©*The Journal of Risk and Insurance*, 2002, Vol. 69, No. 3, 341-371

Fraud Classification Using Principal Component Analysis of RIDITs

Patrick L. Brockett
Richard A. Derrig
Linda L. Golden
Arnold Levine
Mark Alpert

Abstract

This article introduces to the statistical and insurance literature a mathematical technique for an a priori classification of objects when no training sample exists for which the exact correct group membership is known. The article also provides an example of the empirical application of the methodology to fraud detection for bodily injury claims in automobile insurance. With this technique, principal component analysis of RIDIT scores (PRIDIT), an insurance fraud detector can reduce uncertainty and increase the chances of targeting the appropriate claims so that an organization will be more likely to allocate investigative resources efficiently to uncover insurance fraud. In addition, other (exogenous) empirical models can be validated relative to the PRIDIT-derived weights for optimal ranking of fraud/nonfraud claims and/or profiling. The technique at once gives measures of the individual fraud indicator variables' worth and a measure of individual claim file suspicion level for the entire claim file that can be used to cogently direct further fraud investigation resources. Moreover, the technique does so at a lower cost than utilizing human insurance investigators, or insurance adjusters, but with similar outcomes. More generally, this technique is applicable to other commonly encountered managerial settings in which a large number of assignment decisions are made subjectively based on "clues," which may change dramatically over time. This article explores the application of these techniques to injury insurance claims for automobile bodily injury in detail.

Patrick Brockett is the Gus S. Wortham chaired professor of risk management at the University of Texas at Austin. Richard Derrig is senior vice president, Automobile Insurers Bureau of Massachusetts and vice president of research, Insurance Fraud Bureau of Massachusetts. Linda Golden is the Marlene and Morton Meyerson Centennial Professor in Business at the University of Texas. Arnold Levine is professor emeritus in the Department of Mathematics, Tulane University. Mark Alpert is a professor of marketing at the University of Texas.

FRAUD CLASSIFICATION USING PRINCIPAL COMPONENT ANALYSIS RIDITS

Insurance investigators, adjusters, and insurance claim managers are often faced with situations where there is incomplete information for decision making concerning the validity or possible fraudulent status of a particular filed claim. In all circumstances, strategic and tactical decisions must be made anyway, such as whether to pay the claim, refer the file to a special investigative unit (SIU), or even refer the case to the legal department or attorney general's office. Typical of these situations are instances in which an automobile accident occurs and a claim is filed for bodily injury involving injury to the soft tissue (Weisberg and Derrig, 1991, 1992; Weisberg et al., 1994). There may have been no witnesses to the accident and no police report filed, and the claimant may not have sought medical treatment for days following the accident. In these and other similar types of situations, suspicion levels may be high, and being able to classify claims according to their potential for successful negotiation or prosecution for fraud would be useful. Perhaps more important, the ability to separate out the bulk of filed claims that are apparently valid, and to pay them quickly, not only frees up needed investigator resources, but also creates goodwill with the insuring public and helps to dispel slow-payment and bad faith settlement lawsuits. The problem, of course, is that, in a statistical sense, in these situations the criterion variable (in this case, the knowledge of whether fraud actually occurred in a particular case) is not obtainable before decision-making action must be taken and never will be known for the vast majority of the claims.[1] This type of problem is faced not only by fraud investigators, but also by many other managerial decision makers for whom decisions must be made without having access to a sample frame containing the known dependent variable needed to apply traditional discriminant analysis type modeling techniques (probit, logit, feed forward-back propagation neural networks, and so on) or regression analysis.[2]

[1] According to a study of Massachusetts auto insurance, 72 percent of auto injury tort claims filed in 1991 did not result in litigation. Of those that did result in litigation, about 99 percent were settled before an actual jury verdict was reached (Weisberg and Derrig, 1991). Consequently, the number of situations in which one can actually observe the true value of the dependent variable (legal determination of fraud or legally not provable as fraud) is quite small (0.28 percent). This is different from situations such as random audits by the IRS, which can be useful for developing classification functions in that in these latter cases a sample of known results exists upon which to base modeling estimates. Randomly pursuing litigation for fraud in insurance claims can result in excessive legal fees, terrible public relations, regulatory sanctions, and multiple damages!

[2] The underlying available information assumptions necessary to perform PRIDIT analysis differ from (are less stringent than) those required of regression, probit, logit, discriminant analysis, or other "classification" methodologies that rely on having "training samples" from the fraud and the nonfraud groups. The PRIDIT methodology does not require a knowledge of which respondents indulged in fraud in order to implement it; that is, it does not require delineation of a sample of cases in which the occurrence of fraud together with its covariates are known as well as a delineation of a sample of cases in which the absence of fraud together with its covariates are known—this is how it differs from regression, discriminant analysis, and so on. One cannot fit a regression model without a dependent variable. In the artificial intelligence literature the regression, probit, logit, discriminant, and neural network analysis information scenario is known as "supervised learning." One knows the correct classification for some group of respondents and can "supervise," or optimize, the estimation process for selecting parameter values in such a manner that the model optimally distinguishes between

In short, we have a set of claims, some of which are fraudulent and some of which are not, and we do not know which is which. How do we assign claims to each of these two groups? The problem can be stated generally as assignment of members of a population to one of two subgroups. Although that problem is common in managerial settings, for consistency and succinctness, the discussion throughout will be framed in terms of the fraud detection nomenclature.

