi2(13)      = 14340.88
         Prob > chi2  =   0.0000
```

Extra-Appendix 7 – Jarque Bera Normality test

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2) (joint)	Prob>chi2 (joint)
myResiduals	1.4e+05	0.0000	0.0000	.	.

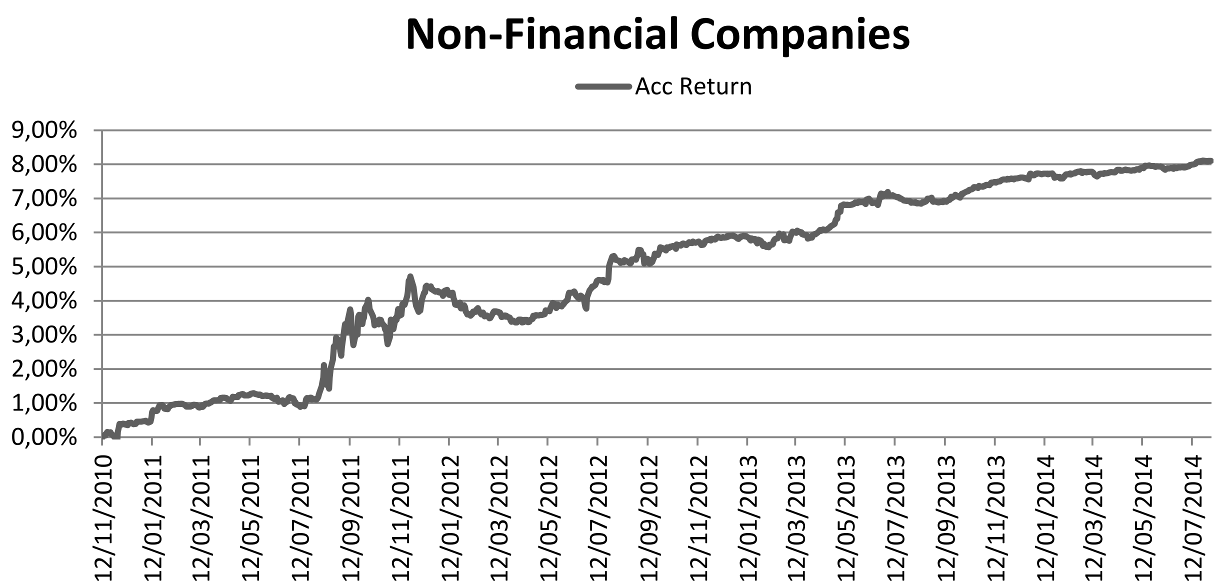
Extra-Appendix 8 – Decision rule's illustration

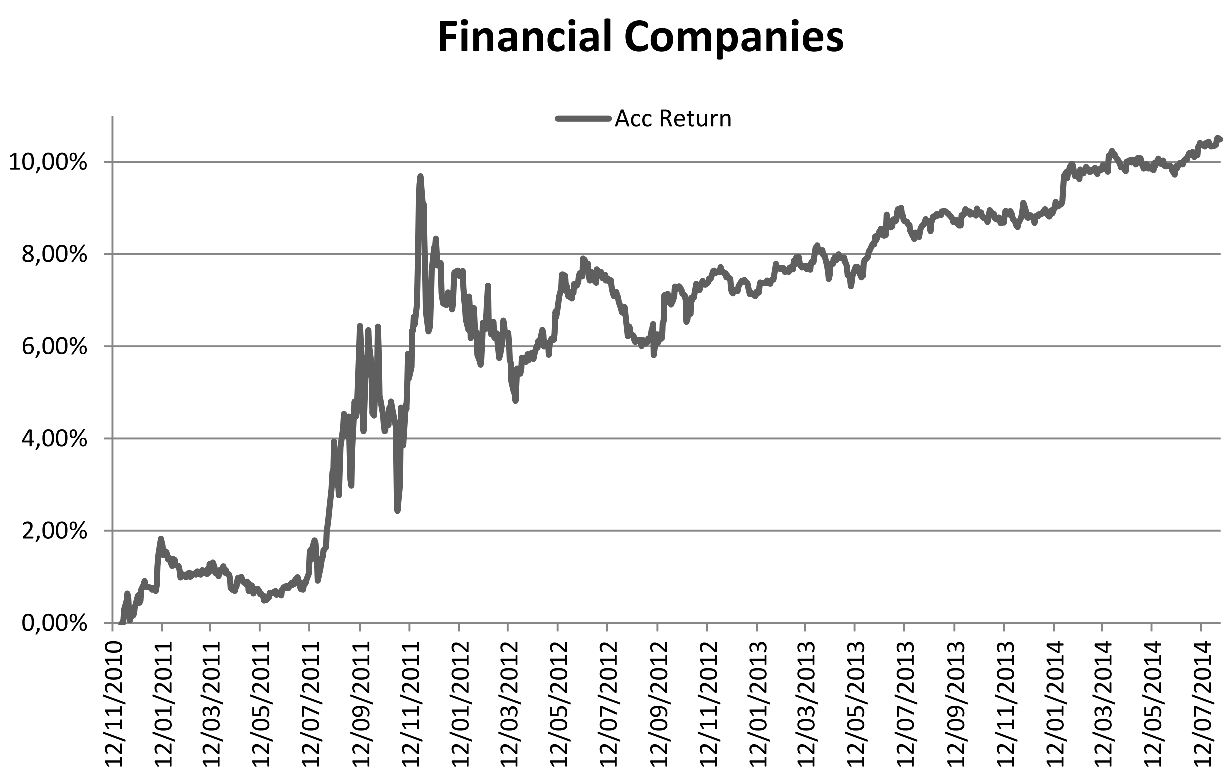
1. If $DR_{t-1} < DSL$, no position taken in day t.
2. If 1. is not verified:
   - 2.1. If $predicted\ CDS_t > (real\ CDS_{t-1} + S \times SD_t)$, long position taken in day t.
   - 2.2. If $predicted\ CDS_t < (real\ CDS_{t-1} - S \times SD_t)$, short position taken in day t.
   - 2.3. No position taken in day t if condition 2.1. and 2.2. are not verified

Where DR is Daily Return, DSL stands for Daily Stop Loss, SD is the Standard Deviation of the CDS spreads and S is a constant value which works as a threshold.

Extra-Appendix 9 – Trading Strategy output

	Non-Financial	Financial	Both
Average Annual Return	2.17%	2.80%	2.31%
Excess Return	1.42%	2.05%	1.56%
Annual SD	1.28%	3.51%	1.60%
Sharpe Ratio	**1.10**	**0.59**	**0.97**
