Predicting Credit Default Swap (CDS) Returns with Machine Learning

Terrence Yidong Zhang

Warrington College of Business Administration, University of Florida

Faculty mentor: Mahendrarajah Nimalendran, Department of Finance, Insurance and Real Estate

Abstract

Credit Default Swaps ("CDS") are contracts that insure one party against default in an underlying financial instrument, usually a bond. Therefore, the price of CDS reflects the perceived risk of default in an underlying financial instrument. This project applied Support Vector Machines ("SVMs") to the prediction of CDS price changes for several individual companies across time. Previous research applying SVMs to predicting CDS prices used historical CDS prices as model inputs. This project proposed and applied several new input variables. Tests over a period of several years, across a group of CDS time-series, indicate that a combined model which uses the new input variables in addition to historical CDS price changes outperforms models that only use historical CDS price changes.

Introduction

Credit Default Swap ("CDS") contracts are financial derivatives whose value is tied to the probability of default in an underlying security ("reference security"). This reference security is typically a fixed income instrument such as a loan or a bond. The exact circumstances which define default events vary from case to case, but in a broad sense, default events occur when the borrower of a loan or bond fails to make interest or principal payments for an extended period of time.

CDS contracts are written on a reference security, where a counterparty relationship exists between a protection buyer and a protection seller. To make an analogy to insurance policies, the protection seller acts as an insurance provider and agrees to compensate the protection buyer in the event of default in the underlying security. In return, the protection buyer makes periodic payments to the protection seller, over the life of the CDS contract. These periodic payments are collectively known as "CDS premiums", and reflect the risk of the underlying security defaulting during the life of the CDS contract.

CDS contracts are typically priced according to the size of the premium associated with each particular contract. The premium paid is commonly quoted as a percentage of the notional value (size) of the CDS contract. This percentage (known as the "CDS spread") is usually stated as an annualized rate and is paid periodically by the protection buyer to the protection seller until a default event occurs or the contract expires. The CDS spread can be interpreted as the price of a CDS contract, and can be used as a basis for comparing various CDS contracts to each other.

Literature Review

As CDS contracts primarily deal with insuring the protection buyer against default, the primary driver of CDS spreads is the probability of default in the reference security. This section highlights previous work regarding the prediction of default probabilities, CDS spreads, and related issues.

Van Gestel et al. (2003) applied Support Vector Machines ("SVMs") to the prediction of credit ratings for financial institutions. Using a dataset of over 3000 banks, each with over 70 possible input variables, Least-Squares SVMs were compared to traditional techniques such as Fisher Discriminant Analysis and Logistic Regression. They conclude that the Least-Squares SVM method outperformed other techniques.

Huang et al. (2004) applied various machine learning methods to predicting credit ratings of individual bonds. The analysis included companies in Taiwan and the United States, and used accounting variables and ratios as input variables. They concluded that SVMs were competitive with Backpropagation Neural Networks ("BNNs") in predictive capability, and that those two machine learning methods generally outperformed Logistic Regression.

Shin et al. (2005) analyzed over 2000 manufacturing firms in Korea, and constructed bankruptcy prediction models using various machine learning methods. They found that SVMs outperformed BNNs, especially over smaller datasets. Possible explanations for the relative outperformance of SVMs are that SVMs have fewer hyperparameters and can generalize better.

Callen et al. (2009) performed cross-sectional analyses of levels and changes in CDS spreads, in addition to a short-window event study of CDS spreads around earnings reports. They describe several interesting characteristics of CDS spreads, such as an inverse relationship between CDS spreads and quarterly earnings.

Gündüz & Uhrig-Homburg (2011) applied SVMs, structural models (the Merton model), and reduced-form (constant intensity) models to single-name CDS bid-ask quotes from January 2001 to December 2004. In the cross-sectional portion of the analysis, SVMs underperformed the Merton model and constant-intensity model. However, in the time-series portion of the analysis, SVMs outperformed the Merton model and constant-intensity model when making one-, five-, and ten-step-ahead predictions for the levels of CDS spreads. Previous values of the CDS time-series served as input variables into the SVMs.

Data and Methodology

This paper expands on the time-series analysis of Gündüz & Uhrig-Homburg (2011) by (1) shifting the objective from predicting the levels of CDS spreads to predicting the changes ("returns") of CDS spreads and by (2) introducing several new input variables ("features").

Data

This project analyzed CDS contracts written on the credit of individual companies ("single-name CDS"). In addition to CDS spread data, the analysis incorporated new features in the form of financial variables, such as stock returns and implied volatilities.