Background and Motivation for the Research

Frequently, a substantial amount of information about a particular filed insurance claim is known, but we do not know how to classify it according to its validity. Standard types of statistical techniques (discriminant analysis, probit, logit, feed forward-back propagation neural networks, and so on) require the insurance investigator or claims adjuster to use existing information on a set of claims previously classified as fraudulent or nonfraudulent to develop a scoring program for new claims (Artís et al., 1999; Weisberg and Derrig, 1998) and require interval-level numerical data for the statistical analysis. For example, a typical application for classical statistical fraud detection would be to use data on fraudulent and nonfraudulent credit card transactions to develop a model that would allow the classification of new claims according to their prospective likelihood of being fraudulent. For a recent approach to modeling fraud in this context, see Belhadji et al. (2000).

More traditional techniques, such as regression, discriminant analysis, and logistic regressions, are no longer useful when the cost of obtaining valid information on the criterion variable (fraud versus nonfraud) is excessive or when such information is impossible to obtain from a practical standpoint. A study using matched claim files of those convicted of fraud and those not convicted of fraud, for example, would very likely still be biased—it is very difficult to prosecute and obtain a conviction for fraud (so those who would be convicted would yield only the most obvious and extreme cases). Moreover, the set of claims classified by the legal system as *nonfraudulent* can be contaminated by unconvicted or unrecognized fraud cases (Caron and Dionne, 1999; Dionne and Belhadji, 1996). In addition, although we might have historical information concerning the proportion of those claims that we suspect are fraudulent to those that are not, information may not be available on the same claim files concerning both the predictor and criterion variables so that classification information is actually exogenous to the data set for modeling fraud likelihood. Given a set of such predictor variables, it is neither practical nor legal to wait for the data collection process to be exhausted in order to obtain a complete set of predictor and criterion variables on the same individuals and make a decision. Moreover, perpetrators of fraud have a "learning curve," so even if all of the obstacles to classical analyses just noted could be overcome, the predictor variables may change significantly over time, rendering these methods powerless.

the known correctly classified subset of cases. Regression, probit, logit, discriminant, and neural network analyses require both the existence and the knowledge of an already classified set of data, whereas PRIDIT analysis does not. In the artificial intelligence literature this latter situation is known as "unsupervised learning," wherein one does not have a training set containing the covariates and the dependent variable. Consequently, one cannot supervise or optimize the estimation process for selecting parameter values in such a manner that the model optimally distinguishes between a known correctly classified subset of cases.

Since which claims are fraudulent and which are not is unknown in the context of this study, the claims adjuster or insurance investigator could use a rigorous methodology to augment investigative procedures within budget, time, and data availability constraints, requiring only data on the gathered fraud predictor variables. One relatively complex method that has been used is the Kohonen Self-Organizing Feature Map[3] applied to so-called fraud, or suspicion of fraud, indicators[4] (Brockett et al., 1998). This article presents a new and simpler nonparametric technique (PRIDIT) that is more easily understood and implemented and can satisfy this managerial need. Applications of the PRIDIT methodology can also extend to classifications finer than the instant binary fraud/nonfraud case. This new methodology provides additional value in its ability to test the consistency of scoring model output with input variable patterns. Specifically, the weights and scores obtained from the PRIDIT methodology are representative of input variable patterns and can be tested for correlation with otherwise determined model scores. For example, if a set of insurance adjusters' subjective assessments of claim file fraud suspicion levels existed, then these could be correlated with the PRIDIT scores for the claim files and hence validated. In addition, PRIDIT score correlations with exogenous variables could be used for "profiling" if desired (for example, target marketing if used in a marketing context, or "red-flagging" in an insurance context). As we shall demonstrate, high absolute correlations indicate consistent modeling scores.

FRAUD INVESTIGATION FROM A STATISTICAL PERSPECTIVE

One potential statistical approach that might be used by fraud investigators to aid decision making in claim handling situations when the dependent variable is unknown is to model the total data set of predictor variables as a probabilistic mixture of two groups, each having some known parametric form, and where the classification or criterion variable (the group membership variable) is itself treated as being unknown or missing. One might then use this unclassified initial sample to obtain the parameter estimates necessary to form classical likelihood ratio or discriminant functions. For example, Ganesalingam and McLachlan (1978) considered the situation where the data obtained arise from a mixture of two univariate normal distributions with common variance σ^2 and means μ_1 or μ_2 depending upon whether the (unclassified) observation was made from group one or two. If the proportion from group one is θ, then the likelihood equation corresponding to this mixture can be explicitly written down and maximized to obtain estimates of the four unknown parameters μ_1 or μ_2, σ^2 and θ. These are then inserted into the likelihood ratio (discriminant) function for possible future classification of new observations. While in theory the situation of (unclassified) multivariate normal observations from each group might be handled analogously, the number of parameters to estimate grows very quickly. More generally,

[3] Like the PRIDIT methodology introduced in this article, the Kohonen Self-Organizing Feature Map methodology uses unsupervised learning; that is, without feeding back knowledge of the dependent variable, in a neural network-type format. However, the Kohonen method is more difficult to implement and fails to give the same information to the fraud investigator as the proposed methodology.

