A Comparison of Artificial Neural Network and Multinomial Logit Models in Predicting Mergers

Nilgun Fescioglu-Unver[a], Başak Tanyeri[b,*]

[a] *TOBB University of Economics and Technology, Department of Industrial Engineering, Ankara, TURKEY*

[b] *Bilkent University, Department of Business Administration, Ankara, TURKEY*

Abstract

A merger proposal occurs when a bidder firm offers to purchase the control rights in a target firm. Predicting who will propose merger bids (bidder candidacy) and who will receive merger bids (target candidacy) is important to measure the price impact of mergers. This study investigates the performance of artificial neural networks and multinomial logit models in predicting bidder and target candidacy. We use a comprehensive dataset that covers the years 1979 to 2004 and includes all deals with publicly listed bidders and targets. We find that both models perform similarly while predicting target and non-merger firms. The multinomial logit model performs slightly better in predicting bidder firms.

Key words: finance, mergers, artificial neural network models, multinomial logit models

1. Introduction

Merger announcements disclose the intent of bidder firms to purchase control rights in a target firm. One of the important merger-related questions is whether bidder and target candidacy is predictable. Hedge funds use investment strategies called 'merger arbitrage' that rely on the prediction of

*corresponding author

URL: basak@bilkent.edu.tr; tel:(90-312) 290 1871; fax:(90-312) 266 4958 (Başak Tanyeri)

bidder and target companies [1]. To understand the possible impact of merger arbitrage strategies, one needs to model and estimate the merger choice and completion. Furthermore, predictability may affect how researchers measure the price impact of mergers. If merger candidacy is predictable, bidder and target shares would reflect the impact of mergers prior to merger announcements. As a result, event study methods that calculate event returns around merger announcements may incorrectly measure the price impact of mergers (Cornett et al., 2011; Tanyeri, 2006). To correctly measure the price impact of mergers, it becomes important to model and estimate predictability. This study investigates the performance of artificial neural network model and multinomial logit model in predicting merger candidacy.

Artificial neural network models are well-suited for finance applications (such as bankruptcy prediction, credit risk assessment, and stock price prediction) since they can successfully represent non-linear relationships between input and output variables. This paper advances the literature on artificial neural networks and mergers by comparing how artificial neural networks and multinomial logistic regressions perform in predicting merger candidacy. This is the first study to use a comprehensive dataset covering the years 1979 to 2004 and all deals with publicly listed bidders and targets. Our sample covers 315,927 observations.

This paper is organized as follows: Section 2 outlines previous studies on merger prediction. Section 3 presents some of the finance area applications of artificial neural networks. Section 4 describes the proxies for merger motives used to model the merger choice, and introduces the multinomial logit models and artificial neural network models used to estimate merger candidacy. Section 5 compares the results obtained with multinomial logit and artificial neural network models. Section 6 concludes the paper.

2. Studies on Merger Candidacy

Previous merger studies use two approaches to model merger candidacy. First approach identifies a single firm characteristic that is used for classification of anticipated and unanticipated merger deals. Second approach

[1] Mitchell and Pulvino (2001) define merger arbitrage, or risk arbitrage, as an investment strategy that attempts to profit from the spread between target stock's price and the offer price for target shares. Merger arbitrageurs realize returns conditional on whether deals are successfully completed (Mitchell and Pulvino, 2001; Mitchell et al., 2007).

develops predictive models of merger candidacy that use multiple firm characteristics to classify merger deals. Using a single variable for classification of anticipated and unanticipated deals is easy to use and interpret. However, predictive models of merger candidacy provide a more comprehensive picture of merger candidacy as they can incorporate multiple variables that measure different merger motives.

The existence of an announced program of mergers (Shipper and Thompson, 1983; Malatesta and Thompson, 1985), merger frequency (Asquith et al., 1983; Loderer and Martin, 1990; Fuller et al., 2002; Ismail, 2005), time elapsed between mergers (Song and Walkling, 2000, 2005) are the variables that the first strand of papers used for classification of anticipated and unanticipated deals. The second strand of papers estimate models of merger candidacy using: logistic models (Dietrich and Sorenson, 1984; Palepu, 1986; Ambrose and Megginson, 1992; Akhigbe et al., 2004; Cremers et al., 2009), probit models (Eckbo et al., 1990; Cornett et al., 2011) and artificial neural network models (Sen and Gibbs, 1994).