CDS spread data. Market participants, such as banks, insurance companies, hedge funds, and other financial institutions, buy and sell CDS contracts over-the-counter[1] ("OTC"). Therefore, time-series data for CDS spreads are sparse and depend on the trading volume of each particular CDS contract.

Bloomberg L.P. is a data provider specializing in financial services, and distributes analytics tools through the Bloomberg Terminal. Bloomberg provides data for CDS spreads, in the form of bid and ask quotes submitted by sell-side banks[2] for various single-name CDS.

This project used Bloomberg to retrieve daily time-series data for CDS contracts on 17 individual companies. The timespan of the data extends from 2010 to 2017 and includes both bid and ask quotes. The 17 individual companies (Table 1) were all domiciled in the United States and were selected based on the availability of both bid and ask quotes for the corresponding CDS spreads. This analysis avoided contracts with low CDS spreads (indicating a low probability of the reference security defaulting), since these contracts are likely to be less volatile than contracts with higher spreads. In addition, this analysis filtered out contracts with large bid-ask spreads, since large bid-ask spreads tend to indicate that the corresponding contract suffers from

greater illiquidity. All CDS spreads considered in this project were for CDS contracts with five-year maturities, since five-year CDS are more commonly traded than other maturities.

financial variables. For each of the 17 individual companies, we retrieved several additional time-series from Bloomberg. These daily time-series are the stock price, the 30-day implied volatility[3]of the stock price, the level and the 30-day implied volatility[4] of the relevant sector index, the level and 30-day implied volatility of the S&P 500 Index, and ask quotes for the Markit CDX North America Investment Grade Index and the Markit CDX North America High Yield Index.

The rationale for considering the stock price data is that stock prices may be more sensitive to information regarding a company's financial health, due to the higher liquidity of the stock market in comparison to the CDS market. Implied volatility tends to rise in anticipation of negative changes to the underlying stock price, and therefore may also communicate information about changes in the probability of default. Broad index data for stocks and CDS contracts are included to capture changes in macroeconomic trends that may affect default risk for an individual company. All data are calculated at a daily frequency over the years 2010 to 2017. Average Bid/Ask Spread is calculated as a percentage of the bid quotes for each company.

Table 1. Individual Companies Analyzed

Company	Industry	CDS Quotes Available	Avg. Bid/Ask Spread
Ford Motor Co	Consumers	1792	5.28%
Kohl's Corp	Consumers	1727	5.81%
Xerox Corp	Technology	1757	5.89%
Staples Inc	Consumers	1707	6.37%
Best Buy Co Inc	Consumers	1698	6.48%
Freeport-McMoRan Inc	Industrials	1697	6.62%
Pitney Bowes Inc	Technology	1763	6.76%
Expedia Inc	Consumers	1770	6.77%
Avnet Inc	Technology	1772	6.99%
Macy's Inc	Consumers	1752	7.00%
The Gap Inc	Consumers	1729	7.06%
Arrow Electronics Inc	Technology	1771	7.56%
Domtar Corp	Industrials	1752	8.00%
Olin Corp	Industrials	1783	8.12%
Motorola Solutions Inc	Technology	1784	7.13%
L Brands Inc	Consumers	1631	6.57%
Nordstrom Inc	Consumers	1759	8.49%

Methodology

The raw dataset for each company in Table 1 exists at a daily frequency, although missing quotes for CDS spreads and implied volatilities lead to each variable having a different length and each raw dataset having a different number of observations.

One benchmark model and three SVMs were created and compared for this project. The "Naïve" benchmark model predicted that the next period's return will be equal to the current period's return. The "Base" SVM used the historical CDS time-series only and had three features, which were the historical returns from one-, two-, and three- periods ago. The "New" SVM used the eight financial variables listed in the Data section as model features. The "Combo" SVM combined the features of the "Base" model and "New" model, for a total of eleven features.

This experiment considered three prediction intervals for the returns of both bid and ask quotes: one-, five-, and ten-day ahead returns. For each prediction interval, the raw datasets were modified such that they would only contain one-, five-, or ten-day returns, respectively. Missing quotes were filtered out, such that for any particular time period in the modified dataset, all variables contain a numeric value. This leads to some discontinuities as certain time periods may be absent from the modified dataset.

All testing was done out-of-sample, in a chronological manner. Every prediction was made after training the SVM and performing hyperparameter tuning with cross-validation (see Table 2). The training and cross-validation periods were constructed such that, for each particular prediction, only data that was available at that point in time was used. This prevented any of the models from getting an advantage by accessing data that was not yet available at the time of the prediction. When each model was retrained, both the training interval and the cross-validation interval were shifted forward to incorporate new data so that the sizes of the intervals remained consistent.

