Neural Networks: An Interdisciplinary Tax Research Methodology

Tina S. Quinn
Arkansas State University

Tonya K. Flesher
University of Mississippi

John D. Johnson
Utah State University

Flyn L. Flesher
Ogletree Deakins

Case-based research has been an important approach to evaluate models of tax law. One area of tax law that has been controversial for decades is the worker classification issue. This study uses artificial intelligence methods to construct a model to help resolve this controversy. Another purpose is to evaluate the performance of neural networks in case-based research. Models were developed using step-wise multiple regression, step-wise logistic regression, discriminant analysis and an artificial neural network (ANN).The ANN model is shown to outperform these other models with its predictive ability. This result points to the importance of ANN in case-based research.

Worker classification for federal tax purposes (both income and employment tax) has been a controversial area for decades. Distinguishing between an employee and independent contractor may be difficult, yet necessary in order to comply with tax laws and regulatory agency requirements. Previous studies concerning this problem fail to account for possible nonlinear and cross product effects that are expected in the data. The purpose of this study is to use artificial intelligence methods to construct a model that will help resolve this controversy and to demonstrate the effectiveness of these methods. Specifically artificial neural networks (ANN), which account for possible cross product effects and provide a flexible nonlinear functional form, are used.

Since 1988, the Internal Revenue Service (IRS) has conducted a high priority campaign to address worker classification issues, with revenue officers performing special audits to uncover misclassification. By 1996, this aggressive enforcement program resulted in 12,983 employment tax audits and the reclassification of 527,000 workers in *Vizcaino v. Microsoft Corp*. 97 F3d 1187, 1202 (CA-9, 1996). As Microsoft learned, the consequences of reclassification can be more pervasive than the amount of assessed taxes, plus any penalty and interest. After the IRS reclassified Microsoft's "freelancers" from independent contractors to employees, several such 'freelancers' sought various employee benefits from Microsoft. Circuit Judge Trott (in his dissent) laments that the IRS's tough enforcement policy not only collects more money for the government, but it has unforeseen consequences of forcing employers to extend retroactive benefits to workers for which they did not contract. Judge Trott sees the effects of reclassification spilling over from the employment tax issue to affect other contracts between companies and workers (1200).

There have been many attempts at clarifying the issue, but neither the Congress, the Treasury Department, nor the IRS have adequately defined employee. The IRS has, however, initiated programs to help alleviate the problem of worker classification (Breault et al. 1997). In 1996, the IRS issued new *Worker Classification Training Materials*, and began a Classification Settlement Program (CSP). The CSP allows taxpayers and tax examiners to resolve worker classification cases as early in the administrative process as possible, reducing taxpayer burden. In 2011, the IRS started the Voluntary Classification Settlement Program (VCSP) that allows eligible taxpayers to obtain relief similar to that available through CSP to taxpayers under audit. VCSP is a voluntary program that allows taxpayers to reclassify their workers as employees for future tax periods with partial relief from federal employment taxes. These attempts and frequent Congressional proposals at resolving the worker classification issue are evidence of this area's long and complex history.

BACKGROUND

The crux of the problem of worker classification is the vagueness inherent in the definition of "common law employee". Section 3121(d) of the Internal Revenue Code (IRC) contains four separate and independent categories of employee. One such category includes employees under the common law rules (rules that have evolved over time through court decisions and custom). In deciding the issue of worker classification, the courts have looked for evidence of control. The Employment Tax Regulations (Treas. Reg. '31.3121) provide that generally a worker is an employee when the principal has the right to control and direct the worker. It is not necessary that control is actually exercised, merely that the right to control exists. Control, however, is an abstract concept.

Revenue Ruling 87-41 (1987-1 C.B. 296) was issued in an effort to clarify the issue of control and sets forth twenty factors that are indicative of presence of control. The ruling gives no guidance as to how many factors must be present or how much weight is given to each factor. These twenty factors are as follows:

1) Instructions.
2) Training
3) Integration.
4) Services rendered personally.
5) Hiring, supervision, and paying assistants.
6) Continuing relationship.
7) Set hours of work.
8) Full time required.
9) Doing work on employer's premises.
10) Order or sequence set.
11) Oral or written reports.
12) Payment by hour, week, or month.
13) Payment of business and/or travel expenses.
14) Furnishing of tools and materials.
15) Significant investment.
16) Realization of profit or loss.
17) Working for more than one firm at a time.
18) Making service available to general public.
19) Right to discharge.
20) Right to terminate.

A very popular method for conducting tax research in the 1970s and 1980s was macro-case analysis. This method flourished and then died. New computer technology provides hope for a resurrection. Macro-case analysis has been defined as the "aggregate analysis of court decisions rendered in an area over a chosen time period" (Misiewicz, 1977, p. 935). Macro-case analysis in tax research may be undertaken for tax compliance purposes or tax planning purposes. Given an issue, the researcher must determine the appropriate court or courts to research and the relevant time span. The relevant cases must be located and the factors to use in the analysis must be determined. The factors may be determined by a review of: (1) court cases; (2) the Internal Revenue Code; (3) Treasury Regulations; (4) IRS Rulings; or (5) other sources. Kramer (1984) describes this technique as quantitative case analysis[1], which he contrasts with traditional qualitative case analysis. The important variables "are statistically inferred from the historical relationships that have existed between a series of facts and judgments in a series of cases decided by the courts on the same issue" (Kramer, 1984, p.20).

Kort (1977) discussed research in political science that analyzed judicial decision making using statistical techniques such as factor analysis, regression and discriminant analysis. Several of these studies dated back to the early 1960s. The first attempt to use quantitative methods as a predictor of judicial decisions was by Kort (1957).

Misiewicz (1977) discussed early examples of accounting research studies using macro-case analysis where non-statistical analyses such as frequency of criteria application and a descriptive matrix were employed. He also suggested that Englebrecht's (1976) dissertation, which used regression models to determine the significance of variables in Tax Court decisions, was the first macro-case analysis dissertation completed in accounting.

Kramer (1984) identified several purposes for statistical analysis of court cases:

1. To determine if the variables contained in the guidelines are, in fact, used by the court in its decision-making process;
2. To determine if the variables receive different weighting in the model;
3. To determine if additional variables (over and above those included in the original guidelines) are important in the decision-making process; and
4. To determine if the guidelines can be simplified and reduced to a smaller number of variables (for example, a subset of the original variables, or a subset of the original variables plus some additional variables) without losing predictive ability.

Kramer (1984, p.21) offered the following critique: "research efforts in this area have proceeded with greater breadth than depth. Researchers have tended to be more interested in using the existing methodology to examine additional problems instead of refining the methodological techniques." Recent advances in computer techniques have made it possible to use more sophisticated techniques.

The first accounting dissertation to use macro-case analysis in tax was completed in 1976. Every year until 1989 at least one dissertation in accounting utilized this method. Following this peak in the 1980s, interest in this type of research disappeared. This sharp decline probably can be explained by examining Kramer's (1984) critique. The research had more breadth than depth. Researchers failed to refine the method and used the existing statistical methods until the list of easily quantifiable topics was exhausted. The results of this study may provide motivation for a revival.

PURPOSE OF STUDY

The problem of worker classification is to resolve the apparent inconsistencies as to which variables courts consider significant. If workers are initially classified correctly, employers can avoid the consequences of reclassification by the IRS. These consequences are more pervasive than the amount of the assessed taxes, plus any penalty and interest. Reclassification can cause violation of the requirements for a qualified pension, profit-sharing or stock bonus plan as well as affect other employee fringe benefit plans. Additionally, there are tax ramifications for workers who are reclassified as employees and lose tax deductions (Burns, 1996, p. 102). This long history of controversy over worker classification with its resultant broad and costly impact has produced unreasonable compliance costs for taxpayer and the government

Case law seems to indicate inconsistencies in the application of the law with respect to the worker classification issue. However, previous case-based research studies concerning this problem fail to take into account possible nonlinear and cross product effects that exist in the data. Therefore one must ask: are the inconsistencies due to the case research methodologies or true inconsistencies in the law? In other words, are there potential nonlinearities or cross product effects that have not been identified? The purpose of this study is to use artificial intelligence methods to construct a model that might lead to a new area of inquiry that will help to resolve this controversy and to show the usefulness of these methods in other tax and legal issues.

In this study, models are developed using step-wise multiple regression, step-wise logistic regression, discriminant analysis and an artificial neural network (ANN) trained by the Genetic Adaptive Neural Network Training (GANNT) algorithm. Multiple regression models are inappropriate due to the limited dependent nature of the models. Logit models resolve this problem and are the most commonly used methodology for this problem. The logit model is a special case of the ANN (when a logistic squashing function is used). However, ANN accounts for possible cross product effects and provides a flexible nonlinear functional form. Thus a superior fit using ANN provides a possible indicator that the cross product or nonlinear effects are important. Previous studies using logit look at limited interaction effects but have not accounted for nonlinearities. The ANN model is shown to outperform these other models with its predictive ability. This provides a basis to indicate that nonlinearities and cross product effects may be crucial and highlights the further importance of ANN in case-based research. ANN represents a class of nonlinear statistical models whose mode of information processing is generally cast in terms of the functioning of the human brain. More specifically the ANN used in this study is a supervised feedforward network referred to as a multilayered perceptron.