3. Finance Applications of Artificial Neural Networks

Artificial Neural Networks (ANN) are mathematical modeling tools which perform complex function mappings (Hornik et al., 1989). ANNs successfully represent complicated and nonlinear relationships between several input and several output variables (Bishop, 1996). ANNs simulate the working principles of the human brain.

An ANN is composed of three layers of neurons. First layer is the *input layer* which has a number of neurons equal to the number of input variables. The third layer is the *output layer* which has the neurons that represent the output variables. The second layer resides between these two layers and is called as the *hidden layer*. The hidden layer can be composed of single or multiple layers. In each layer, there are several neurons. The neurons in the input layer are connected to the hidden layer neurons. Hidden layer neurons are connected to the neurons in the output layer through a network. Each of the links in this network has a weight. Training phase determines the weights of the links are determined using the training dataset. The resulting network represents the relationship between the input and output variables.

Finance applications of Artificial Neural Networks include bankruptcy prediction, credit risk assessment, stock-market and foreign exchange rate forecasting (Fadlalla and Lin, 2001; Huang et al., 2007; Fethia and Pasiourasb, 2010). Models of bankruptcy prediction determine the financial health or failure of a new firm using characteristics of their sample of failed and non-failed firms. Zhang et al. (1999) compared performances of ANNs and logistics regression in predicting bankruptcy. Odom and Sharda (1990); Perez (2006) modeled bankruptcy forecasting problem as a classification problem and predicted bankruptcy using ANNs. Shin et al. (2005) compared use of support vector machines and neural networks for bankruptcy prediction.

Credit-risk assessment models use current and historic information about a customer to determine whether he/she can be granted credit. Tsai et al. (2009) and Yeh and Lien (2009) compared the prediction capabilities of ANNs and other data mining techniques for the Taiwanese financial institutions. Trinkle and Baldwin (2007) developed ANN models and used them to generate first-order interpretable credit models that can explain the credit decision. Stock market and exchange rate forecasting models have been used to detect stock price manipulations (Ogut et al., 2009) and to optimize portfolio formation (Ko and Lin, 2008; Freitas et al., 2009).

The only study to use ANN for prediction of mergers is Sen and Gibbs (1994). Sen and Gibbs (1994) compared the performance of ANNs and logistic regression using 117 target-observations and 2,545 non-target observations covering the years from 1980 to 1985. While Sen and Gibbs (1994) predict only target and non-target candidacy, this paper develops models for both merger and bidder candidacy that rely on a larger set of independent variables. Furthermore, the sample used in this study covers the years from 1979 to 2004 with 5,207 bidder observations, 2,641 target observations, and 308,079 non-merger firm observations.

4. Research Method

This section develops models to predict the merger choice of firms. At any point of time, managers make a choice between three alternatives: (i) to propose a bid to attain control rights in another company (being a bidder), (ii) to solicit/receive bids for control rights in their company (being a target), (iii) to neither propose nor solicit bids (being a non-merger firm). Several variables can predict bidder and target candidacy. Section 4.1 explains these variables. Sections 4.2 and 4.3 develop the artificial neural network models and multinomial logit models to be used in estimation of merger candidacy.

4.1. Sampling Frame and Description of Variables

We follow the strategy of Tanyeri (2006) and Cornett et al. (2011) to construct the sample of merging and non-merging firms and to develop predictors of merger candidacy. The sample of merging firms (bidders and targets) are from Security Data Company's US Mergers and Acquisitions database and cover the period from 11/16/1977 to 12/30/2004. We restrict the merging sample to include those deals that show an intent to transfer control rights. To be included in the sample, bidders must hold less than fifty percent of outstanding target shares before the merger announcement and bidders must propose to hold more than fifty percent of outstanding target shares after the merger [2]. Sample firms are nonfinancial US enterprises due to the differences in regulatory environment and the lack of data availability for foreign and financial firms. We also require sample firms to be public companies.