All analysis was carried out in the R programming language, using the RStudio environment. The R package "e1071" was used to perform support vector regression. All three SVMs in this project were constructed as epsilon-insensitive regressions using the radial basis function ("RBF") kernel. Three hyperparameters (gamma, cost, and epsilon) were tuned during cross-validation.

Table 2. Experiment Design, by prediction interval

	One-day	Five-day	Ten-day
Training Periods	1000	75	25
Cross-validation Periods	30	30	25
Predictions Made	164	121	46
Retraining Frequency	5	1	1

The numbers in this table are for each individual time-series. Retraining Frequency refers to the number of predictions made before the model was retrained on new data.

Results

The four models made predictions at each prediction interval for both bid and ask time-series, over each of the 17 companies. To make inferences regarding prediction accuracy, the actual returns were then regressed against the predictions in the form of $A_t = \alpha + \beta * (P_t)$, where A_t and P_t refer to the actual return for time t and the predicted return for time t, respectively.

From these linear regressions, the beta coefficients and adjusted r-squared values were calculated and compared to each other. A beta coefficient closer to 1 indicates better prediction performance (on average) for the method being analyzed, while a higher r-squared value indicates more explanatory power.

For a particular prediction interval and quote type (bid or ask), each of the four models had 17 beta values and 17 r-squared values (one for each individual company). Therefore, to make pairwise comparisons between two models, Wilcoxon Rank-Sum tests were conducted on both the beta and r-squared variables. The p-values of these tests are shown in Tables 3 through 13.

For the pairwise tests on beta, the null hypothesis stated that neither model had a beta that was significantly closer to 1 than the other model. For the pairwise tests on r-squared, the null hypothesis stated that neither model had a higher r-squared than the other model. The results of these pairwise comparisons are discussed below.

One-Day Ahead Predictions

When comparing the models on the basis of beta, the "Combo" model performed the best across both bid and ask quotes, while the "Naïve" model performed the worst (Table 3, Table 4). However, on the basis of r-squared, both the "Naïve" and "Combo" models outperformed the other models (Table 5, Table 6).

For each value shown in the tables, the model in the corresponding row performed better than the model in the column.

Table 3. Beta, ask quotes

P-values	Naïve	Base	New	Combo
Naïve				
Base	1.74%			
New	0.00%			
Combo	0.00%	2.24%		

Only significant p-values are shown.

Table 4. Beta, bid quotes

P-values	Naïve	Base	New	Combo
Naïve				
Base	0.00%			
New	0.01%			
Combo	0.00%	3.98%	5.44%	

Only significant p-values are shown.

Table 5. R-squared, ask quotes

P-values	Naïve	Base	New	Combo
Naïve		3.57%	6.01%	
Base				
New				
Combo				

Only significant p-values are shown.

Table 6. R-squared, bid quotes

P-values	Naïve	Base	New	Combo
Naïve		1.98%	2.53%	
Base				
New				
Combo		0.19%	0.87%	

Only significant p-values are shown.

Five-Day Ahead Predictions

When comparing on the basis of beta, the "Combo" model stands out as having the best performance (Table 7 and Table 8). On the basis of r-squared, the results are less conclusive, but still suggest that both the "New" and "Combo" models outperformed the other models (Table 9, Table 10).

For each value shown in the tables, the model in the corresponding row performed better than the model in the column.

Table 7. Beta, ask quotes

P-values	Naïve	Base	New	Combo
Naïve				
Base				
New	0.55%	7.97%		
Combo	0.01%	0.47%		

Only significant p-values are shown.

Table 8. Beta, bid quotes

P-values	Naïve	Base	New	Combo
Naïve				
Base				
New				
Combo	0.00%	1.01%	6.62%	

Only significant p-values are shown.

Table 9. R-squared, ask quotes

P-values	Naïve	Base	New	Combo
Naïve				
Base				
New				
Combo		8.71%		

Only significant p-values are shown.

Table 10. R-squared, bid quotes

P-values	Naïve	Base	New	Combo
Naïve				
Base				
New	4.43%	7.27%		
Combo	7.27%			

Only significant p-values are shown.

Ten-Day Ahead Predictions

When comparing on the basis of beta, both the "New" and "Combo" models show greater performance (Table 11, Table 12). Comparisons on the basis of r-squared are less clear-cut. For instance, none of the pairwise comparisons were significant on an r-squared basis for ask quotes. But the results still suggest that the "New" and "Combo" models may have slightly higher performance (Table 13).