LITERATURE REVIEW

Worker Classification

Dave N. Stewart conducted an extensive study of the employee/independent contractor issue (1980). Stewart analyzed 148 tax cases litigated in the District Court and Court of Claims from 1940 through 1979 to build a mathematical model to identify the variables used by the federal courts in deciding this issue. Discriminant analysis, ordinary least squares (OLS) regression analysis, and a non-linear logit analysis were all used to identify the relevant variables deemed the most important by federal judges. The non-linear logit procedure resulted in a five-variable model that correctly classified 97.3 percent of the cases in sample. The five variables in the model were Profit or Loss, Supervision, Independent Trade, Permanent Relationship, and Integration. The discriminant analysis resulted in a six-variable model that correctly classified 96.6 percent of the cases. The discriminant model contained the same five variables as the logit plus the additional variable of Hiring of Assistants. The OLS step-wise procedure produced a seven-variable model that correctly classified 95.3 percent of the cases. This model contained the same six variables as the discriminant model plus a seventh variable, Controlling the Place of Work.

Robinson and Hulen (1996) performed a logit analysis on 321 determination letters issued by the IRS in fiscal year 1990. The authors coded the factors in the rulings based upon the questions contained in Form SS-8 (Determination of Employee Work Status for Purposes of Federal Employment Taxes and Income Tax Withholding). The Form SS-8 divides the common law factors into parts and sub-parts, resulting in 88 variables. The authors found seven variables to be significant in modeling IRS decisions. The variables included Control and Supervision; Direction and Methods; Routine or Schedule; Type of Pay; Holding Out to the Public; Prohibition Against Competing with the Firm; and Filing Reports. These variables correctly modeled IRS decisions for 97 percent of the rulings.

Martindale and Price (Working Paper) performed a logistic regression on 168 court cases to determine if there exists a subset of the IRS 20 factors that can be used to distinguish between employees and independent contractors. A forward step-wise procedure produced a seven-variable model that correctly classified 97.62 percent of the cases. The variables were Instructions, Set Hours, Mode of Payment, Investment, Risk (profit or loss), Integration, and Continuing Relationship.

Cushing and Arguea (1999) used logistic regression to estimate a model using data obtained from Private Letter Rulings issued by the IRS from 1988 through 1993. The authors used the IRS's 20 factors as independent variables and the resulting model had an overall prediction accuracy rate of 98.5 percent that dropped to 91.4 percent when a holdout sample was used. A neural network model produced similar results, but the logistic regression model was adopted because it was more accurate on the holdout sample. The final model was estimated using a stepwise procedure and five of the 20 factors entered the model. The five factors are: Set working hours; Required reports; Employer provided tools and materials; Profit or loss; and Services available to the public.

O'Neil and Nelsestuen (1993) conducted a simple, non-statistical study involving the analysis of letter rulings issued in response to inquiries made by eleven workers of one firm during 1991. The purpose of the analysis was to discern patterns in the criteria mentioned in the determinations and whether those patterns were associated with observable characteristics of the job in question. The only "pattern" found was that the nine independent contractor classifications were based solely upon the 20-factor test in Rev. Rul. 87-41 while the classification of two workers as employees was based primarily on pre-1987 revenue rulings. The authors suggested that the factors applicable to a worker's case might be made to "fit" into the facts of a previous letter ruling at the discretion of the IRS.

Empirical research in the employer-employee relationship area is limited; however, numerous articles have been published in accounting journals in an attempt to clarify the problem. These articles are descriptive in nature, giving an overview of the problem and developing guidelines to determine employee vs. independent contractor status. This study contributes to the empirical research.

Neural Networks

Zelezny (1999) describes ANNs and compares the models to statistical techniques and expert systems. He notes that linear regression models assume a linear relationship and no interaction between the input variables, assumptions that may often be violated. ANN, however, can represent linear or non-linear relationships and inherently include interaction effects of the input. No *a priori* assumption as to functional form of the relationship is necessary when using ANNs. Zelezny concludes that ANNs can be useful tools for both the internal and external auditor.

Ramamoorti, et al, investigate whether neural networks can help enhance the internal auditor’s risk assessment process. The auditor may be faced with large amounts of both qualitative and quantitative data, making risk assessment both complex and difficult. The authors used data consisting of qualitative and quantitative risk factor information about the academic and administrative departments at the University of Illinois at Chicago. A Delphi study was conducted using experienced internal auditors who used the risk factor information to assign risk rankings to 141 departmental units. These risk factor rankings were the target output values for the neural network models. The authors used three neural network vendor software packages to train and test the models. The training set used 70% of the data with a test set of 30% of the data. The results of the neural network models were compared to stepwise multiple regression and logistic regression, focusing on the top 25 riskiest departments. The study results indicate the neural network models’ performance were superior when compared to the traditional statistical techniques. The authors conclude that internal auditors could benefit from using neural network technology for risk assessment.

O’Leary (1998) provides a meta-analysis of the use of neural networks to predict corporate failure. He reviewed and compared fifteen papers. The findings were compared on similarity of comparative solutions, number of correct classifications, impact of hidden layers and impact of percentage of bankrupt firms. O’Leary concluded that neural networks generated results at least as good as discriminant analysis, logit probit, and ID3. The author also found characteristics that influenced the quality of the neural network models. The training proportion of bankrupt firms, deviation from a single hidden layer and the time frame involved seemed to have a negative influence on the quality of neural network models.

Busta and Weinberg (1998) introduce an analytical review procedure that measures the degree to which a data set's digit distribution deviates from a Benford digit distribution. In theory, Benford's law states that for many types of data, the digits of the numbers are distributed in a predictable pattern. Accounting data is of such a type that would follow such a pattern, however, manipulated numbers would not be expected to follow a Benford pattern. The authors used a neural network to distinguish between "normal" and "manipulated" data. The degree to which a data set's digit distribution deviates from a Benford digit distribution can indicate potential manipulation and point to the need for further audit testing. The authors used 800 simulated data sets composed of 200 two-digit numbers with varying contamination distributions to train the neural network. A holdout sample of 800 data sets was used to test the model. The neural network classified 70.8% of the 800 data sets correctly. The results were sensitive to the degree of contamination: the more contamination, the more accurate the model. The authors conclude that neural network results compare favorably to traditional analytical review procedures.

Fanning and Cogger (1998) investigated the use of neural networks for detecting management fraud in published financial statements from over 200 US companies. The authors identified twenty variables to be most relevant concerning the indication of fraudulent financial reporting. These variables were used to classify the statements as either (probably) fraudulent or (probably) non-fraudulent. The results of the neural network model were compared to results from stepwise logistic regression, stepwise linear discriminant analysis and stepwise quadratic discriminant analysis. The neural network model accurately predicted 63 percent of the holdout sample while none of the other models had prediction accuracies greater than 52 percent. The authors suggest that neural networks offer superior ability to standard statistical techniques in detecting fraudulent financial statements.

METHODOLOGY

This study utilized a neural network, ordinary least squares (OLS) regression, discriminant analysis, and logistic regression (Logit) in a macro-case analysis of court cases. Judicial cases involving the employer-employee relationship provided the data for this study. The first case tried in the courts in this area was in 1940. The cases in the sample spanned the years 1940 to 1993 inclusive. Cases after 1993 would have resulted from audits initiated during a new era of compliance by the IRS and might have corrupted the data. Courts at both the trial level and appellate level were included. Several cases dealt with multiple, distinct sets of employees. Therefore, the total number of usable observations was 202. The court ruled against the IRS in 118 (58.4%) of the observations. Forty observations were randomly selected as a holdout sample, leaving 162 observations used to train the neural networks and develop the OLS, discriminant analysis, and logit models.

The cases were briefed noting the ruling, and the existence/non-existence of 23 factors. The 20 factors in Rev. Rul. 87-41 were used as well as three other factors. A preliminary study of 36 tax cases revealed three other factors often considered by the Court. These factors were: The relationship the parties think they are creating (a common law factor disregarded in Rev. Rul. 87-41); the offering of fringe benefits to the worker; and the degree of skill possessed by the worker (very high or very low).

Each of the 23 factors was rephrased into the form of a question (see Table 1). The text of each case was analyzed to determine the answers to the questions. These answers were used as the independent variables in this study. If a factor was not mentioned or the answer could not be determined, this was noted. The answers to the questions could be yes, no or 0 if not mentioned. These variables were then assigned a numeric value (1 if the answer was yes, -1 if the answer was no, and 0 if not mentioned) and used as independent variables for the prediction techniques. The dependent variable in this study was the ruling by the court. There were two possible outcomes for each observation; the court could have held that the employer-employee relationship existed or that the relationship did not exist. There were no compromises, that is, either the worker was an employee or the worker was not an employee. The dependent variable was coded as 1 if the court ruled that the employer-employee relationship existed and 0 if held not to exist.