We apply the sampling criteria used to construct the merging-firms sample to compile the sample of nonmerging firms. We compile a sample of US, nonfinancial firms using the CRSP-COMPUSTAT merged database. The sample covers 110 quarters from the third quarter of 1977 to fourth quarter of 2004. We map the merging sample onto the CRSP-COMPUSTAT data for identification of bidders, targets and non-merger firms. A firm-quarter is defined as: a bidder-quarter if the firm attempts at least one merger bid in the next financial-statement-release quarter, a target-quarter if the firm receives at least one bid in the next quarter, and a non-merging firm-quarter if the firm neither proposes nor receives any bids in the next quarter. We also require that the firms have non-missing data for variable construction and windsorize the variables at the 1st and 99th percentiles to reduce the effect of outliers. These filters produce 2,530 firms proposing 5,400 bids in 5,207 firm quarters, 2,352 firms receiving 2,706 bids in 2,641 firm quarters, and 11,010 firms neither proposing nor receiving bids in 308,079 firm quarters.

Table 1 summarizes the data set used in this study. First rows average the book value of assets (in million dollars) of bidder, target, and non-merger firms in each year. Second rows list the number of bidder, target, and non-merger firms in each year. The second half of the sample (covering the years 1991 to 2004) is richer than the first half (covering the years 1979 to 1990) in terms of merging firms. There are, on average 262 bidders and 125 targets

[2] We use the SDC variables 'menumain' and 'formc' to identify the deals that show an intent to transfer control rights.

per year in the second half and 128 bidders and 74 targets per year in the first half. Bidders prove largest (on average 3,538 million dollars) in terms of book value of assets. Non-merging firms (on average 962 million dollars) are larger than targets (on average 1,421 million dollars). The size distribution indicates that the larger sample firms buy the smaller firms.

We review theories on merger motives to develop predictors for merger candidacy. Theoretical models establish that managers may engage in mergers to generate shareholder value and/or to protect opportunistic benefits that managerial positions enable them to enjoy. Managers may create shareholder value by: (i) increasing efficiency of human and financial capital; (ii) attaining economies of scale and scope; and (iii) increasing market power (Gort, 1969; Holmes and Schmitz, 1995; Fluck and Lynch, 1999; Jovanovic and Rousseau, 2002). Incentive conflicts between managers and shareholders may also lead to mergers when opportunistic managers focus on generating value for themselves at the expense of shareholders (Jensen and Meckling, 1976; Jensen, 1986; Datta et al., 2001; Shleifer and Vishny, 2003; Hartzell et al., 2004; Rhoades-Kropf and Viswanathan, 2004; Jensen, 2005).

Eight variables [3], namely *sales shock, square of sales shock, asset size, asset growth, sales growth, concentration ratio, resource-growth mismatch, and return on assets (ROA)*, proxy for merger motives to generate shareholder value. *Sales shock* (defined as the absolute value of the difference between the two-year median industry [4] sales growth rate and the two-year median sales growth rate for all firms listed in our sample) is a proxy for economic disturbances that may motivate mergers (Gort, 1969; Maksimovic and Phillips, 2001; Andrade et al., 2001). The *square of sales shock* allows for non-linearity in the sales shock variable. The proxies for the desire to reduce costs by increasing economies of scale and scope through mergers are *asset size* (defined as the log of total assets), *asset growth* (defined as the two-year growth rate of assets), *sales growth* (defined as the two-year growth rate in sales) (Gort, 1969; Palepu, 1986; Ambrose and Megginson, 1992; Moeller et al., 2005). The ease of entry and exit into the industry may predict merger candidacy and is measured by *concentration ratio* (defined as the sum of sales of the largest four firms (in terms of sales) divided by total

[3]Interested readers may refer to Cornett et al. (2011) and Tanyeri (2006) for the definitions and in-depth discussions about the variables used in this study

