PERFORMANCE MODELS FOR LIQUID BOND PORTFOLIO MANAGEMENT

Alessandra Orsoni and David Iglesias

Kingston University – Faculty of Business and Law

A.Orsoni@kingston.ac.uk (Alessandra Orsoni)

Abstract

Financial traders play in a really mature market where competition makes the availability of up-to-date information essential for managers to make the correct decisions when the portfolio policies are set. This paper proposes intelligent models of trader's performance under variable market conditions to support the effective hedging of the portfolio and ensure the overall profitability of the trading desk. The research examines liquid bond trading in a London-based investment bank and develops intelligent models to assess and improve the performance of their trading team. Intelligent models based on AI techniques have been widely used in financial applications to forecast financial indices over time. This research differs from the existing work in that it takes these indices as input and uses this information to estimate traders' performance, based uniquely on their individual history of performance. Furthermore, this is not a forecasting task, it is a correlation/association task, as the range of possible market conditions is defined and the model is built and trained on data available over the entire range. Another novel aspect is the idea of using AI to model the behaviour of a particular individual while the existing work ranging from ANNs to agent-based simulation targets the behaviour of the average individual within a category, so the idea is not to model the behaviour of an average trader, but the behaviour of one trader in particular. Agent-based simulaton offers the opportunity to model the evolution of individuals through their interaction with the environment and with other agents, however the rules of evolution have to be predefined and coded into the agent upon creation, in this research instead the behavioural pattern of the individual is extracted directly from his/her own history. Finally, the possibility to model the adaptability/evolution of the individual trader and his/her ability to learn from experience is accounted for in the modular design of the ANNs, as ANN modules may be regularly re-trained on recent data or even continually trained to keep up with the individual changes in trading performance.

Keywords: intelligent systems, decision support, financial models, liquid bond trading.

Presenting Author's biography

ALESSANDRA ORSONI is a Senior Lecturer in the School of Business Information Management at Kingston University (Kingston, UK). She received both her MS in Mechanical Engineering and her ScD in Engineering Systems Design and Innovation from the Massachusetts Institute of Technology (Cambridge, MA). Prior to joining Kingston University she was a research associate in the Department of Materials Science and Metallurgy at the University of Cambridge (Cambridge, UK) and worked as an assistant professor in the Department of Production Engineering at the University of Genova (Genova, Italy).

1. Introduction

Liquid bonds are debt securities which trade frequently with variable commission. Liquidity is the term used to describe how easy it is to sell a trade. Highly liquid bonds include U.S. Treasuries, which trade billions of dollars every day; however, liquid bonds also include the bonds of a company (viewed as) close to bankruptcy, which only appeal to speculators and, thus, trade much less frequently. Liquidity has a direct impact on the commission paid to trade a bond, which unlike stocks, rarely trade on a fixed commission schedule.

The study presented in this paper is based upon a liquid bond trading desk operating in a London-based investment bank. The context of operation is a highly competitive market with tight margins and a large portfolio to control. Depending on the volatility of the market, coverage actions may be needed to protect the portfolio. These include for instance limiting the risk for some of the traders or changing the limit on their notional amount, considering that some traders may be more profitable in specific market conditions due to experience, knowledge and different degrees of risk adversity. In order to profitably run the trading desk, managers need to quickly adjust the individual traders' responsibilities and the portfolio's hedges in response to changes in the market conditions.

In consideration of the identified needs, the research seeks to establish a correlation between the market conditions and the performance of the individual trader. To this effect the research examines the use of different techniques including regression, exponential smoothing, and AI techniques based on Artificial Neural Networks (ANNs).

2. Variable Definition: I/O

Important challenges for the proposed modelling work are in the choice of the indices required to describe the market conditions and their volatility, as well as in the measurement of each trader's performance. Due to the size of the liquid bond market it would be impossible to recover all the data required for a complete representation of the market. However, extensive discussion with banking experts suggested that the market conditions could be described using the price of a single bond characterised by the highest possible liquidity. Because there is always offer and demand for highly liquid bonds, under this premise, market conditions could be said to be accurately represented by the bond's price. After careful consideration of the available bonds, the 10Y German Bond was selected for use in this research. The German government is a very reliable bond issuer, additionally the selected bond has AAA seniority and the highest liquidity in the market. Bloomberg is the source system used in the bank as a reference provider of market prices. The list of prices requested to Bloomberg encompassed the past two years. 2005 and 2006 market prices were collected and saved in a two-data columns format: Value Date and Market Price measured in euro. Market volatility and market direction also need to be defined as part of the required input to the performance model. In the trading environment historical volatility is measured by the standard deviation. The volatility associated to each market price was then calculated as the standard deviation across the preceding 50 days. The direction of the market is simply defined as the price differential between two consecutive days (e.g. between today's and yesterday's price). The designated set of input for the representation of the market conditions includes the price, volatility and direction measures defined above. Traders' profitability was the most difficult set of variables to represent. Banks are reluctant to show any set of figures that could give clues on their performance. However, the authors reached a confidentiality agreement with the bank managers ensuring discretion at any time when the data was required. For this purpose, figures had to be altered in such a way that the original correlation between traders' performance and bond prices could be maintained, and any private data like traders names was deleted from the set of data.