TABLE 1
INDEPENDENT VARIABLES

QUESTION
1. Does the principal have the right to require that the worker comply with instructions about how the work is to be performed?
2. Does the principal require that the worker undergo training?
3. Are the worker's services an integral part of the principal's business?
4. Does the principal require that the worker render the services personally?
5. Does the principal hire, supervise, and pay the worker's assistants?
6. Does a continuing relationship exist between the principal and the worker?
7. Does the principal set the hours of work?
8. Does the principal require the worker to work full time?
9. Is the work either performed on the principal's premises or a route specified by the principal?
10. Does the principal set the order or sequence of the work performed?
11. Does the principal require the worker to submit oral or written reports, or is the worker subject to inspections?
12. Is the worker paid by the hour, week, or month?
13. Does the principal pay the worker's business and/or travel expenses?
14. Does the principal furnish the tools and materials?
15. Does the principal make the significant investment in the facilities?
16. Can the worker realize a profit or suffer a loss?
17. Does the worker work for more than one firm at a time?
18. Does the worker make his/her services available to the general public?
19. Does the principal have the right to fire the worker without incurring a liability?
20. Does the worker have the right to quit without incurring a liability?
21. Did at least one of the parties intend that the employer-employee relationship be established?
22. Does the principal offer fringe benefits to the worker?
23. Does the worker possess either a very high degree of skill or very low degree of skill?

Prediction Techniques

OLS, discriminant analysis, and logistic regression (logit) analysis were performed and a neural network was trained using the Genetic Adaptive Neural Network Training (GANNT) Algorithm developed by Dorsey, Johnson, and Mayer (Dorsey, et al).

OLS: Despite the theoretical limitations of using OLS with a dichotomous dependent variable, a step-wise OLS model was estimated for comparison purposes.

Discriminant Analysis: While acknowledging that some of the classic assumptions may be violated in this study, [2] a step-wise discriminant analysis was performed using the MAHAL method, which selects variables for inclusion based upon the Mahalanobis' distance between the groups (SPSS 1988, 460). Mahalanobis' distance is a generalized distance measure where D^2 is computed as the squared distance from point X (a specific case) to the group centroid (an imaginary point with coordinates that are the group's mean on each of the variables) (Klecka 1989, 16,44). Discriminant analysis has been a popular methodology for macro-case analysis.

Logit: In order to overcome the theoretical limitations of using ordinary least squares (OLS) with a dichotomous dependent variable, another method was needed. The solution is to specify a nonlinear probability model instead of a linear model. One such nonlinear model is the logit model (Aldrich and Nelson 1984, 31).

The logit model involves transforming a probability function into a cumulative logistic probability function (Pindyck and Rubinfeld 1981, 287). Assume that $\sum b_n x_n = Z$, then the expression $P=\exp(Z)/(1+\exp Z)$ (commonly referred to as the logistic function) is continuous and can take on any value from 0 to 1. Thus the logit model constrains the probability (P) to values from 0 to 1 even though the $\sum b_n x_n$ is not so constrained.

While the OLS model seeks to minimize the sum of squared errors in estimates of the coefficients of the independent variables, the logit model uses a maximum likelihood technique to estimate the coefficients. According to Aldrich and Nelson (1984), the conceptual difference between OLS and Maximum Likelihood Estimation (MLE) is that OLS is concerned with picking parameter estimates that produce the smallest sum of squared errors in fitting the model, while MLE is concerned with picking parameter estimates that imply the highest probability or likelihood of obtaining the observed Y. The logit model is appropriate when the dependent variable is dichotomous, the functional relationship is non-linear, and the independent variables are categorical. It should be mentioned that the nonlinear function is a special case of nonlinear functions while the ANN provides a more flexible alternative.

Neural Networks: Neural Networks (NN) comprise a class of nonlinear statistical models. In this class of models, input nodes ("sensors") send signals along connections that augment or pare the signal by a NN weight. A hidden processing node collects these weighted signals and passes the result through a cumulative distribution function (c.d.f.) (such as a sigmoid function), producing an output activation. The hidden nodes provide signals that are sent through additional weighted connections to the network output. These nonlinear models have been shown to be universal approximators (Funahashi (1989) and Hornik et al. (1989)). The ability of the network to be a universal approximator allows for the computation of *any* continuous function using linear summations and a single properly chosen nonlinear cumulative distribution function (c.d.f.). The arrangement of simple nodes into a multilayered framework produces a mapping between inputs and outputs consistent with any underlying functional relationship regardless of its "true" functional form. The significance of having a general mapping between the input and output vectors is that it eliminates the need for unjustified *a priori* restrictions commonly used to facilitate estimation (e.g., the Gauss Markoff assumptions in regression analysis). Indeed, according to Hornik et al. (ibid., p. 359), "multilayer feedforward networks with as few as one hidden layer are . . . capable of universal approximation in a very precise and satisfactory sense."

Being universal approximators means that the resulting model accurately depicts the underlying relationship between the input and output data and that the approximation technique is robust across different data sets and underlying functional relationships. They, therefore, accurately depict the relationship between inputs and outputs regardless of the data patterns, or underlying functional forms, and any lack of success in approximating the true relationship between inputs and outputs "must arise from inadequate learning, insufficient numbers of hidden units or the lack of an underlying deterministic relationship between input and target" (Hornik et al., 1989, p. 363).

Genetic Versus Backpropagation Training Algorithms

As in the Salchenberger et al. study, researchers have traditionally trained NNs using the backpropagation algorithm developed by Werbos (1974), Parker (1985), LeCun (1986), and Rumelhart (1986a, 1986b). Wasserman (1989) and Hecht-Nielsen (1990), however, have outlined various problems with the algorithm, including the tendency of the network to become trapped in local optima. Neural networks incorporating a genetic algorithm have been shown to overcome the problem of local optima as well as to outperform backpropagation in the selection of proper weights (Sexton et al., 1997). For this reason, we use the genetic adaptive neural network training (GANNT) algorithm developed by Dorsey et al. (1994). This training algorithm searches the weight space without use of any gradient information. The weights are coded in real valued strings whose fitness is determined by their effectiveness. Beginning with a set of randomly selected strings, the GANNT algorithm performs a global search rather than a local one. Unlike gradient descent methods, the GANNT algorithm does not become stuck in local minima. Also, the objective function does not have to be differentiable. Dorsey et al. (1994) show empirically that the genetic algorithm performs very well on a large class of problems with genetic network architectures. Thus, they demonstrate that the genetic-algorithm-based training method for the selection of the appropriate weight matrices overcomes many shortcomings of backpropagation and lends to the function approximation abilities of the neural network.

THE GENETIC ALGORITHM

The neural network configuration used in this study was a multilayer perceptron with 23 input nodes in the input layer, one hidden layer with five nodes, and one output layer with a single node. The network was trained using the Genetic Adaptive Neural Network Training (GANNT) algorithm.

As its name implies, the genetic algorithm seeks a solution in a manner similar to the Darwinian process of natural selection ("survival of the fittest"). The terminology even borrows from the field of genetics. The GANNT algorithm is based upon three processes: Reproduction, crossover, and mutation.

The algorithm may best be explained by an example. Assume a neural network with three input nodes (one for each of three independent variables), one hidden layer with two nodes, and an output layer with one node. Each input node is connected to each node in the hidden layer and each node in the hidden layer is connected to the output node, making a total of eight connections. Each connection has a weight, a total of eight weights.

In GANNT, these eight connections with their corresponding weights are called a "chromosome string." Each individual weight is called a "gene". That is, the matrix of weights is a chromosome string while each weight is a gene on the chromosome string. The researcher selects the number of chromosome strings to be used, which must be even (a requirement for the crossover stage of the algorithm).

This research used 30 chromosome strings and 100,000 iterations. An additional procedure was performed with an objective function to minimize the sum of squared errors (SSE) plus the number of connections. The procedure randomly selects a connection to be set to zero. If the objective function does not increase, the connection is left at zero and another connection is randomly selected to be set to zero and the procedure is repeated. If the objective function increases, the connection is reinstated and another connection is selected at random and the procedure is repeated. This process effectively eliminates those connections and ultimately those inputs that do not contribute to the final output (Johnson et. al., Working

Paper). At the end of the process, the chromosome string with the smallest SSE was selected as the solution to test the hold-out sample.