[4]Each industry covers all firms with the same two-digit SIC code. The one-digit SIC code is used when there are fewer than five firms in an industry.

industry sales) (Gort, 1969; Eckbo et al., 1990). Firms with a mismatch in capital resources growth opportunities may engage in mergers. We measure this mismatch with the *resource-growth mismatch* indicator (defined as an indicator which takes on the value one (zero) if the two-year sales growth is larger (smaller) than the industry median and the ratio of long-term debt to total assets is lower (higher) than the industry median) (Palepu, 1986; Ambrose and Megginson, 1992; Fluck and Lynch, 1999). The match quality between bidders and targets may also predict candidacy. We use *Return on Assets (ROA)* (defined as the book value of net income before extraordinary items divided by total assets) as a proxy for match quality (Lang et al., 1989; Palepu, 1986; Ambrose and Megginson, 1992; Maksimovic and Phillips, 2001; Akhigbe et al., 2004).

Three variables, namely *cash ratio, prior mergers, industry mergers*, measure managerial motives to protect opportunistic benefits through mergers. Large cash reserves as measured in *cash ratio* (defined as cash and marketable securities divided by total assets) enable managers to pursue personal benefits and as such would enable managers to propose an empire-building merger and would motivate managers to desist takeovers. The *prior mergers* variable (defined as the number of times a firm proposes or receives a merger bid in the prior two years excluding the current bid) accounts for the empire-building motives of managers (Shipper and Thompson, 1983; Malatesta and Thompson, 1985; Asquith et al., 1983; Loderer and Martin, 1990; Holmes and Schmitz, 1995; Fuller et al., 2002; Ismail, 2005). Clustering of mergers across time and industry is well-documented. Mergers may be motivated by a desire to avoid risk by joining the herd. We measure merger clustering in time and industry using *industry mergers* (defined as the number of industry firms that made or received a bid divided by the total number of industry firms; this ratio is cumulated for the past two years).

Mispricing of shares may affect investment decisions; hence merger decisions (Myers and Majluf, 1984). Two alternative hypotheses exist on whether managers use their private information about mispriced shares to act in the best interests of shareholders or to protect opportunistic benefits. Hansen (1987); Rhoades-Kropf and Viswanathan (2004); Eckbo et al. (1990) state that managerial beliefs about stock overvaluation may motivate stock-financed mergers. These mergers intend to generate long run value for pre-merger shareholders at the expense of post-merger shareholders. Jensen (2005) states that managerial beliefs about stock overvaluation may motivate mergers financed with overvalued equity when managers want to generate and/or protect opportunistic benefits. Three variables, share turnover (defined as the number of shares of stock traded divided by the total shares outstanding), price run-up (defined as the two-year change in stock price) and information asymmetry (defined as an indicator that is one if the market-to-book value [5] is higher than the industry median and the firm's share turnover is lower than its industry median), are proxies for managerial motives to take advantage of its information advantage.

4.2. Artificial Neural Networks Model

Performance of an ANN depends on the number of hidden layers, number of neurons in each hidden layer, training function and the data used to train the network. The sample has a significant imbalance in the target, bidder and non-merger classes. The sample of 315,927 firm-quarters consists of 2641 target-quarters, 5207 bidder-quarters, and 308,079 non-merger-quarters. The number of non-merger quarters is almost 60 times more than the number of bidder quarters and 117 times more than the target quarters. This kind of imbalance has negative effects on the performance of learning algorithms which assume a balanced class distribution (Liu et al., 2008). There are several methods to overcome these negative effects. In this study, we use the under-sampling method. This method under-samples the class with the highest number of elements to the size of the minimum class (Liu et al., 2008). This study separates the data into two sets: training and validation, and test sets. Then, this study balances the training and validation set by under-sampling, and reduces the size of non-merger quarters to the size of the target quarters.