[4] Fraud indicators in this automobile insurance context are often binary response statements about the circumstances of a claim, with a positive response indicating increased likelihood of fraud (Canadian Coalition Against Insurance Fraud, 1997).

the EM algorithm and imputation of missing data techniques could be used in this statistical parametric approach (see Rubin, 1976).

However, several major differences exist between the types of problems addressed by fraud investigators investigating bodily injury claims and the types of problems solved by the above statistical techniques. Some of the major differences impact the usefulness of the statistical mixture methodology for fraud detection and are described below.

1. **The type of data:** Often claim files contain data collected in a categorized fashion or variables of interest categorically grouped into a fixed number of ordinal categories. Hence, the data requiring analysis by fraud investigators are, statistically speaking, discrete and ordinal, but with no natural metric scale. (The useful variable obtained from the subjective assessments of a "claims-wise" claimant is an example [Weisberg and Derrig, 1998]). Most standard parametric statistical methods require interval-level data, often continuous and normally distributed (however, see Rubin, 1976).

2. **The model:** The underlying stochastic processes that give rise to the unclassified observations for the two groups (fraud versus nonfraud) are generally of unknown parametric form. In particular, a simple multivariate stochastic model with only a few parameters to estimate from each group cannot be reasonably assumed at the advent of the analysis.

3. **The classification function:** The fraud investigator not only has an interest in being able to discriminate between the groups, but also wants to obtain a one-dimensional "suspicion level score" for each claim. The investigator desires that this score allow for ranking claim files and for performing a correlation analysis with each of several exogenous variables, such as demographics and behavioral data or fraud assessment records of the claims adjusters. Thus, for fraud investigation, the classification *function* is at least as important as the classification *rule,* since it can lead to the potential for obtaining a measure of external validity of the methodology. In most clustering techniques or Kohonen feature maps (which also use unsupervised learning), an ordinal score is not associated with each claim file to indicate how strongly each claim is associated with membership in the fraud group. This ability to obtain claim scores capable of being used by the investigators and adjusters in their analysis and decision making can aid significantly in classifying claims into fraud and nonfraud groups.

4. **The validity of the instrumental variables:** Since each variable costs money and time to gather, the fraud investigator is also generally interested in determining the relative worth of each predictor variable for discriminating between the fraud and nonfraud groups. This would be the analog of a coefficient in a regression, or the standardized discriminant coefficient in a supervised learning context were the dependent variable actually an observable dichotomous variable. This situation is desired because certain variables are more costly or time-consuming to obtain (such as depositions and independent medical examinations) and should be evaluated relative to their worth for distinguishing fraud from nonfraud claim files. Moreover, adding variables shown to be of dubious value can actually decrease the ability of the analysis to distinguish cases.

We shall show mathematically and empirically that PRIDIT analysis (using principal component analysis in conjunction with RIDIT scoring) achieves the above goals even though the technique does not presuppose the delineation of group membership prior to analysis (uses unsupervised learning) and uses rank-ordered categorical data that may not be interval level.

In the next section, we present the mathematical foundations of the PRIDIT methodology, discussing the scoring of qualitative or quantitative ordinal variables. We show how to use PRIDIT analysis to obtain a theoretically justified measure of "variable discriminatory power" for each individual fraud indicator variable within a claim file and an overall individual fraud suspicion-level score for each claim file. We then demonstrate how to attain all the goals of the analysis described in the Introduction. We then present empirical results of an application of PRIDIT to fraud assessment of bodily injury claims in automobile insurance, with a discussion and conclusions. Observe that the PRIDIT methodology has been used successfully in a number of other fields, including epidemiology.[5]

AN OVERVIEW OF PRIDIT FOR DISCRETE ORDINAL DATA

PRIDIT analysis is, essentially, a technique that scores the contents of a claim file on a set of predictor (independent) variables individually and as a whole. Through a computational algorithm, all claim files can be ranked in order of decreasing suspicion levels and, if desired, can be assigned to group membership (fraud/nonfraud) on the basis of this scoring. Simultaneously, and equally important, a measure of individual predictor variable discriminatory power can be obtained. Details on the mathematical foundations of the algorithms are given in Appendix A.