A formal description of the GANNT algorithm used in the present study (the description draws on the work of Dorsey and Mayer, 1994; and Dorsey, Johnson and Mayer, 1994) follows.

$$\max_{\vec{\gamma}\in\Xi,\,\vec{\beta}\in\Xi} f(\xi) = -\left((\tilde{y}_i - f(\vec{\gamma}_i, \vec{\beta}_i \mid \tilde{x}))\_(\tilde{y}_i - f(\vec{\gamma}_i, \vec{\beta}_i \mid \tilde{x}))\right)$$

To solve the problem:

Let

Ξ = A finite parameter space, $\Xi \subset R^k$, and $f{:}\,\Xi \to R$,

Ξ^m = $\displaystyle\mathop{\mathrm{X}}_{i=1}^{m} \Xi$,

$\vec{\gamma}_{ji}$ = a vector $\{\gamma_{j1i}, \gamma_{j2i}, \gamma_{j3i}, \ldots, \gamma_{j\,k-(MH+1)\,l}\}$ where $\gamma_{jli} \in \Xi^m$,

$\vec{\beta}_l$ = a vector $\{\beta_{i1}, \beta_{i2}, \ldots, \beta_{iMH}, \beta_{i0}\}$ where $\beta_{ij} \in \Xi$,

ξ_i = a vector $\{\vec{\gamma}_{1i}', \vec{\gamma}_{2i}', \ldots, \vec{\gamma}_{MH'i}, \vec{\beta}_i\}$ thus $\xi \subset R^k$ and $\xi \in \Xi$,

$\tilde{\xi}_i$ = a vector $\{\vec{\gamma}_{1i}', \vec{\gamma}_{2i}', \ldots, \vec{\gamma}_{MH'i},\}$ thus $\tilde{\xi} \subset R^{k-(MH+1)}$ and $\tilde{\xi} \in \Xi$,

$\tilde{\xi}_{ij}$ = an element of the vector $\tilde{\xi}_i$ where $j=1, \ldots, (MI+1)*MH$,

Ω = a probability space,

$\tilde{X}$ = $\{x_1, x_2, \ldots, x_{MI}, 1\}$ where $x_i \subset R^N$ and N is the total number of observations, MH is the total number of hidden nodes, $MI=((k-(MH+1))*MO)/MH-1$ is the number of input nodes (excluding bias),

$\tilde{y}$ = $\{y_1, y_2, \ldots, y_{MO}\}$ where y_i is observed value of the output s.t. $y_i \subset R^N$ and MO is the number of output nodes (in this case $MO=1$),

$\psi(\lambda)$ = a differentiable c.d.f. such as $(1+e^{-\lambda})^{-1}$, and $G^g : \Omega \times \Xi^m \to \Xi^m$,

where G^g denotes a process on a set of m candidate solutions (vectors of Ξ) corresponding to the g^{th} generation. The iteration process can be written schematically as follows:
$G^0 \to G^1 \to \ldots \to G^{c-1} \to G^c$, where convergence is achieved in the c^{th} generation. Iterations are terminated by a stopping rule such as: Stop when

$$\left| \max_{\xi\in G^g} f(\xi) - \max_{\xi\in G^{g-1}} f(\xi) < \delta \right|$$

and

$$\left| \operatorname*{argmax}_{\xi 0 G^g} f(\xi) - \operatorname*{argmax}_{\xi 0 G^{g-1}} f(\xi) < \varepsilon, \right|$$

The process G^g where $g=(1, ,\dots, c)$ draws m (an even number) weight vectors $\vec{\gamma}_{ji}$ uniformly from Ξ, and applies them each as parameters in the feedforward network

$$f(\vec{\gamma}_i, \vec{\beta}_i \mid \tilde{x}) = \beta_0 + \sum_{j=1}^{MH} \beta_j \psi\left(\tilde{x}\vec{\gamma}_{ji}{}'\right),$$

and then the optimal values of $\vec{\beta}_i$ $\forall$ i ($i=1, \dots, m$) are found using the least squares estimator

$$\vec{\beta}_i = (\psi(\tilde{x}\vec{\gamma}_{ji}{}')'\psi(\tilde{x}\vec{\gamma}_{ji}{}'))^{-1}\psi(\tilde{x}\vec{\gamma}_{ji}{}')'\tilde{y}_i.$$

An error vector $\vec{\varpi}_i$ $\forall$ i ($i=1, \dots, m$) is then generated where $\vec{\varpi}_i = \tilde{y}_i - f(\vec{\gamma}_i, \vec{\beta}_i \mid \tilde{x})$.

The objective function in this case is the sum of squared errors SSE_i $\forall$ i, where

$$SSE_i = \vec{\varpi}_i{}'\vec{\varpi}_i.$$

The process G^g where $g=(1,\dots, c)$ uses SSE_i to generate a fitness value $\Im(SSE)_i$ $\forall$ i in m which is essentially a selection probability. In particular, the selection probabilities determine which members of G^1 contribute offspring to the second generation, G^2, through the complete process G^c. The ξ_i most likely to contribute are those corresponding to the largest values of $\Im(SSE)_I$. One possible fitness function suggested by Dorsey and Mayer (1994) is

$$\Im(SSE_i) = 10 * \left(\frac{(SSE_{\max} - SSE_i)}{\sum_{j=1}^{m}(SSE_{\max} - SSE_j)} \right).$$

Note that the $\Im(SSE)_i$ chosen is required to be strictly increasing and non-negative. The non-negativity requirement ensures that the probabilities are well defined. The requirement that $\Im(SSE)_i$ is strictly increasing ensures that the most promising members of G^1 (the larger $f(\xi_i)^g$) are given the best chance of contributing to G^2.

The algorithm then proceeds as follows: Use the selection probabilities $\Im(SSE)_i$ to draw m parent vectors $\tilde{\xi}$ with replacement. The set $H^1 = \{\tilde{\xi}'_1, \dots, \tilde{\xi}'_m\}$ denotes the resulting vectors. Draw two vectors $\tilde{\xi}'_1$, $\tilde{\xi}'_2$, uniformly from H^1. Select an integer I from 0 to $k\text{-}MO*(MH+1)$ at random. Create a third and fourth vector by crossing over $\tilde{\xi}'_1$ and $\tilde{\xi}'_2$ at the I^{th} position as follows:

$$\tilde{\xi}_1'' = (\tilde{\xi}_{1,1}', \dots, \tilde{\xi}_{1,I}', \tilde{\xi}_{2,I+1}', \dots, \tilde{\xi}_{2,k-(MH+1)}'),$$

$$\tilde{\xi}_2'' = (\tilde{\xi}_{2,1}', \dots, \tilde{\xi}_{2,I}', \tilde{\xi}_{1,I+1}', \dots, \tilde{\xi}_{1,k-(MH+1)}'),$$

$\tilde{\xi}'_1$ and $\tilde{\xi}'_2$ are not replaced in H^1 (the "reproduction pool"). Repeat the uniform draws until $H^1 \subset \{\varnothing\}$ ($m/2$ times) and, thereby, generate the m vectors $\tilde{\xi}''_i$, $i = 1, \dots, m$.

Construct the set $G^2 = \{\tilde{\xi}'_1, \dots, \tilde{\xi}'_m\}$.This portion of G^g is commonly referred to as "reproduction and crossover" in the genetic algorithm literature. It is through these steps that the desirable traits from G^1 are passed on to G^2. This last step of the process is called "mutation". For each element of $\tilde{\xi}'$ pick a scalar $\tilde{\xi}'$ at random from Ξ_j. Let Y be the outcome of a Bernoulli trial and specify $\zeta = Prob(Y_i =$

1) and $1-\zeta = Prob(Y_i = 0)$. Generate $(k\text{-}MO*(MH+1))m$ observations on Y. Replace $\tilde{\xi}'_{ij}$ in $\tilde{\xi}'_i$ with $\tilde{\xi}''_{ij}$ iff $Y_i = 1$ on the corresponding trial. As was stated above, the process $G^g : \Omega\times\Xi^m \rightarrow \Xi^m$ continues step by step until convergence is achieved in the c^{th} generation.

Neural Net vs. Logit

Given the following equation for the neural network: $f(\vec{\gamma}_i, \vec{\beta}_i \mid \tilde{x}) = \beta_0 + \sum_{j=1}^{MH} \beta_j \psi(\tilde{x}\vec{\gamma}_{ji}')$,

where $\beta_0 = 0$, $MH=1$, $\beta_1 = 1$, $\beta_{2,MH} = 0$ and $\psi(\lambda) = (1+e^{-\lambda})^{-1}$, the neural net can be expressed as:

$$y = \frac{1}{1 + e^{-\sum_{i=1}^{N+1} w_i X_i}}$$

$$P = \frac{e^{(Z)}}{(1+e^{(Z)})} = \frac{1}{\frac{(1+e^{(Z)})}{e^{(Z)}}} = \frac{1}{\frac{1}{e^{(Z)}} + \frac{e^{(Z)}}{e^{(Z)}}} = \frac{1}{1+e^{-(Z)}}$$

The logit model as stated above where $\sum b_n x_n = Z$,
which is equivalent to the constrained neural network model. This means that the logit model is a special case of the neural network. Thus even if the logit model is correct, the neural net should give an equivalent result. Prior studies including a study by Hansen, et. al. have concluded that logit offers superior results. Since the logit model is a special case of the neural network this implies that the differences would have to be explained by a failure of the training method to yield as efficient parameter values as the maximum likelihood method of the logit model. Since the backpropagation method used in this study is a gradient technique, a global search technique was adopted for the neural network that has been shown to produce superior parameter estimates (Sexton et. al.).