The input layer of the ANN model has 13 nodes, one node for each of the following variables: *sales shock*, *asset size* [6], *asset growth*, *sales growth*, *four-firm four-firm concentration ratio*, *resource-growth mismatch*, *return on assets*, *cash ratio*, *prior mergers*, *industry mergers*, *share turnover*, *price run-up*, *information asymmetry*. The definitions of these variables are given in Section 4.1. The output layer contains three neurons, one for target, one for bidder and one for non-merger firm indicators. The ANN model has a

[5] Market-to-book ratio is the ratio of the closing price of the firm's common stock multiplied by the number of common shares outstanding to the book value of stockholder's equity.

[6] The asset size variable is log of assets in the multinomial logit model and is the actual million dollar value in ANN model.

single hidden layer. We used MATLAB to implement the ANN model. The network is a feed-forward backpropagation network with tan-sigmoid transfer function for hidden layer and linear transfer function for the output layer. Network is trained with the Levenberg-Marquardt backpropagation method. The results of the artificial neural network model is presented in Section 5.1.

4.3. Multinomial Logit Model

The multinomial logit model uses the under-sampled data, as described in section 4.2. Equation 1 computes the probability of being a target, bidder or non-merger firm. Using the variables outlined in Section 4.1 (denoted X), STATA estimates β, the variable coefficients, and computes the probability of each merger choice for all sample firms in each quarter. Let Y_i denote the denote the choice of firm i and $Prob(Y_i = j)$ denote the probability that firm i chooses choice j ($j = 0, 1, 2$):

$$Y_i = \begin{cases} 0: & \text{firm neither proposes nor receives a bid in quarter } t+1 \\ 1: & \text{firm proposes a bid in quarter } t+1 \\ 2: & \text{firm solicits or receives a bid in quarter } t+1 \end{cases}$$

$$Prob(Y_i = j) = \frac{e^{\beta_j' X_i}}{1 + \sum_{j=1}^{2} e^{\beta_j' X_i}} \tag{1}$$

and

$$Prob(Y_i = 0) = \frac{1}{1 + \sum_{j=1}^{2} e^{\beta_j' X_i}}$$

A firm in any quarter is identified as a bidder if the probability of proposing a bid is greater than the probabilities of receiving a bid and of not engaging in merger activity. Similarly, a firm in any quarter is identified as a target (non-merger) if the probability of soliciting a bid (not engaging in merger activity) is greater than the probabilities of proposing a bid and of not engaging in merger activity (receiving a bid). The results of the multinominal logit model is presented in Section 5.2.

5. Results

We use 10-fold cross validation to estimate the performance of the ANN and Multinomial Logit Models. For *m-fold* cross validation, we separate the data into m-pieces and train/estimate the models on the $(m-1)$ pieces of data and test on the i^{th} piece. Later, we combine the correct estimation percentages on the m pieces (Liu, 2006). This yields the average correct estimation percentage of all the m models. Correct estimation percentage is calculated as $100 * NumModelDetectedAsClass_i/NumTotalClass_i$ where $NumModelDetectedAsClass_i$ is the number of cases the model detect correctly as belonging to class i and $NumTotalClass_i$ is the total number of cases belonging to class i.

5.1. ANN Model Results

We compute the correct estimation percentages of the ANN model for different number of neurons in the hidden layer. Table 2 presents the average results of the 10 fold cross validation on the test dataset. The first column presents the number of nodes in the hidden layer and the following columns present target, bidder, non-merger and overall correct detection accuracy (in percentages). Models show different performances in detecting target and bidder firms. Target detection accuracy vary between 32.46 and 40.25, bidder detection accuracy varies between 43.76 and 51.83 and non-merger firm detection accuracy varies between 49.74 and 53.74. Table 2 shows that the ANN model with 10 nodes in the hidden layer performs better than the other models in terms of overall correct detection percentage.

Table 3 represents the classification percentages for ANN model with 10 nodes in the single hidden layer. The model correctly identifies target, bidder and non-merger firms with 40.25, 45.21 and 53.24 percent accuracy respectively. The highest accuracy is for non-merger firms and the lowest accuracy is for target firms.