ASSIGNING NUMERICAL SCORES TO THE CATEGORIES OF QUALITATIVE ORDINAL VARIABLES

The basic task faced in using PRIDIT analysis is to develop a relative ranking of each of the claim files according to an underlying latent variable (fraud likelihood in this case) when the criterion variable (actual determination of fraud or nonfraud) is not observed even for a subsample. Various scoring techniques for ranked ordered categorical data have been investigated both empirically and theoretically by Brockett (1981), Brockett and Levine (1977), and Golden and Brockett (1987). Natural integer scoring (that is, increasing integer scoring wherein one simply assigns, for example, the numbers 1, 2, 3, 4, and 5 to the five permissible categorical responses for a variable) is a usual "default" method used by most researchers and practitioners who then apply classical statistical methods. In reality, however, the use of these classical statistical methods presupposes that the underlying input data are interval-level metric data (often with a specific statistical distribution as well), and hence presupposes that this natural integer scoring method creates interval-level metric data with "equal interval" spacing between categories. In fact, natural integer scoring may lead to qualitatively different results from, say, scoring questions to approximate normality or some other scoring method when the data are ordinal and non-interval-level. When used in statistical analysis, integer scoring can result in different variables being labeled as "significant" than would occur if a different scoring method were used on the same data set. In

[5] A brief description of an epidemiological study using this methodology is given subsequently.

an empirical study, Golden and Brockett (1987) found that another scoring method called RIDIT scoring, introduced by Bross (1958), produced the best results for several different distinct types of standard statistical analysis and, overall, was a superior scoring method for rank-ordered categorical variables, outperforming natural scoring in several respects.

In choosing a scoring mechanism incorporating the ranked nature of the responses and also the empirical response probabilities observed for the categories, we follow the development outlined in Brockett and Levine (1977) and Brockett (1981). This section summarizes this development. Intuitively, the scoring objective is to quantify the level of fraud suspicion produced by a categorical characterization of a particular indicator variable in the claim file. In addition, one desires to simultaneously obtain an overall fraud suspicion score for each entire claim file.

Let k_t denote the number of ranked response categories available for fraud indicator variable t, and denote the observed response proportions for the entire set of claim files by $\hat{p}_t = (\hat{p}_{t1}, \ldots, \hat{p}_{tk_t})$. Assume that the response categories are ordered in decreasing likelihood of fraud suspicion so that a higher categorical response indicates a lesser suspicion of fraud. For the categorical option i to variable t, assign the numerical value or score:

$$B_{ti} = \sum_{j<i} \hat{p}_{tj} - \sum_{j>i} \hat{p}_{tj} \quad i = 1, 2, \ldots, k_t. \tag{1}$$

We shall call this score the RIDIT score for the categorical response value.[6]

This procedure transforms any set of categorical responses into a set of numerical values in the interval $[-1, 1]$ which reflect the relative "abnormality" of the particular response. For example, for a binary response fraud indicator variable (in natural integer scoring, we might write that 1 = yes and 2 = no) a "yes" might be more indicative of potential fraud than a "no." Assuming, for example, that 10 percent of the claim files indicate a "yes" on this variable and 90 percent a "no" response, calculate the scores B_{t1} ("yes") = -0.9 and B_{t2} ("no") = 0.1. This scoring mechanism then produces numerical values to assign to each category—we would use -0.9 instead of 1 and 0.1 instead of 2 in the analysis. Just like natural integer scoring, the RIDIT score calculated in Equation (1) increases as the likelihood of fraud decreases, but unlike natural integer scoring, it also reflects the extent or degree to which the particular response is abnormal, and in which direction. Another binary response fraud indicator variable with 50 percent of the claim files indicating a "yes" on this variable and 50 percent a "no" response would have resulted in the same natural integer scores as above, but would result in RIDIT scores of -0.5 and 0.5, respectively, indicating to the analyst that a "yes" on the first indicator variable is more abnormal than that of a "yes" on the second indicator variable (-0.9 versus -0.5). Table 1 shows how the numerical transformation works for the medical treatment indicator variables used in the Weisberg and Derrig (1999) study.[7]

Effectively, this scoring method produces a variable score for a claim indicator variable which is positive when most claims result in a "lower" ranked category for this

[6] This score is actually a linear transformation of the RIDIT scoring method first introduced by Bross (1958) for epidemiological studies, but the version in Equation (1) is more convenient for our analysis.

[7] The description of all fraud indicators used in the study is attached as Appendix B.

TABLE 1
Computation of PRIDIT Scores

Variable	Variable Label	Proportion of "Yes"	B_{t1}("Yes")	B_{t2}("No")
Large number of visits to chiropractor	TRT1	44%	−0.56	0.44
Chiropractor provided 3 or more modalities on most visits	TRT2	12%	−0.88	0.12
Large number of visits to a physical therapist	TRT3	8%	−0.92	0.08
MRI or CT scan but no inpatient hospital charges	TRT4	20%	−0.80	0.20
Use of "high-volume" medical provider	TRT5	31%	−0.69	0.31
Significant gaps in course of treatment	TRT6	9%	−0.91	0.09
Treatment was unusually prolonged (> 6 months)	TRT7	24%	−0.76	0.24
Independent medical examiner questioned extent of treatment	TRT8	11%	−0.89	0.11
Medical audit raised questions about charges	TRT9	4%	−0.96	0.04

variable (that is, most claims are more likely to be fraudulent than the instant claim) and negative if most claim files have a "higher" ranked response category (are more likely to be nonfraudulent than the instant claim) for that variable. Note, for example, that a "yes" on TRT9 in Table 1 is much more extreme and indicative of fraud than is a "yes" on TRT1, even though both would have the same natural integer score.