RESULTS

OLS Step-Wise Model

The step-wise regression analysis produced a model with ten variables and a constant. Eight of these variables are significant at the .05 level. The significant variables are Variables 1 (instructions), 6 (continuity), 9 (employer premises), 11 (required reports), 13 (expenses paid), 14 (tools and materials), 15 (investment), and 16 (profit or loss).

Table 2 contains the results of the step-wise OLS regression analysis. The model has a multiple R of .8404, an R^2 of .7064, and an adjusted R^2 of .6868. The R^2 indicates that approximately 71% of the variation in Y can be explained by the regression equation, while the adjusted R^2 indicates approximately 69% of the variation in Y explained by the model. The F value is 36.308 and significant at the .0000 level. This model correctly predicted 150 (92.59%) of the 162 observations in the training set and 30 (75%) of the 40 observations in the test set. In spite of the problems inherent with the use of a dichotomous dependent variable and qualitative independent variables, the OLS model performed reasonably well overall. The step-wise model had an overall accuracy rate of 89.11%.

TABLE 2
OLS STEP-WISE MODEL

VARIABLE	COEFFICIENT	T STATISTIC	SIG T
V1 Instructions	.157985	4.882	*.0000
V6 Continuity	.074602	2.355	*.0198
V9 Employer premises	.057355	2.019	*.0452
V11 Reports	.107394	3.098	*.0023
V12 Mode of payment	.054608	1.703	.0905
V13 Expenses paid	.105884	3.045	*.0027
V14 Tools & materials	.092006	2.706	*.0076
V15 Investment	.075683	2.116	*.0360
V16 Profit or loss	-.086980	-2.331	*.0211
V18 Serves public	-.087562	-1.967	.0510
Constant	.554231	14.890	.0000

*Significant at the .05 levels.

Discriminant Analysis Model

The discriminant analysis procedure produced a model with thirteen variables and the constant. The variables are Variables 1(instructions), 14 (tools and materials), 13 (expenses paid), 11 (required reports), 18 (serves public), 12 (mode of payment), 15 (investment), 6 (continuity), 16 (profit or loss), 9 (employer premises), 22 (fringe benefits), 19 (discharge), and 20 (terminate).

The variables that entered the model were the variables that maximized the difference in group means. The discriminant function coefficients do not have tests of significance as do regression coefficients, but the size of the F-ratio and Wilks' lambda may be used as an indication of the significance of the variables. Therefore, it appears that the predominant factor is Variable 1 (instructions).

The model has a χ^2 of 198.20 that is significant at the .0000 level and the canonical correlation is 0.8515073. This canonical correlation squared $(0.8515073)^2$ equals .725064682, indicating that approximately 72.5% of the variance in the dependent variable can be explained by the model (Hair, et al, 1987). Table 3 contains the results of the discriminant analysis. The variables are listed in the order in which they entered the model. The model correctly predicted 153 (94.44%) of the observations in the training set and 35 (87.50%) of the observations in the test set. The model had an accuracy rate of 93.07% over the entire data set.

TABLE 3
DISCRIMINANT ANALYSIS MODEL

VARIABLE	COEFFICIENT	WILKS' LAMBDA	F Ratio
V1 Instructions	0.6618970	0.59514	20.578
V14 Tools & materials	0.4335871	0.44931	8.2595
V13 Expenses paid	0.4013377	0.40110	6.5069
V11 Reports	0.4468700	0.36623	8.3683
V18 Serves public	-0.3093316	0.34637	2.3604
V12 Mode of payment	0.1784834	0.33183	1.5531
V15 Investment	0.3155052	0.31885	3.9540
V6 Continuity	0.3537766	0.31077	6.1707
V16 Profit or loss	-0.4261557	0.30166	6.4826
V9 Employer premises	0.2091945	0.29373	2.6776
V22 Fringe benefits	0.3022897	0.28829	2.7392
V19 Discharge	0.4010003	0.28396	5.2917
V20 Terminate	-0.5297546	0.27494	4.8591
Constant	0.6177887		

Step-Wise Logistic Regression

Step-wise logistic regression produced a model with seven variables and a constant. The seven variables are Variables 1 (instructions), 6 (continuity), 11 (required reports), 13 (expenses paid), 14 (tools and materials), 16 (profit or loss), and 19 (discharge). All seven variables are significant at the .05 level based upon the Wald statistic.

The model has a goodness of fit χ^2 of 89.640. The log-likelihood ratio is 72.84, indicating that the variables in the model explain approximately 73% of the variance in the dependent variable. Table 4 contains the results of the step-wise logistic regression.

This model correctly predicted 151 (93.21%) of the 162 observations in the training set. The model correctly predicted 33 (82.50%) of the forty observations in the test set and has an accuracy rate of 91.09% over the entire data set.

TABLE 4
STEP-WISE LOGIT MODEL

VARIABLE	COEFFICIENT	WALD	SIG WALD
V1 Instructions	1.6593	9.0823	*.0026
V6 Continuity	1.2267	6.6418	*.0100
V11 Reports	1.8799	8.3154	*.0039
V13 Expenses paid	1.5003	8.4048	*.0037
V14 Tools & Materials	1.4035	8.9180	*.0028
V16 Profit or loss	-1.9651	9.6681	*.0019
V19 Discharge	1.0348	3.9393	*.0472
Constant	0.6967	1.2531	.2630

Neural Network Results

The neural network correctly predicted 154 (95.06%) of the 162 observations in the training set and correctly predicted 36 (90%) of the 40 observations in the test set. The accuracy rate was 94.06% over the full data set, with 190 of the 202 observations predicted correctly. The neural network has no tests of significance for the independent variables. The final model included twelve variables and one node in the hidden layer.

The twelve variables remaining in the neural network model are Variables 1 (instructions), 4 (personal service), 6 (continuity), 8 (full time), 11 (required reports), 3 (expenses paid), 15 (investment), 16 (profit or loss), 19 (discharge), 20 (terminate), 21 (intent of parties), and 23 (degree of skill). Due to the elimination of those variables making no contribution to the final results, the neural network results are included in the comparison of the step-wise methods.

Comparison of Prediction Accuracy

Table 5 contains the prediction accuracy rates of the techniques. The step-wise OLS, and step-wise logit had accuracy rates of 89.11%, and 91.09%, respectively, over the full data set indicating similar performance. The discriminant analysis, and neural network had accuracy rates of 93.07%, and 94.06%, respectively, over the full data set, indicating similar performance.

The neural network model has the best overall prediction accuracy (94.06%). The neural network outperformed all models in correctly predicting the test set (90%) and had a higher accuracy rate (94.06%) over the entire data set.

TABLE 5
COMPARISON OF PREDICTION ACCURACY

Prediction Technique	Training Set	Test Set	Full Data Set
Step-wise OLS	92.59%	75.00%	89.11%
Discriminant Analysis	94.44%	87.5%	93.07%
Step-wise Logit	93.21%	82.50%	91.09%
Neural Network	95.06%	90.00%	94.06%

Comparison of Step-Wise Methods

The primary objective of this study was to determine what variables the courts considered significant when deciding the issue of employer-employee relationship. The twenty factors listed in Rev. Rul. 87-14, plus the three factors gleaned from a review of the cases, were used to develop models using various methodologies. Step-wise methods eliminate the variables of lesser significance, leaving only the more important variables in the model. Practitioners and taxpayers alike could benefit from a more parsimonious model.

Table 6 contains a comparison of the variables considered important by the step-wise methods (OLS, logit, discriminant analysis, and neural network). A comparison of the results of the step-wise methods revealed only five variables remained in all four step-wise models. The Variables 1 (instructions), 6 (continuity), 11 (required reports), 13 (expenses paid), and 16 (profit or loss) were considered significant by all four prediction techniques.

TABLE 6
COMPARISON OF STEP-WISE METHODS

VARIABLE	OLS	D.A.	LOGIT	N.N.
V1 Instructions	X	X	X	X
V4 Personal service				X
V6 Continuity	X	X	X	X
V8 Full time				X
V9 Employer premises	X	X		
V11 Reports	X	X	X	X
V12 Mode of payment	X	X		
V13 Expenses paid	X	X	X	X
V14 Tools & materials	X	X	X	
V15 Investment	X	X		X
V16 Profit or loss	X	X	X	X
V18 Serves public	X	X		
V19 Discharge		X	X	X
V20 Terminate		X		X
V21 Intent				X
V22 Fringe benefits		X		
V23 Skill				X

SUMMARY AND RECOMMENDATIONS

The neural network had the highest overall prediction rates, followed by discriminant analysis, step-wise logit, and step-wise OLS, respectively. Since classification of workers is a dichotomous decision, the neural network results were superior to the results of the other prediction techniques in this study due to its higher prediction accuracy rate.