5.2. Multinomial Logit Results

Table 4 presents the average results of multinomial logit estimations of the 10 fold cross validation. Multinomial logit regressions estimate the probability of a firm proposing a bid, soliciting a bid, and neither proposing nor receiving a bid in the next quarter in the 10% of the data designated as test data in each validation fold. The rows of Table 4 are the real identities of the observations (bidder, target, non-merger) and the columns are the estimated identities of the observations. The model correctly identifies target, bidder and non-merger firms with 40.32, 52.20 and 55.77 percent accuracy respectively. Similar to results of ANN model, the highest accuracy is for non-merger firms and the lowest accuracy is for target firms.

6. Conclusion

This paper compares the performance of artificial neural networks and multinomial logistic models in predicting merger candidacy. Both models perform similarly while predicting target and non-merger firms. The multinomial logit model performs slightly better in predicting bidder firms.

Multinomial logit models yield coefficient estimates for the variables that have economic meaning. ANN models work as a blackbox and do not automatically reveal coefficient estimates. However, there exist different rule-extraction techniques which can interpret the ANN weights, and explain the relationship between the input and output variables (Taha and Ghosh, 1999).

Multinomial logit models estimate linear models. ANN model handles non-linear relationships between the independent and dependent variables. The ANN model is also powerful in handling large number of input variables and variables with interactions among each others.

This paper uses the explanatory variables that Tanyeri (2006) and Cornett et al. (2011) construct, to estimate merger candidacy. There may exist other independent variables which can be extracted from financial statements. However, logit models are restricted (by degrees of freedom) in the number of variables they can handle whereas ANN models are not as restricted. Future studies will investigate the predictive power of ANN and multinomial logit models when the number of independent variables and the complexity of the problem increase.

References

Akhigbe, A., Madura, J., Whyte, A. M., 2004. Partial anticipation and gains to bank merger targets. Journal of Financial Services 25, 55–71.

Ambrose, B., Megginson, W., 1992. The role of asset structure, ownership structure, and takeover defenses in determining acquisition likelihood. Journal of Financial and Quantitative Analysis 27, 575–589.

Andrade, G., Mitchell, M., Stafford, E., 2001. New evidence and perspectives on mergers. Journal of Economic Perspectives 15, 103–120.

Asquith, P., Bruner, R. F., Mullins, D. V., 1983. The gains to bidding firms from merger. Journal of Financial Economics 11, 121–139.

Bishop, C., 1996. Neural Networks for Pattern Recognition. Oxford University Press.

Cornett, M. M., Tanyeri, B., Tehranian, H., 2011. The effect of merger anticipation on bidder and target firm announcement period returns. Journal of Corporate Finance.
URL `http://dx.doi.org/10.1016/j.jcorpfin.2010.10.004`

Cremers, K., Nair, V., John, K., 2009. Takeovers and cross-section of returns. Review of Financial Studies 22, 1409–1445.

Datta, S., Iskandar-Datta, M., Raman, K., 2001. Executive compensation and corporate acquisition decisions. Journal of Finance 56, 2299–2333.

Dietrich, J., Sorenson, E., 1984. An application of logit analysis to prediction of merger targets. Journal of Business Research 12, 393–412.

Eckbo, E. B., Maksimovic, V., Williams, J., 1990. Consistent estimation of cross-sectional models in event studies. Review of Financial Studies 3, 343–365.

Fadlalla, A., Lin, C., 2001. An analysis of the applications of neural networks in finance. Interfaces 31 (4), 112–122.

Fethia, M. D., Pasiourasb, F., 2010. Assessing bank efficiency and performance with operational research and artificial intelligence techniques: A survey. European Journal of Operational Research 204, 189–198.

Fluck, Z., Lynch, A., 1999. Why do firms merge and then divest? a theory of financial synergy. Journal of Business 72, 319–346.

Freitas, F. D., DeSouza, A. F., de Almeida, A. R., 2009. Prediction-based portfolio optimization model using neural networks. Neurocomputing 72, 2155–2170.

Fuller, K., Netter, J., Stagemoller, M., 2002. What do returns to acquiring firms tell us? evidence from firms that make many acquisitions? Journal of Finance 57, 1763–1793.