# Positive Returns	508	463	524
# Negative Returns	463	508	448
Maximum Daily Positive Return	0.57%	1.67%	0.77%
Maximum Daily Negative Return	-0.43%	-1.45%	-0.55%


Non-Financial Companies
Acc Return
9,00%
8,00%
7,00%
6,00%
5,00%
4,00%
3,00%
2,00%
1,00%
0,00%
12/11/2010
12/01/2011
12/03/2011
12/05/2011
12/07/2011
12/09/2011
12/11/2011
12/01/2012
12/03/2012
12/05/2012
12/07/2012
12/09/2012
12/11/2012
12/01/2013
12/03/2013
12/05/2013
12/07/2013
12/09/2013
12/11/2013
12/01/2014
12/03/2014
12/05/2014
12/07/2014


Financial Companies
Acc Return
10,00%
8,00%
6,00%
4,00%
2,00%
0,00%
12/11/2010
12/01/2011
12/03/2011
12/05/2011
12/07/2011
12/09/2011
12/11/2011
12/01/2012
12/03/2012
12/05/2012
12/07/2012
12/09/2012
12/11/2012
12/01/2013
12/03/2013
12/05/2013
12/07/2013
12/09/2013
12/11/2013
12/01/2014
12/03/2014
12/05/2014
12/07/2014


Both: Financial & Non-Financial Companies
Acc Return
9,00%
8,00%
7,00%
6,00%
5,00%
4,00%
3,00%
2,00%
1,00%
0,00%
12/11/2010
12/01/2011
12/03/2011
12/05/2011
12/07/2011
12/09/2011
12/11/2011
12/01/2012
12/03/2012
12/05/2012
12/07/2012
12/09/2012
12/11/2012
12/01/2013
12/03/2013
12/05/2013
12/07/2013
12/09/2013
12/11/2013
12/01/2014
12/03/2014
12/05/2014
12/07/2014