2. Data Pre-processing

Data integrity, in terms of maintaining coherence between value dates in the different data sets and avoiding missing figures in some of the fields, is extremely important to this modeling work. To ensure coherence, an Excell-based function was defined to check the date columns of each set for integrity against the others in the input sets. A new column was added in all data sets for these purposes: the market direction, market volatility, and traders P&L sets all had an added in column. Each row in that new column would be set to zero, if the date in the corresponding row was found to be the same as in the other data sets, it would be set to one if the date was found to be different. This procedure was applied to the entire data set and summed at the end. A sum of zero would indicate date integrity (i.e. no discrepancies). A sum greater than zero would indicate the presence of date discrepancies among the data sets. Market direction, market volatility and traders P&L data sets produced a sum larger than zero.

Some of the value dates existing within the P&L data set, where missing in the market data sets. Mainly, discrepancies were due to the different bank holidays in different countries (UK and Germany). As an example the trading date 2006-01-02 existed, but the market data was unavailable for that date. Therefore, no market direction or volatility were available, additionally both of them would depend on value in previous dates which could cause further discrepancies during the modelling stage.

Finally, once the causes of mismatch were identified, all sets of data were filtered, deleting any row of incomplete data (i.e. characterized by missing fields in one set or another). The reason for deleting missing value dates was to ensure a correspondence between patterns of profitability and the market conditions across all sets of data. Other approaches could have been adopted to deal with missing dates, for instance Jason E. Kutsurelis (1998), chooses "the average closing between of the day preceding and following the missing day" as a figure to represent the missing date. This approach was not viable in the case of this research, because mapped patterns needed to be real ones. Hence the decision to delete incomplete records rather than replacing missing entries with 'artificially determined' figures.

For manipulation purposes, all the data sets were integrated within a unique Excell-based sheet consisting of fourteen columns. Eleven columns contained the traders' P&L, another two contained the market volatility and the mark direction, respectively. The last column was used to index the data by date. The total number of rows within the sheet was 234 to represent the data encompassing the year 2006. All the monetary figures were expressed in units of euro.

3. Statistical Modelling Techniques

Prior to AI-based modelling, the researchers looked at statistical evidence of a correlation between the performance of the individual traders and the market conditions. For demonstration purposes correlation and regression analysis techniques were applied to the most promising trader in terms of correlation strength. The results from this preliminary modelling have been used as reference to comparatively assess the benefits of modelling based on neural networks.

For a selected sample trader, identified as trader A, the Pearson correlation coefficient between performance and both market price and volatility was calculated using SPSS.

As shown in Table 2, the correlation coefficient calculated to assess the strength of the relationship between trader A's P&L and the market direction was -0.551. Experts indicated as significant any correlation higher than 0.5. However, the correlation between trader A's P&L and the market volatility was only -0.04, indicating that there was not a significant correlation. Therefore, pattern association for this variable could only be determined from local correlations in small ranges of market conditions.

Table 1: Correlation Results for Trader A

Correlations

		P&L trader A	Market value	Market volatility
P&L trader A	Pearson Correlation	1	-.551**	-.040
	Sig. (2-tailed)		.000	.544
	N	234	234	234
Market value	Pearson Correlation	-.551**	1	-.514**
	Sig. (2-tailed)	.000		.000
	N	234	234	234
Market volatility	Pearson Correlation	-.040	-.514**	1
	Sig. (2-tailed)	.544	.000	
	N	234	234	234

**. Correlation is significant at the 0.01 level (2-tailed).

Additionally, the possibility to have a lag between the market conditions and the traders' P&L was tested (see table 2). SPSS-based analysis demonstrated the absence of a lag between traders' profitability and market conditions. This result confirmed the "real time" impact that events in the liquid bond market have on traders' performance.