Moreover, consistent with the ranked-order categorical nature of the variables, the scoring method is monotonically increasing (higher numerical scores corresponding to higher ordered categorical classification options and hence a higher likelihood of being nonfraudulent) and each score is "centered" overall so that the expected value $\sum_i \hat{P}_{ti} B_{ti} = 0$ for each fraud indicator variable (t). Unlike natural integer scoring, however, this scoring method does not presuppose an equal distance between the categories. Together with some other intuitively reasonable assumptions,[8] one can prove that these characteristics of a scoring method essentially characterize the scoring system B_{ti} (see Brockett, 1981; Brockett and Levine, 1977).

Since all variable scores are scaled to the same [−1,1] scale, predictor variables with vastly different numbers of potential categories into which a claim might be classified become readily comparable so that a high answer on a question with ten classification options is not viewed as twice as influential on a summative overall fraud suspicion score as a high classification option on a variable with five classification

[8] The assumptions leading to B_{ti} are detailed in Brockett and Levine (1977) and Brockett (1981), and include the assumption that B_{ti} should increase in the categorical option (i).

options. In fact, Brockett (1981) proves mathematically that our scoring scheme makes the empirical frequency curve for the variable as close to uniform on $[-1, 1]$ as possible, using the Kolmogorov-Smirnof distance measure, so all the predictor variables are initially on equal footing in terms of distribution. This scoring technique is analogous to ipsative methods (Clemens, 1956), but without the necessity of assuming equal intervals between categories. It eliminates the necessity of assigning integer values in an ad hoc fashion (as is done with natural integer scoring) and improves the statistical characteristics of the resulting scored data for subsequent standard statistical analysis, whatever it is. More important for the purposes of this article, it opens the door to creating an entirely new methodology for quantitative fraud detection for the entire claim file. This is discussed in the next section.

ASSESSING THE DISCRIMINATORY POWER OF THE PARTICULAR FRAUD VARIABLES AND OBTAINING OVERALL FRAUD SUSPICION SCORES FOR ENTIRE CLAIM FILES

Let $\boldsymbol{F} = (f_{it})$ denote the matrix of individual PRIDIT variable scores for each of the $t = 1, 2, \ldots, m$ variables, for each of the $i = 1, 2, \ldots, N$ claim files—that is, $f_{it} = B_{tk}$ if claim file i contains categorical response level k to variable t. Obtain an overall suspicion score for each claim file by simply adding the respective individual variable scores.[9] In matrix notation, let $\boldsymbol{W}^{(0)} = (1, 1, \ldots, 1)'$, the prime denoting transpose. Then the vector of simple overall summative fraud suspicion scores obtained for each claim file in matrix notation is $\boldsymbol{S}^{(0)} = \boldsymbol{F}\boldsymbol{W}^{(0)}$. Now, by taking the normalized scalar product of the set of claim file overall summative fraud suspicion scores with their individual variable t scores, we get a measure of consistency of indicator variable t with the overall fraud suspicion scores for the claim files (see Sellitz, 1964). This measure is similar in nature to the Cronback α measure of reliability used in questionnaire analysis to assess the consistency of individual questions with the overall questionnaire score. Thus, $\boldsymbol{W}^{(1)} = \boldsymbol{F}'\boldsymbol{S}^{(0)}/||\boldsymbol{F}'\boldsymbol{S}^{(0)}||$ can be viewed as a system of "weights" for the individual variables in the set of claim files, where the components of $\boldsymbol{W}^{(1)}$ give the normalized product of the indicator variable t with overall claim file scores and measure the consistency of the individual variable being weighted to the overall score for the claim file. A larger value for $\boldsymbol{W}^{(1)}$ indicates a larger consistency of this variable t with overall suspicion level determined for the entire claim file.

It now makes sense, since we know which variables are more consistent, or "better judges," of overall fraud suspicion for the set of claim files, to give higher weight in our analysis to these "better" variables and hence to calculate a "weighted" claim file score for each claim file, giving higher weight to better judges (fraud variables) as follows. Using the components of $\boldsymbol{W}^{(1)}$ as variable weights yields a weighted overall vector of fraud suspicion scores $\boldsymbol{S}^{(1)} = \boldsymbol{F}\boldsymbol{W}^{(1)}$ for the set of claim files. However, there is no need to stop now. Using this "better" assessment of overall fraud suspicion for each claim, we can now obtain an ever better measure of overall fraud suspicion score for the claim file by "correlating" the individual scores with this new better overall

[9] This initial value of suspicion scores corresponds to the naive model of counting the number of fraud indicators. This simple counting method for indicator variables is common in several areas (such as in an early warning system for insurer insolvency used previously by the Texas Department of Insurance). As we shall see, this simple summative scoring method for claim files can be improved by further analysis.

score. This in turn can be correlated again with the individual variable scores to get even better weights, $\boldsymbol{W}^{(2)} = \boldsymbol{F}'\boldsymbol{S}^{(1)}/||\boldsymbol{F}'\boldsymbol{S}^{(1)}||$. New weights can then be used to get a better weighted summative claim file score, and this new set of overall fraud suspicion scores can then again be correlated with the individual variable t scores to get new weights, and so on.