The step-wise prediction methods indicated that five variables were the most significant. Variables 1 (instructions), 6 (continuity), 11 (required reports), 13 (expenses paid), and 16 (profit or loss) were included in all the step-wise models. Congress could legislate a definition of employee that incorporates these five variables.

Suggestions for Future Research

One issue that was not addressed by this study was whether the choice of legal forum was a relevant factor in the outcome of the cases. It might be useful to determine if different courts exhibited different decision-making behavior. Future research in the area of the employer-employee relationship might investigate whether inconsistencies exist among the courts in applying the common law rules.

The neural network outperformed the other predictive techniques and produced superior results in this study. However, more research is needed to further validate neural network methodology as an alternative to traditional methodologies. A Monte Carlo study could be undertaken to compare the neural network performance to the performances of the more traditional methodologies such as OLS, discriminant analysis, and logit.

CONCLUSION

Although Congress, the Treasury Department, and the IRS have failed to adequately define employee, a feasible test is possible. Rev. Rul. 87-41 was issued in an effort to alleviate the problem, but the twenty factors listed are too numerous to be practically applied. Practitioners need a more concise definition with fewer factors in order to adequately advise clients when classifying workers.

The use of neural networks in macro-case analysis is a feasible alternative to the traditional methodologies. The neural network was superior in predicting the outcome of the cases. No *a priori* assumptions as to the functional form of the relationship were necessary. The field of macro-case analysis could be revived using NN methodology.

The field of accounting research could benefit from the use of neural network methodologies. The problems inherent in accounting research pose no obstacle to the use of neural networks. The neural network outperformed the traditional methodologies in predicting the existence of the employer-employee relationship. Neural networks have several advantages over other methods when conducting macro-case analysis. They perform well at pattern recognition. No *a priori* assumptions need be made at as to the functional form of the relationship; the neural network will find the functional form. A neural network inherently includes any interaction effects, and small sample size poses little obstacle for a neural network. More accounting research should be undertaken comparing neural networks with the traditional methodologies.

ENDNOTES

1. Quantitative judicial analysis is a branch of jurimetrics (a term developed by Loevinger in 1949). See Loevinger, 1949, P. 455.
2. Specifically, each group should have been drawn from a population with a multivariate normal distribution on the discriminating variables. However, previous research has shown that "discriminant analysis is a rather robust technique which can tolerate some deviation from these assumptions" (Klecka 1989, 61)
3. Technically speaking, for a wide class of nonlinear functions NNs can provide arbitrary approximations to arbitrary functions in a variety of normed function spaces (e.g., functions in Lp spaces and functions in Sobolev spaces with a Sobolev norm) provided a sufficient number of hidden nodes (see Lee et al., 1993).

REFERENCES

Aldrich, J. H. & Nelson , F. D..(1984). *Linear Probability, Logit & Probit Models*, Quantitative Applications in the Social Sciences Series: No. 07-045. Beverly Hills: Sage Publications.

Breault, A., Ager, D. H. & Gribens ,B.K. (1997). New initiatives for resolving worker classification issues. *Taxation for Accountants*, 58 (March) 185-186.

Burns, J. O. & Freeman, T.A. (1996). Avoiding IRS reclassification of workers as employees. *The Tax Adviser*, February, 102-109.
Busta, B. & Weinberg, R. (1998). Using Benford's Law and neural networks as a review procedure. *Managerial Auditing Journal*, Volume 13, Number 6. 356-366.
Cushing, W. W., Jr., & Arguea, N.M. (1999). Logit analysis of the employee classification for tax purposes. *Journal of Financial and Strategic Decisions*, Vol. 12. http://www.studyfinance.com/jfsd/index.html.
Dorsey, R.E., Johnson, J.D. & Mayer, W.J. The Genetic Adaptive Neural Network Training (GANNT) algorithm for generic feed forward artificial neural systems. Working paper, School of Business, University of Mississippi, University, Mississippi.
Dorsey, R.E., Johnson, J.D. & Mayer, W.J. (1994). A genetic algorithm for the training of feedforward neural networks. *Advances in Artificial Intelligence in Economics, Finance, and Management* (J.D. Johnson and A.B. Whinston, eds.). Vol. 1. Greenwich, CT: JAI Press Inc., pp. 93-111.
Fanning, K. & Cogger, K. (1998). Neural network detection of management fraud using published financial data. *International Journal of Intelligent Systems in Accounting, Finance and Management*, 7, No. 1, 21-42.
Funahashi, K.I. (1989). On the approximate realization of continuous mappings by neural networks. *Neural Networks* 2 (3): 183-192.
Hair, J. F., Anderson, R.E. & Tatham, R. L. (1987). *Multivariate Data Analysis: With Readings*, Second Edition, New York, New York: Macmillan Publishing Company.
Hansen, J. V., McDonald, J.B., & Stice J. D. (1992). Artificial intelligence and generalized qualitative-response models: An empirical test on two audit decision-making domains. *Decision Sciences*, 23, (3), (May/June): 708-723.
Hecht-Neilson, R. (1990). *Neurocomputing*. Reading, MA: Addison-Wesley.
Hornik, K., Stinchcombe, M. & White, H. (1989). Multilayer feed-forward networks are universal approximators. *Neural Networks* 2(5):359-366.
Johnson, J. D., Sexton, R. S. & R. E. Dorsey, Untitled working paper, College of Business, University of Mississippi, University, Mississippi.
Klecka, W. R. (1989). *Discriminant Analysis*, Quantitative Applications in the Social Sciences Series: No. 07-019, Beverly Hills: Sage Publications.
Kort, F. (1977). *A Special and a General Multivariate Theory of Judicial Decisions*, Sage Professional Papers in American Politics, Vol. 3, No. 04-033, Beverly Hills: Sage Publications.
LeCun, Y. (1986). Learning processes in an asymmetric threshold network. *Disordered Systems and Biological Organization* (E. Bienenstock, F. Fogelman Souli and G. Weisbuch, eds.). Berlin: Springer Verlag, pp. 233-240.
Lee. T.H., White, H. & Granger, C.W.J. (1993). Testing for neglected nonlinearity in time series model: A comparison of neural network methods and alternative tests. *Journal of Econometrics* 56(3): 269-290.
Martindale, B. C. & Price, J. E. Federal Employment Status Determination. Working Paper.
O'Leary, D.E. (1998). Using neural networks to predict corporate failure. *International Journal of Intelligent Systems in Accounting, Finance and Management*. 7 (3): 187-197.
O'Neil, C. J. & Nelsestuen, L. (1993). "Employee or independent contractor status? Conflicting Letter Rulings continue controversy. *Tax Notes*, (May 17): 961-972.
Parker, D. (1985). *Learning Logic*. Technical Report TR-87. Cambridge, MA: Center for Computational Research in Economics and Management Science, MIT.
Ramamoorti, S., A.D. Bailey, A.D. Jr., & Traver, R.O. (1999). Risk assessment in internal auditing: A neural network approach. *International Journal of Intelligent Systems in Accounting, Finance and Management*, 8 (3): 159-180.
Robinson, J. & Hulen, M. (1996). Employees or independent contractors? Simplified rules for classifying workers. *National Public Accountant*. (March): 28-32.

Salchenberger, L.A., Cinar, E.M. & Lash, N.A. (1992). Neural Networks: A new tool for predicting thrift failures. *Decision Sciences* 23 (July/August): 899-916.
Sexton, R., Dorsey, R., & Johnson, J. (1998). Toward global optimization of neural networks: A comparison of the genetic algorithm and backpropagation. *Decision Support Systems* 22: 171-185
Stewart, D.N. (1980). Employee or independent contractor: An examination of the relevant variables employed by the federal courts in deciding the question. Ph.D. dissertation, University of Florida.
Stone, M. & Rasp, J. (1991). Tradeoffs in the choice between Logit and OLS for accounting choice studies. *The Accounting Review* 66: 170-187.
Vizcaino v. Microsoft Corp., 97 F3d 1187 (CA-9, 1996)
Wasserman, P.D. (1989). *Neural Computing: Theory and Practice*. New York: Van Nostrand Reinhold.
Werbos, P. 1974. Beyond regression: New tools for prediction and analysis in the behavioral sciences. Unpublished Ph.D. thesis, Harvard University.
Zelezny, A. (1999). The application of artificial neural networks in the fields of auditing. Masters Thesis. University of Vienna. Vienna, Austria.