Table 2: Correlation Results for Trader A

Cross Correlations

Series Pair: date with P&L trader A

Lag	Cross Correlation	Std. Error[a]
-7	-.546	.066
-6	-.545	.066
-5	-.546	.066
-4	-.549	.066
-3	-.548	.066
-2	-.546	.066
-1	-.546	.066
0	-.551	.065
1	-.537	.066
2	-.519	.066
3	-.501	.066
4	-.483	.066
5	-.464	.066
6	-.447	.066
7	-.434	.066

a. Based on the assumption that the series are not cross correlated.

Table 2 shows trader A's cross correlation, as dependent on the lag. Lag, is here intended as the time taken by any event in the market to impact on traders' results. The highest degree of correlation was found for the analysis done with 0 lag. This finding confirmed the researchers' expectation that any impact on the markets has immediate consequences on traders' results. Experience shows that, while figures on the data set are updated on a daily basis, market events affect traders almost instantaneously.

Based on the correlation results just shown for trader A, it made sense to develop a regression model to link trader A's P&L and the market price (as volatility was not deemed significant for this trader).

SPSS was used again for regression analysis, the corresponding linear regression equation is reported below (equation 1):

$$P = 883806.44 - 351828.63 \times Mp \qquad (1)$$

Where P indicates Trader's performance and Mp indicates market price.

Tables 3 to 5 show the SPSS output based on which, the parameters for equation 1 were determined.

Table 3: Summary of Regression Model

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.551[a]	.304	.301	20275.74609	.048

a. Predictors: (Constant), Market value

b. Dependent Variable: P&L trader A

Table 4: ANOVA Table for Regression

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.2E+010	1	4.165E+010	101.323	.000[a]
	Residual	9.5E+010	232	411105879.3		
	Total	1.4E+011	233			

a. Predictors: (Constant), Market value

b. Dependent Variable: P&L trader A

Table 5: Regression Coefficients

Coefficients[a]

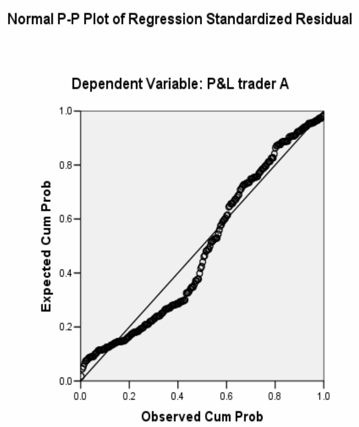
Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	883806.4	82228.891		10.748	.000
	Market value	-351829	34952.491	-.551	-10.066	.000

a. Dependent Variable: P&L trader A

Finally, the residual plots showed a normal distribution of the residuals (see figure 1). Residual were normally distributed and their Durbin – Watson statistical measurement threw an autocorrelation of 0.048. This low autocorrelation demonstrates that no important explanatory variables are missing, therefore the model is deemed valid.

Figure 1 shows the residual plot for the regression model presented in equation 1.

Figure 1: Residual Plot



However, when the model was tested, accuracy was very low under certain market conditions. When the liquid bond market volatility was high the accuracy decreased drastically, whereas when the market stayed stable the accuracy of the model increased. Therefore, more powerful correlation techniques (e.g. AI-based ones) would be required for the purposes of this project.

4. ANN-based Modelling Techniques

Artificial Neural Networks (ANNs) have been extensively applied and tested on complex problems of association, correlation and pattern recognition, and their ability to perform well on these tasks has been widely demonstrated (Rumelhart et al., 1986; Fahlmann, 1988; Hassoun, 1995; Mosca et al., 1997; Ren et al., 1998; Al-Dabass et al., 1997 and 1999). ANNs are also known to perform very well when modeling a single unknown function, even in complex cases where the actual function cannot be mathematically defined (Rumelhart et al., 1986; Fahlmann, 1988; Hassoun, 1995; Bruzzone and Orsoni, 2003; Saunders et al., 2004). Because patterns of human behaviour are expected to fall in this category of problems, where an analytical solution/representation cannot be produced, the research will rely on the associative capabilities of ANNs. A key strength of ANN-based models is their ability to extract patterns from historical data, so an important requirement for the implementation of the model is the availability of suitable data sets (Mosca et al., 1997). These include a training data set and a test set. The former is used by the ANNs to establish a correlation and the latter is used by the researchers to assess their ability to generalize (i.e. to provide valid output for input values not used for training purposes - Rumelhart et al., 1986). For this study, an ANN module was designed using an Excel-based application (see Appendix I for design details) to establish a correlation between the performance of trader A and the market conditions. The set of data describing the market conditions and the trader's performance were fed to the ANN as learning and test data sets (10% of the available data was used for testing purposes).