The theorem in Appendix A ensures that this mathematical process converges, and it shows that this limiting weight for fraud indicator variable t is actually proportional to a discriminatory power measure A_t developed in Appendix A. Specifically, the limiting variable weight $\hat{\boldsymbol{W}}^{(\infty)}$ is the first principal component of $\mathbf{F}'\mathbf{F}$, which is a consistent estimate of principal component $\hat{\boldsymbol{W}}^{(\infty)}$ of $\mathbf{E}[\mathbf{F}'\mathbf{F}]$, the tth component of which is explicitly

$$W_t^{(\infty)} = \frac{A_t}{(\mu_1 - U_{tt})\sqrt{\sum_{s=1}^{m} A_s^2/(\mu_1 - U_{ss})^2}}, \tag{2}$$

where μ_1 is the largest eigenvalue of $E[\boldsymbol{F}'\boldsymbol{F}]$, and for each s, U_{ss} is the uniqueness variance in a factor analysis model of RIDIT scored fraud suspicion variables $E[\boldsymbol{F}'\boldsymbol{F}]$. Here

$$\mathrm{A_t} = \sum_{\mathrm{i=1}}^{\mathrm{k_{t-1}}} \sum_{\mathrm{j>i}} \left\{ \pi_{ti}^{(1)}\pi_{tj}^{(2)} - \pi_{ti}^{(2)}\pi_{tj}^{(1)} \right\}, \tag{3}$$

where $\pi_{tj}^{(1)}$ is the proportion of the fraud or group 1 claims that fall into category j on fraud indicator variable t and $\pi_{tj}^{(2)}$ is the proportion of the nonfraud or group 2 claims that fall into category j on fraud indicator variable t.

The theorem proven in Appendix A is interesting for several reasons. Since the limiting weights are proportional to $\mathrm{A}_t/(\mu_1 - U_{tt})$, and because we may estimate this limiting weight by $\hat{\boldsymbol{W}}^{(\infty)}$, the first principal component of $\boldsymbol{F}'\boldsymbol{F}$, μ_1 by $\hat{\mu}_1$, the largest eigenvalue of $\boldsymbol{F}'\boldsymbol{F}$, and the uniqueness variances via a factor analysis subprogram yielding the diagonal uniqueness variance matrix $\hat{U}$, we may easily estimate the vector of relative question discriminatory power values by $(\hat{\mu}_1 I - \hat{U})\hat{W}^{(\infty)}$. Thus, as long as an underlying discriminating combination of variables exists, up to a common multiplier, an estimate of A_t, the contingency table measure of association for variable t, can be found using principal component and factor analysis without ever having to know which claim files belong to which groups and in which proportions! Our measure of variable importance is thus validated.

It should be emphasized that using principal component analysis to weight variables is not new (Daultrey, 1976). However, in general there is no guarantee that the result will have any meaningful statistical interpretation. We have shown here that by using our particular scoring system a useful interpretation exists in terms of the discriminatory power A_t of the variable and also an interesting connection between contingency table analysis, principal component analysis, and factor analysis. There is no guarantee that this occurs with other scoring systems, such as the natural integer scoring so commonly used.

To aid in fixing ideas, we interject here an example from an epidemiological study of the health effects of rural-to-urban migration in Senegal (Benyouseff et al., 1974). One goal of the study was to test the idea that people who do not adapt to urban life are at greater risk for disease. To this end, a questionnaire was designed consisting of 59 ordinal questions about different types of social behavior that were indicators of adaptation, such as regular employment, frequenting clubs, use of radios, and language spoken. The questionnaire was meant to classify the migrants into two groups, adapters and non-adapters, according to whether the scores were positive or negative. An example of a question was "Who are your close friends?" with these ordered categorical responses: (1) only those from my village, (2) only from my tribe, (3) about the same number from my tribe and the town, and (4) mostly from the town. Three hundred migrants were interviewed. The resulting PRIDIT analysis produced 7 questions out of the 59 with weight A_t not significantly different from 0. By examining the relationship of PRIDIT scores with other health-related indicators, the non-adapters were found to be at significantly greater risk for disease than the general population of migrants.