Gort, M., 1969. An economic disturbance theory of mergers. Quarterly Journal of Economics 83, 624–642.

Hansen, R., 1987. A theory for the exchange medium in mergers and acquisitions. Journal of Business 60, 75–95.

Hartzell, J., Ofek, E., Yermack, D., 2004. What's in it for me? ceos whose firms are acquired. Review of Financial Studies 17, 37–61.

Holmes, T., Schmitz, J., 1995. On the turnover of business firms and business managers. Journal of Political Economy 103, 1005–1038.

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Networks 2 (5), 359–366.

Huang, W., Lai, K. K., Nanamoru, Y., Shouyang, W., Yu, L., 2007. Neural networks in finance and economics forecasting. International Journal of Information Technology & Decision Making 6 (1), 113–140.

Ismail, A., 2005. Will multiple acquirers ever learn? the us evidence from single versus multiple acquirers? Working Paper, American University of Beirut.

Jensen, M., Meckling, W., 1976. Theory of the firm: Managerial behavior, agency costs, and ownership structure,. Journal of Financial Economics 3, 305–360.

Jensen, M. C., 1986. Costs of free cash flow, corporate finance, and takeovers. American Economic Review 76, 323–329.

Jensen, M. C., 2005. Agency costs of overvalued equity. Financial Management (34), 5–15.

Jovanovic, B., Rousseau, P., 2002. The q-theory of mergers. American Economic Review 92, 198–204.

Ko, P.-C., Lin, P.-C., 2008. Resource allocation neural network in portfolio selection. Expert Systems with Applications 35, 330–337.

Lang, L., Stulz, R., Walkling, R., 1989. Managerial performance, tobin's q and the gains from successful tender offers. Journal of Financial Economics 24, 137.

Liu, J., Hu, Q., Yu, D., 2008. A comparative study on rough set based class imbalance learning. Knowledge-Based Systems 21, 753–763.

Liu, Y., 2006. Create stable neural networks by cross-validation. In: International Joint Conference on Neural Networks.

Loderer, C., Martin, K., 1990. Corporate acquisitions by listed firms. Financial Management 19, 17–33.

Maksimovic, V., Phillips, G., 2001. The market for corporate assets: Who engages in mergers and asset sales and are there efficiency gains? Journal of Finance 56, 2019–2065.

Malatesta, P., Thompson, R. A., 1985. Partially anticipated events. Journal of Financial Economics 14, 237–250.

Mitchell, M., Pedersen, L. H., Pulvino, T., 2007. Slow moving capital. American Economic Review 97, 215–220.

Mitchell, M., Pulvino, T., 2001. Characteristics of risk and return in risk arbitrage. Journal of Finance 56, 2135–2175.

Moeller, S., Schlingenmann, F., Stulz, R., 2005. Wealth destruction on a massive scale: A study of acquiring-firm returns in the recent merger wave. Journal of Finance 60, 757–782.

Myers, S., Majluf, N., 1984. Corporate financing and investment decisions when firms have information that investors do not have. Journal of Financial Economics, 187–221.

Odom, M., Sharda, R., 1990. A neural network model for bankruptcy prediction. In: Proceedings of the IEEE International Conference on Neural Networks. pp. 163–168.

Ogut, H., Doganay, M. M., Aktas, R., 2009. Detecting stock-price manipulation in an emerging market: The case of turkey. Expert Systems with Applications 36, 11944–11949.

Palepu, K. G., 1986. Predicting takeover targets. Journal of Accounting and Economics 8, 3–35.

Perez, M., 2006. Artificial neural networks and bankruptcy forecasting: a state of the art. Neural Computing & Applications 15, 154–163.

Rhoades-Kropf, M., Viswanathan, S., 2004. Market valuation and merger waves. Journal of Finance 59, 2685–2718.

Sen, T. K., Gibbs, A. M., 1994. An evaluation of the corporate takeover model using neural networks. Intelligent Systems in Accounting, Finance and Management 3, 279–292.