For each value shown in the tables, the model in the corresponding row performed better than the model in the column.

Table 11. Beta, ask quotes

P-values	Naïve	Base	New	Combo
Naïve				
Base				
New	3.99%	7.11%		
Combo	3.24%	0.53%		

Only significant p-values are shown.

Table 12. Beta, bid quotes

P-values	Naïve	Base	New	Combo
Naïve				
Base				
New	0.64%	0.75%		
Combo	9.50%	2.24%		

Only significant p-values are shown.

Table 13. R-squared, bid quotes

P-values	Naïve	Base	New	Combo
Naïve				
Base				
New		7.27%		
Combo		7.97%		

Only significant p-values are shown.

The model that consistently performed better across all three prediction intervals was the "Combo" model, with the "New" model performing second best. This conclusion is more clearly supported by comparisons on the basis of beta, whereas the results of comparisons on the basis of r-squared were a lot less definitive.

The beta variables for the one-day prediction interval were mostly positive and relatively close to 1, across all four models. For the five-day prediction interval, the beta variables began to drift towards 0 and further away from 1. This trend continued for the ten-day prediction interval, where many beta values turned negative. This decrease in predictive power also occurred when looking at adjusted r-squared values, which tended to decrease across the board as the prediction interval lengthened. This suggests that returns over longer time intervals are more difficult to predict.

Conclusion

This project analyzed the returns of CDS spreads (bid and ask quotes) across time. 17 individual companies were considered and four different models were assigned the task of predicting returns across three different time intervals. Previous work in the literature suggested that the historical CDS time-series alone can predict the future path of CDS spreads. The results of this project add to the literature by suggesting that the best performing model ("Combo" model) is in fact one that conducts a support vector regression using a set of features including stock returns, changes in implied volatility, and the historical CDS time-series itself.

Areas for further study on this topic can be discussed in terms of modifications to the machine learning methods and improvements in experiment design.

From a machine learning perspective, other machine learning methods such as Neural Networks, Deep Learning, or Boosted Trees should be applied to the prediction of CDS spreads. In addition, very little time in this project was dedicated to feature selection or feature engineering. Future work can apply a different set of features to each single-name CDS, based on the industry classification or financial characteristics of the underlying company.

From an experiment design perspective, better conclusions could be drawn by gathering a more comprehensive dataset across more companies and industry sectors. CDS contracts with maturities other than five-years should also be tested. With a larger dataset, more controls can be applied to the dataset. For instance, CDS on highly profitable firms are likely to behave differently than CDS on less profitable firms. The same idea applies to firms with strong credit ratings and firms with weak credit ratings. Finally, if CDS spreads can be predicted with a reasonable degree of accuracy, the feasibility of creating a trading strategy around those predictions should be assessed. This presents a tough challenge since CDS contracts are traded over-the-counter and detailed transaction data is difficult to obtain.

References

Callen, J. L., Livnat, J., Segal, D. "The Impact of Earnings on the Pricing of Credit Default Swaps." *Accounting Review* 84 (2009): 1363-1394

Gündüz, Y., Uhrig-Homburg, M. "Predicting credit default swap prices with financial and pure data-driven approaches." *Quantitative Finance* 11.12 (2011): 1709-1727

Huang, Z., Chen, H., Hsu, C., Chen, W., Wu, S. "Credit rating analysis with support vector machines and neural networks: a market comparative study." *Decision Support Systems* 37.4 (2004): 543-558

Shin, K., Lee, T., Kim, H. "An application of support vector machines in bankruptcy prediction model." *Expert Systems with Applications* 28.1 (2005): 127-135

Van Gestel, T., Baesens, B., Garcia, J., & Van Dijcke, P. "A Support Vector Machine Approach to Credit Scoring." *Bank en Financiewezen* 2. (2003): 73-82

Endnotes

[1]While CDS contracts trade over-the-counter, various exchange operators and financial institutions (CME Group, ICE, DTCC, etc.) provide a variety of clearing and data collection services for CDS trades.

[2]Source: Bloomberg Help Desk.

[3]Implied volatilities are derived from the prices of options contracts. Bloomberg uses the Black-Scholes model to calculate implied volatilities, using linear interpolation across strike price and time. This paper used the implied volatility derived from at-the-money option contracts.

[4]Because we did not have access to data on sector index options, implied volatilities for the sector indices are approximated using options on the SPDR Exchange Traded Funds that correspond to each particular sector index.