This study illustrated three uses of the PRIDIT analysis also relevant to the analysis of fraud. First, it enabled a reduction in the number of factors (variables) needed to divide respondents into relevant groups, a boon to future investigations into this matter; second, it provided a criterion for group membership through the score of each individual. Finally, it provided a quantitative measure of a qualitative variable (adaptability); this quantitative measure provided the capability to determine correlations with other quantitative measure such as prevalence of a disease and morbidity.[10]

CLASSIFYING CLAIMS BY PRIDIT SCORES

With respect to classification, consider two cases for the proportion of group 1 claims, θ known and θ unknown. When θ is known, arrange the N claim files via their one-dimensional scores $S = \sum_{t=1}^{m} W_t^{(\infty)} X_t$ and then classify the first Nθ claim files into the high fraud suspicion group 1. Here X_t is the claim file's calculated score obtained for variable t, $X_t = \sum_{i=1}^{k_t} B_{ti} I[\text{cateogry } i \text{ given}]$ where I_A is the indicator of the set A.

If θ is unknown, break the two groups according to positive or negative overall scores and classify claims into the low-suspicion group if the overall fraud suspicion score is positive.[11] Refer to these two methods as the ranking method and the algebraic sign method, which roughly correspond to priors proportional and priors equal, respectively, in a discriminant analysis (although, of course, PRIDIT analysis does not require data on actual group membership). In any event, at this stage the claim files have all been linearly ordered in terms of their potential suspicion levels and simultaneously a measure of worth for each individual fraud indicator has been obtained, as desired at the onset of the analysis. Whether using the ranking method or the algebraic sign method, the analyst can quickly dispose of numerous claim files that are apparently non-fraudulent, pay the claims, and focus attention and resources in a hierarchical

[10] Note that the use of regression or other supervised learning type techniques would be impossible in this unsupervised learning environment (as is also the case in fraud for bodily injury claims). Only exogenous validation is possible.

[11] Rigorous statistical methods also exist for estimating θ and for reassigning membership in the two groups based on the estimate. See Levine (1991) for details. However, by this methodology, only those with scores close to zero may be reassigned. For practical purposes, this reassignment does not affect the procedures proposed here.

manner on the most suspicious claims for subsequent investigation. In many applications the ability to dispose of (pay) the multitude of apparently valid claims is an important cost-effective step forward and one which can be automated (via PRIDIT analysis) for even further economic savings.

AN EXAMPLE USING THE TREATMENT VARIABLES FROM THE AUTOMOBILE INSURERS BUREAU

The study reported in Weisberg and Derrig (1998) tracked 65 fraud indicators across 127 bodily injury claims (see Appendix B for a listing).[12] In addition to the treatment variables shown above in Table 1 (9), indicator variables pertained to the accident (19), the claimant (11), the insured (8), the injury (12), and lost wages (6). For each of 127 claim files, senior claim managers (adjusters) recorded the presence (yes) or absence (no) of each of the 65 fraud indicators. In addition, for comparison purposes, adjusters recorded the overall assessment or suspicion scores for each individual claim file on a zero-to-ten scale for each fraud indicator category as well as an overall claim suspicion score. Senior investigators from the Insurance Fraud Bureau of Massachusetts (investigators) independently reviewed the same files, together with the adjusters' indicator choices, and produced their own zero-to-ten suspicion assessment of each claim file. These sample data allow illustration of the PRIDIT technique when θ is known,[13] along with a consistency test of model suspicion scores.

Table 2 shows the PRIDIT $\hat{W}^{\infty}$ weights for the nine treatment[14] variables described in Table 1. Also shown are the regression model weights using the adjusters' suspicion score as the dependent variable and natural integer scoring for the variables.[15]

Notice that indicator variables TRT6 and TRT7 carry the heaviest weight in scoring the 127 claim responses on their own response pattern (the factor analysis estimation of weights by PRIDIT), but they carry no significant weight in aligning the response pattern to the suspicion scores from the adjuster coders. This is the same type of inconsistency between fraud indicators and suspicion score assignments revealed by unsupervised neural network clustering (Brockett et al., 1998). This type of inconsistency typifies systematic weaknesses when subjective, or multiple observer, data are used.[16] Variables may be "missed" or "understated" by human adjusters due to a

[12] These fraud indicator and suspicion score data also underlie the Brockett et al. (1998) neural network clustering of claims.

[13] In this case, θ was known in the sense that 62 claims were assessed to be fraud by at least one coder (27 by adjuster but not investigator, 16 by investigator but not adjuster, and 19 by both adjuster and the investigator) and 65 claims were randomly chosen from claims assessed by all coders as nonfraud (Weisberg and Derrig, 1998).

[14] Appendix B shows the 65 weights when considered by category and when all indicators are considered at once. Table 1 and Appendix B give the verbal description of the variables.

[15] Significant regression coefficients are used as weights in a scoring model on claims subsequent to the training (regression) data.