Shin, K.-S., Lee, T. S., Kim, H.-j., 2005. An application of support vector machines in bankruptcy prediction model. Expert Systems with Applications 28, 127–135.

Shipper, K., Thompson, R. A., 1983. Evidence on the capitalized value of merger activity for acquiring firms. Journal of Financial Economics 11, 85–119.

Shleifer, A., Vishny, R. W., 2003. Stock market driven acquisitions. Journal of Financial Economics 70, 295–311.

Song, M. H., Walkling, R. A., 2000. Abnormal returns to rivals of acquisition targets: A test of the 'acquisition probability hypothesis'. Journal of Financial Economics 55, 143–171.

Song, M. H., Walkling, R. A., 2005. Anticipation, acquisitions and bidder returns. Working Paper, FMA European Meeting, Siena.

Taha, I. T., Ghosh, J., 1999. Symbolic interpretation of artificial neural networks. IEEE Transactions on Knowledge and Data Engineering 11, 448–464.

Tanyeri, B., 2006. Unpleasant shocks or welcome surprisees? what information is conveyed in merger announcements. Ph.D. thesis, Boston College.

Trinkle, B. S., Baldwin, A. A., 2007. Interpretable credit model development via artificial neural networks. Intelligent Systems In Accounting Finance and Management 15, 123–147.

Tsai, M.-C., Lin, S.-P., Cheng, C.-C., Lin, Y.-P. L., 2009. The consumer loan default predicting model - an application of dea-da and neural network. Expert Systems with Applications 36, 11682–11690.

Yeh, I.-C., Lien, C.-h., 2009. The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. Expert Systems with Applications 36, 2473–2480.

Zhang, G., Hu, M., Patuwo, B., Indro, D., 1999. Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. European Journal of Operational Research 116, 16–32.

Table 1: Asset size and distribution of bidders, targets, and non-merger firms across years

year	Bidder	Target	Non-merger
1979	1,992	869	806
	5	3	3,931
1980	2,831	573	870
	40	24	7,553
1981	1,510	396	951
	105	53	6,880
1982	1,205	500	1,080
	176	41	7,010
1983	1,136	523	1,131
	261	55	7,791
1984	1,527	761	804
	264	85	11,034
1985	2,150	1,025	783
	83	101	12,196
1986	2,446	380	860
	116	124	12,430
1987	4,294	639	917
	112	114	12,045
1988	5,052	1,179	1,023
	109	116	12,186
1989	4,064	529	1,075
	155	102	12,702
1990	3,518	2,342	1,173
	114	74	12,644
1991	1,981	251	1,268
	180	61	12,440
1992	3,060	609	1,313
	183	57	12,468
1993	2,248	453	1,393
	187	84	12,699
1994	2,411	536	1,351
	234	98	13,433
1995	2,248	581	1,347
	330	143	14,068
1996	4,197	1,153	1,333
	343	139	14,864
1997	3,084	874	1,455
	374	216	15,302
1998	3,226	1,251	1,552
	361	228	15,488
1999	5,558	1,431	1,714
	355	249	15,034
2000	8,553	1,406	2,029
	280	157	13,925
2001	5,373	1,756	2,384
	218	106	13,114
2002	7,712	1,456	2,443

Table 2: Classification accuracy for ANN models with different number of nodes

	Correct Detection Percentage			
# nodes	Target	Bidder	Non-merger	Overall
7	37.78	47.75	52.95	46.16
10	40.25	45.21	53.54	46.33
15	32.46	51.83	53.82	46.03
20	40.17	48.31	49.74	46.07
25	38.61	43.76	53.74	45.37

Table 3: Classification accuracy for ANN model with 10 nodes

	Correct Detection Percentage		
Identity Estimate	Target	Bidder	Non-merger
Target	40.25	22.48	37.28
Bidder	24.92	45.21	29.86
Non-merger	28.29	18.17	53.54

Table 4: Classification accuracy for Multinomial Logit Model

Identity Estimate	Target	Bidder	Non-merger
Target	40.32	22.54	37.14
Bidder	19.91	52.20	27.89
Non-merger	24.29	19.94	55.77