APPENDIX A
TABLE OF CASES

Aberdeen Aerie No. 24 of Fraternal Order of Eagles v. U.S., 50 F.Supp. 734 (D.Ct.Wa., 1943).
Adams v. Commissioner, 43 TCM 1203 (1982).
Air Terminal Cab, Inc. v. U.S., 478 F.2d 575 (CA-8, 1973).
G.W. Allen & Co. v. Henslee, 86 F.Supp. 295 (D.Ct.Tn., 1949).
American Consulting Corporation v. U.S., 311 F.Supp. 715 (D.Ct.Pa., 1970).
American Homes of New England, Inc. v. U.S., 173 F.Supp. 857 (D.Ct.Ma., 1959).
American Institute of Family Relations v. U.S., 79-1 USTC &9364 (D.Ct.Ca., 1979).
Apaca, Inc. v. U.S., 77-1 USTC &9139 (D.Ct.N.M., 1976).
Aparacor, Inc. v. U.S., 556 F.2d 1004 (Ct.Cls., 1977).
Associated Bicycle Service, Inc. v. U.S., 128 BR 436 (Bkt.Ct.In., 1991).
Atlantic Coast Life Ins. Co. v. U.S., 76 F.Supp. 627 (D.Ct.S.C., 1948).
Avis Rent A Car System, Inc. v. U.S., 503 F.2d 423 (CA-2, 1974).
Azad v. U.S., 388 F.2d 74 (CA-8, 1968).
Barrett v. Phinney, 278 F.Supp. 65 (D.Ct.Tx., 1967).
Bartels v. Birmingham, 59 F.Supp. 84 (D.Ct.Ia., 1945).
Beatty v. Halpin, 267 F.2d 74 (CA-8, 1959).
Beaverdale Memorial Park v. U.S., 47 F.Supp. 663 (D.Ct.Ct., 1942).
Beckwith v. U.S., 67 F.Supp. 902 (D.Ct.Ma., 1946).
Ben v. U.S., 139 F.Supp. 883 (D.Ct.N.Y., 1956).
Ben Construction Corporation v. U.S., 312 F.2d 78 (Ct.Cls., 1963).
Beyer v. U.S., 78-1 USTC &9307 (Ct.Cls., 1978).
Bilenas v. Commissioner, 47 TCM 217 (1983).
Biltgen v. Reynolds, 58 F.Supp. 909 (D.Ct.Mn., 1943).
Bonded Insulation & Construction Co., Inc. v. U.S., 131 F.Supp. 635 (D.Ct.N.J., 1955).
Bonney Motor Express v. U.S., 206 F.Supp. 22 (D.Ct.Va., 1962).
Bothke v. Commissioner, 39 TCM 826 (1980).
Breaux and Daigle, Inc. v.U.S., 900 F.2d 49 (CA-5, 1990).
Henry Broderick, Inc. v. Squire, 69 F.Supp. 109 (D.Ct.S.D., 1946).
Burnetta v. Commissioner, 68 TC 387 (1977).
Burruss v. Early, 44 F.Supp. 21 (D.Ct.Va., 1942).

Cannon Valley Milling Co. v. U.S., 59 F.Supp. 785 (D.Ct.Mn, 1945).
Cape Shore Fish Co., Inc. v. U.S., 330 F.2d 961 (Ct.Cls., 1964).
Capital Life & Health Insurance Co. v. Bowers, 90 F.Supp. 600 (D.Ct.S.C., 1950).
Capital Trawlers, Inc. v. U.S., 216 F.Supp. 440 (D.Ct.Me., 1963).
Cassady v. Commissioner, 52 TCM 1130 (1986).
Chase Manufacturing, Inc. v. U.S., 446 F.Supp. 698 (D.Ct.Mo., 1978).
Club Hubba Hubba v. U.S., 239 F.Supp. 324 (D.Ct.Hi., 1965).
Combined Metals Reduction Co. v. U.S., 53 F.Supp. 739 (D.Ct.Ut., 1943).
Consolidated Housecraft, Inc. v. U.S., 170 F.Supp. 842 (D.Ct.N.Y., 1959).
Co-op Cab Co. v. Allen, 82 F.Supp. 695 (D.Ct.Ga., 1947).
Crawford Packing Company v. U.S., 228 F.Supp. 549 (D.Ct.Tx., 1962).
Critical Care Registered Nursing, Inc v. U.S., 776 F.Supp. 1025 (D.Ct.Pa., 1991).
Crossett Lumber Co. v. U.S., 79 F.Supp. 20 (D.Ct.Ar., 1948).
Culp v. Commissioner, 47 TCM 1112 (1984).
Cutler v. U.S., 180 F.Supp. 360 (Ct.Cls., 1960).
De-Raef Corporation v. U.S., 70 F.Supp. 264 (Ct.Cls., 1947).
Diaz v. U.S., 90-1 USTC &50,209 (D.Ct.Ca., 1990).
Dillon v. Commissioner, 56 TCM 1038 (1989).
Edwards v. U.S., 168 F.Supp. 955 (Ct.Cls., 1958).
Eicher v. Commissioner, 48 TCM 1013 (1984).
Emard v. Squire, 58 F.Supp. 281 (D.Ct.Wa., 1945).
Entner v. U.S., 67 F.Supp. 684 (D.Ct.Oh., 1946).
Etherton v. U.S., 278 F.Supp. 568 (D.Ct.Tn., 1967).
Everhart v. U.S., 71-1 USTC &9368 (D.Ct.N.C., 1971).
Ewing v. Vaughn, 169 F.2d 837 (CA-4, 1948).
Farm & Home Modernization Corp. v. U.S., 138 F.Supp. 423 (D.Ct.N.Y., 1956).
Filipidis v. U.S., 71-1 USTC &9197 (D.Ct.Md., 1970).
Fleeman v. U.S., 175 F.Supp. 336 (D.Ct.Oh., 1959).
Gamal-Eldin v. Commissioner, 55 TCM 582 (1988).
General Investment v. U.S., 823 F.2d 337 (CA-9, 1987).
General Wayne Inn, Inc. v. Rothensies, 47 F.Supp. 391 (D.Ct.Pa., 1942).
Gensler-Lee v. U.S., 70 F.Supp. 675 (D.Ct.Ca., 1946).
Gilmore v. U.S., 443 F.Supp. 91 (D.Ct.Ma., 1977).
Gramercy Construction Co. v. U.S., 64-1 USTC &9333 (D.Ct.N.Y., 1964).
Gregg v. Commissioner, 12 TCM 478 (1953).
Gulf Oil Corporation v. U.S., 57 F.Supp. 376 (D.Ct.Pa., 1944).
Hamlin v. U.S., 74-2 USTC &9578 (D.Ct.Ks., 1974).
Harris v. Commissioner, 36 TCM 1426 (1977).
Darrell Harris, Inc. v. U.S., 91-1 USTC &50,271 (D.Ct.Ok., 1991).
Henry v. U.S., 452 F.Supp. 253 (D.Ct.Tn.,1978).
Herman v. Commissioner, 52 TCM 1194 (1986).
Hirsch v. Rothensies, 56 F.Supp. 92 (D.Ct.Pa., 1944).
Illinois Tri-Seal Products, Inc. v. U.S., 353 F.2d 216 (Ct.Cls., 1965).
Imholte v. U.S., 118 BR 103 (Bkt.Ct.Ak., 1990).
Indian Refining Co. v. Dallman, 31 F.Supp. 455 (D.Ct.Il., 1940).
International Exterminator Corp. v. U.S., 305 F.Supp. 676 (D.Ct.Tx., 1969).
Investors Heritage Life Insurance Co. v. U.S., 79-1 USTC &9246 (D.Ct.Ky., 1979).