As indicated in Appendix I, the ANN design consists of a back-propagation network with one input layer, one hidden layer and one output layer architecture. The learning rate was experimentally set to 0.05 and the momentum was set to 0.5. For training purposes, the option to present the network with randomised input was not selected. The training mode was defined to be sequential starting with 100 training cycles. Finally, the validation set consisted of 10% randomly selected rows from the original set of examples.

The ANN design was manually tuned based on trial and error to improve the accuracy of its output. Different network architecture options, training options and validation set options had to be set and tested, prior to achieving an acceptable degree of accuracy for the prototype model.

For the trained network, the average error was 14%, which is satisfactory considering the impact of market volatility, however it is still somewhat high considering the intended use of the model (i.e. to provide performance guidelines to managers in the liquid bond trading desk). The error was however much lower than the one calculated for the regression model, and there is certainly scope for improvement with the availability of a more flexible development environment, capable among other things to implement more advanced algorithms. For the purposes of demonstrating the methodology the results are deemed satisfactory, and prone to further improvement: ultimately the study has demonstrated that traders' P&L can be estimated for any given market data conditions using ANN-based models. Estimates well matched reality trends under different values of volatility and market direction.

5. Conclusion and Future Work

This research has proposed a decision model based on ANNs to establish a relationship between the performance of a trader and the market conditions. Initial experiments based on a prototype ANN-model demonstrate the viability of the approach. Further work will look at using neural models (one for each trader on the team) as components of a DSS. The system will use the performance estimates for each trader under the current market conditions, to recommend management actions when setting a new policy for the portfolio. These include limiting the risk and/or the notional amount for the traders who are likely to perform the worst under the current market conditions.

References

Al-Dabass D., P. Vindlacheruvu, and D. J. Evans,1997. Parallelism in neural nets. International Journal of Parallel Algorithms and Applications 11(3-4) pp 169-185.

Bruzzone A.G. and A. Orsoni, 2003. AI and simulation-based techniques for the

assessment of supply chain logistic performance. In Proceedings of the IEEE Annual Simulation Symposium (ANSS 2003), Orlando, FL, March 30-April 3 pp. 154-164.

Fahlmann S. E. 1988. An empirical study to of learning speed in back-propagation networks", CMU Technical Report, CMU-CS88, June.Hassoun M.H. 1995. Fundamentals of Artificial Neural Networks, Cambridge MA: MIT Press.

Kutsurelis J.E. (1998) Forecasting Financial Markets Using Neural Networks: An Analysis of Methods and Accuracy, Naval Postgraduate School, Monterey, CA.

Mosca R., P. Giribone, A.G. Bruzzone, A. Orsoni and S. Sadowski, 1997. Evaluation and analysis by simulation of a production line model built with back-propagation neural networks. International Journal of Modelling and Simulation 17(2) pp.72-77.

Ren M., D. Al-Dabass, and D. Su, 1998. A three-layer hierarchy for off-line recognition of Chinese characters. International Journal of Computer Processing of Oriental Languages (CPOL), 11(4) June 1998, pp409-426.

Rumelhart D.E., G.E. Hinton and R.J. Williams 1986. Learning representations by back-propagating errors", Nature, 323(9), pp 533-536.

Saunders J., A. Battersby, and D. Al-Dabass, 2004. An intelligent systems model of the piano tuner", UKSIM2004, In proceedings of the UK Simulation Society Confrence, St Catherine's College, Oxford, 29 - 31 March, pp75-80.

(Appendix I follows)

APPENDIX I – ANN'S DESIGN PARAMETERS

Network ArchitectureOptions

Number of Inputs *(bewtween 2 and 50)*	**2**	Number of Outputs *(between 1 and 10)*	**1**
Number of Hidden Layers *(1 or 2)*	**1**	Hidden Layer sizes *(Maximum 20)*	**Hidden 1** 16 / **Hidden 2** 1
Learning parameter *(between 0 and 1)*	**0.05**	Initial Wt Range (0 +/- w): w =	**0.6**
Momentum *(between 0 and 1)*	**0.5**		

Training Options

Total #rows in your data *(Minimum 10)*	**191**	No. of Training cycles *(Maximum 500)*	**100**
Present Inputs in Random order while Training ?	**NO**	Training Mode *(Batch or Sequential)*	**Sequential**

Save Network weights **With least Training Error**

Training / Validation Set **Partition data into Training / Validation set**

If you want to partition, how do you want to select the Validation set ?
Please choose one option **1**
Please fill up the input necessary for the selected option

Option 1 : Randomly select	10%	**of data as Validation set** (between 1% and 50%)
Option 2: Use last	5	**rows of the data as validation set**

Save model in a separate workbook? **NO**