Henry J. Jackson v. Commissioner, 12 TCM 437 (1953).
Jackson v. Phinney, 266 F.Supp. 835 (D.Ct.Tx., 1967).
Jagolinzer v. U.S., 150 F.Supp. 489 (D.Ct.R.I., 1957).
James v. Commissioner, 25 TC 1296 (1956).
E.C. Jones v. U.S., 79-1 USTC &9120 (D.Ct.Tx., 1978).
M.H. Jones v. U.S., 76-2 USTC &9607 (D.Ct.Ok., 1976).
Kentucky Cottage Industries, Inc. v. Glenn, 39 F.Supp. 642 (D.Ct.Ky., 1941).
Kershner v. Commissioner, 14 TC 168 (1950).
Kiddie, Inc. v. Commissioner, 69 TC 1055 (1978).
Kirkconnell v. U.S., 347 F.2d 260 (Ct.Cls., 1965).
Kizzier v. U.S., 598 F.2d 1128 (CA-8, 1979).
Klinger v. U.S., 78-2 USTC &9668 (D.Ct.Ia., 1978).
Kroblin Refrigerated Xpress, Inc. v. U.S., 78-2 USTC &9809 (D.Ct.Ia., 1978).
Kurio v. U.S., 281 F.Supp. 252 (D.Ct.Tx., 1968).
W.O. Lakie, Inc., v. U.S., 70 F.Supp. 665 (D.Ct.Mi., 1946).
Lanigan Storage & Van Company, Inc. v. U.S., 389 F.2d 337(CA-6, 1968).
Levin v. Manning, 124 F.Supp. 192 (D.Ct.N.J., 1952).
Lieb v. U.S., 438 F.Supp. 1015 (D.Ct.Ok., 1977).
Loeb v. U.S., 209 F.Supp. 22 (D.Ct.La., 1962).
Los Angeles Athletic Club v. U.S., 54 F.Supp. 702 (D.Ct.Ca., 1944).
Mainer v. Commissioner, 55 TCM 1168 (1988)
March v. Commissioner, 42 TCM 283 (1981).
Martin v. U.S., 267 F.Supp. 268 (D.Ct.Fl., 1967).
Martin v. Wichita Cab Co., 46-2 USTC &9325 (S.Ct.Ks., 1946).
Matt v. Commissioner, 59 TCM 472 (1990).
Mav Freight Service, Inc. v. U.S., 462 F.Supp. 503 (D.Ct.Ut., 1978).
McAdams Drilling Company v. U.S., 62-2 USTC &9812 (D.Ct.Ok., 1962).
William C. McCombs Company, Inc. v. U.S., 436 F.2d 979 (Ct.Cls., 1971).
McCormack v. Commissioner, 52 TCM 1321 (1987).
McCormick v. U.S., 531 F.2d 554 (Ct.Cls., 1976).
McGowan v. Lazeroff, 148 F.2d 512 (CA-2, 1945).
McQuire v. U.S., 349 F.2d 644 (CA-9, 1965).
Meisinger v. Commissioner, 11 TCM 584 (1952).
Mervis v. U.S., 187 F.Supp. 248 (D.Ct.La., 1960).
Metropolitan Roofing & Modernization Company, Inc. v. U.S., 125 F.Supp. 670 (D.Ct.Ma., 1954).
M.F.A. Mutual Insurance Company v. U.S., 314 F.Supp. 590 (D.Ct.Mo., 1970).
Millard's Incorporated v. U.S., 146 F.Supp. 385 (D.Ct.N.J., 1956).
Mladinich v. U.S., 379 F.Supp. 117 (D.Ct.Ms., 1974).
A.D. Moore v. U.S., 92-2 USTC &50,401 (D.Ct.Mi., 1992).
Moore v. U.S., 68-2 USTC &9661 (D.Ct.Tx., 1968).
Harry D. Morgan Signs, Inc. v. U.S., 67-2 USTC &9742 (D.Ct.Or., 1967).
Morish v. U.S., 555 F.2d 794 (Ct.Cls., 1977).
Mutual Trucking Co. v. U.S., 51 F.Supp. 114 (D.Ct.Oh., 1943).
Neill v. U.S., 271 F.Supp. 445 (D.Ct.Ms., 1967).
Nevin, Inc., v. Rothensies, 58 F.Supp. 460 (D.Ct.Pa., 1945).
O'Brien v. Commissioner, 79 TC 776 (1982).
O'Hara Vessels, Inc. v. Hassett, 60 F.Supp. 672 (D.Ct.Ma., 1942).

Orange State Oil Co. v. Fahs, 52 F.Supp 509 (D.Ct.Fl., 1942).
Pack v. U.S., 434 F.Supp. 232 (D.Ct.Tn., 1977).
Packard v. Commissioner, 63 TC 621 (1963).
Party Cab Co. v. U.S., 75 F.Supp. 307 (D.Ct.Il., 1947).
Powers v. U.S., 424 F.2d 593 (Ct.Cls., 1970).
Professional & Executive Leasing, Inc. v. Commissioner, 89 TC 225 (1987).
Pulver v. Commissioner, 44 TCM 644 (1982).
R & H Corporation v. U.S., 255 F.Supp. 870 (D.Ct.Pa., 1966).
Radcliff v. Commissioner, 59 TCM 61 (1990).
Radio City Music Hall Corporation v. U.S., 50 F.Supp. 329 (D.Ct.N.Y., 1942)
Radovich v. Commissioner, 13 TCM 1121 (1954).
Ralls, Inc. v. U.S., 470 F.2d 579 (Ct.Cls., 1972).
Rayhill v. U.S., 364 F.2d 347 (Ct.Cls., 1966).
Reserve National Insurance Co. v. U.S., 74-1 USTC &9486 (D.Ct.Ok., 1974).
Rice v. Commissioner, 47 TCM 852 (1984).
Ringling Bros-Barnum & Bailey Combined Shows v. Higgins, 189 F.2d 865 (CA-2, 1951).
Schwing v. U.S., 65 F.Supp. 227 (D.Ct.Pa., 1946).
Security Roofing & Construction Co., Inc. v. U.S., 163 F.Supp. 794 (D.Ct.Ma., 1958).
Service Trucking Co., Inc. v. U.S., 347 F.2d 671 (CA-4, 1965).
Shreveport Laundries, Inc. v. U.S., 84 F.Supp. 435 (D.Ct.La., 1949).
Silver v. U.S., 131 F.Supp. 209 (D.Ct.N.Y., 1954).
Simpson v. Commissioner, 64 TC 974 (1964).
Paul Smith v. U.S., 64-2 USTC &9511 (D. Ct.Tx., 1963).
Spicer Accounting, Inc. v. U.S., 918 F.2d 90 (CA-9, 1991).
Spirella Co., Inc. v. McGowan, 52 F.Supp. 302 (D.Ct.N.Y., 1943).
Springfield Productions, Inc. v. Commissioner, 38 TCM 74 (1979).
Standard Chemical Manufacturing Co. v. U.S., 78-1 USTC &9431 (D.Ct.Nb., 1978).
Standard Life & Accident Insurance Company v. U.S., 75-1 USTC &9352 (D.Ct.Ok., 1975).
Standard Oil Co. v. Glenn, 52 F.Supp. 755 (D.Ct.Ky., 1943).
Star Fish & Oyster Company, Inc. v. U.S., 223 F.Supp. 402 (D.Ct.Al., 1963).
Star Oil Company, Inc. v. U.S., 78-1 USTC &9160 (D.Ct.Ok., 1977).
Steffens v. Commissioner, 49 TCM 54 (1984).
Sterns v. Clauson, 122 F.Supp. 795 (D.Ct.Me, 1954).
Stone v. U.S., 55 F.Supp. 230 (D.Ct.Pa., 1943).
Tapager v. Birmingham, 75 F.Supp. 375 (D.Ct.Ia., 1948).
Ten Eyck Co., Inc. v. U.S., 41 F.Supp. 375 (D.Ct.N.Y., 1941).
Texas Co. v. Higgins, 32 F.Supp. 428 (D.Ct.N.Y., 1940).
Theresa Enterprises, Inc. v. U.S., 78-2 USTC &9630 (D.Ct.Ca., 1978).
Thor Company v. U.S., 173 F.Supp. 65 (D.Ct.Ma., 1959).
Thornton v. U.S., 102 F.Supp. 554 (Ct.Cls., 1952).
Tidwell v. U.S., 63 F.Supp. 609 (D.Ct.Tn., 1945).
Titanium Ores Corporation v. U.S., 205 F.Supp. 606 (D.Ct.Md., 1962).
T.L. Squared, Inc. v. U.S., 74-1 USTC &9260 (D.Ct.Oh., 1974).
Tomlin v. U.S., 70 F.Supp. 677 (D.Ct.Ca., 1946).
Tristate Developers, Inc. v. U.S., 549 F.2d 190 (Ct.Cls., 1977).
U.S. v. Conforte, 624 F.2d 574 (CA-9, 1980).
U.S. v. Erickson, 78-1 USTC &9164 (D.Ct.Ws., 1977).

U.S. v. Polk, 550 F.2d 566 (CA-9, 1977).
U.S. v. Rasbury, 92-1 USTC &50,195 (D.Ct.Al., 1992).
W.M. Webb, Inc. v. U.S., 271 F.Supp. 249 (D.Ct.La., 1967).
Whalen v. Harrison, 51 F.Supp. 515 (D.Ct.Il., 1943).
Willard Storage Battery Co. v. Carey, 103 F.Supp. 7 (D.Ct.Oh., 1952).
Willard Sugar Co. v. Gentsch, 59 F.Supp. 82 (D.Ct.Oh., 1944).
Williams v. U.S., 38 F.Supp. 536 (D.Ct.Il., 1941).
Wolfe v. U.S., 570 F.2d 278 (CA-8, 1978).
Woolson Spice Co. v. U.S., 85 F.Supp. 242 (D.Ct.Oh., 1949).
World Market Centers, Inc. v. IRS, 77-1 USTC &9277 (D.Ct.Tx., 1976).
Yearwood v. U.S., 55 F.Supp. 295 (D.Ct.La., 1944).
Yellow Cab Corporation v. U.S., 210 F.Supp. 64 (D.Ct.N.J., 1962).
Zipley v. U.S., 156 F.Supp. 141 (D.Ct.Pa., 1957).