[16] The internal inconsistency of the claims adjusters and claims investigators is also noted in Brockett et al. (1998) and occurs, in part, because investigators and adjusters focus differently on variables. The PRIDIT method is, however, nonsubjective and does not focus on specific variables, but rather gives higher weight to variables that seem to distinguish the high group (nonfraud) from the low group (fraud), yielding an ultimate claim file score. In addition, the regression weights reflect, in part, the subjective evaluation of the investigator or adjuster. In

Table 2
Weights for Treatment Variables

Variable	PRIDIT Weights $W^{(\infty)}$	Regression Weights
TRT1	0.30	0.32***
TRT2	0.19	0.19***
TRT3	0.53	0.22***
TRT4	0.38	0.07
TRT5	0.02	0.08*
TRT6	0.70	−0.01
TRT7	0.82	0.03
TRT8	0.37	0.18***
TRT9	−0.13	0.24**

Regression significance shown at 1 percent (***), 5 percent (**), or 10 percent (*) levels.

number of subjective reasons. Such inconsistencies also highlight the need for a "tracking" mechanism to reveal when significant inconsistencies arise (new response patterns compared to old response patterns or to old model scores). That issue is addressed in the next section.

Table 3 shows a sample of the first ten of 127 claims, their transformed treatment variable values, and their final weighted score. Negative scores would indicate class 1 (potential fraud) claims. Note that in addition to the two-way classification, the claims can be ranked in a manner similar to ranking by suspicion scores.

Thus, in a totally objective and automated manner, one can decide to investigate claim 3 first, claim 7 next, and pay the rest of the 10 claims (or if resources allow, investigate files in increasing PRIDIT score order until resources are exhausted). Note that this evaluation is made on the basis of the data alone, with no need to hire costly adjusters or investigators to examine all files, score them, and rank-order them.[17] Significant economic savings as well as internal consistency can be so achieved. As an adjunct to

fact, the turnover rate of adjusters and investigators (whose subjective opinions are dependent upon training and economic conditions other uncontrollable variables) means that when a model such as regression is trained using the subjective opinions of a particular evaluator, the change of this evaluator may make the parametric "trained" model no longer applicable to a newly hired evaluator. PRIDIT analysis does not have this weakness and can even be used for continuous monitoring to determine when significant changes in the underlying claim variable weights or RIDIT scores have occurred, necessitating a recalibration of the model. The PRIDIT weights can thus be thought of as more reliable.

[17] Note, however, that both objective fraud indicators and subjective fraud indicators exist, and that experienced professionals are still needed to decide how to score certain of the individual subjective fraud indicator variables, but the level of expertise needed to score the individual variables is less. Moreover, psychological literature has established that raters are more reliable for rating particular attributes or specific components of an overall multiattribute problem than they are for performing the overall problem, so there is still a gain to be gotten here.

TABLE 3
PRIDIT Transformed Indicators, Scores, and Classes

Claim	TRT1	TRT2	TRT3	TRT4	TRT5	TRT6	TRT7	TRT8	TRT9	Score	Class
1	0.44	0.12	0.08	0.2	0.31	0.09	0.24	0.11	0.04	0.07	2
2	0.44	0.12	0.08	0.2	−0.69	0.09	0.24	0.11	0.04	0.07	2
3	0.44	−0.88	−0.92	0.2	0.31	−0.91	−0.76	0.11	0.04	−0.25	1
4	−0.56	0.12	0.08	0.2	0.31	0.09	0.24	0.11	0.04	0.04	2
5	−0.56	−0.88	0.08	0.2	0.31	0.09	0.24	0.11	0.04	0.02	2
6	0.44	0.12	0.08	0.2	0.31	0.09	0.24	0.11	0.04	0.07	2
7	−0.56	0.12	0.08	0.2	0.31	0.09	−0.76	−0.89	0.04	−0.10	1
8	0.44	0.12	0.08	−0.8	−0.69	0.09	0.24	0.11	0.04	0.02	2
9	−0.56	−0.88	0.08	0.2	0.31	0.09	0.24	0.11	−0.96	0.05	2
10	−0.56	0.12	0.08	0.2	0.31	0.09	0.24	0.11	0.04	0.04	2

the evaluation by experienced insurance adjusters or investigators, the PRIDIT analysis can be used to create a "first pass" through the data to better focus the investigators' or adjusters' attention to certain claim files for analysis and can give quantitative back-up to their own subjective assessments, helping them do their own jobs better.

CONSISTENCY TEST FOR SUBJECTIVE RANKINGS

The PRIDIT technique provides several relevant output values that can be used to exogenously test the consistency of subjective rankings, or classification of claims, with fraud indicator data.[18] Some examples are as follows:

1. Subjective scores (ranks) for claims files obtained from investigators or adjusters can be compared to PRIDIT scores (ranks) by Pearson (Spearman) correlation.
2. Subjective bivariate class membership for claims (fraud/nonfraud) can be compared to PRIDIT classes (negative/positive) by contingency table analysis.
3. Fraud indicator discriminatory power for subjective scores and classes (regression coefficients) can be compared by correlation to PRIDIT weights to distinguish internal discrimination (PRIDIT) from external, possibly inconsistent, discrimination.
4. Exogenously derived relationships (regressions) of fraud indicator patterns to suspicion scores or classes can be compared to PRIDIT scores, ranks, or classes by correlation and contingency table analysis to reveal the extent of consistency in subjective evaluations of indicators and subjective scores, ranks, or classes.

We proceed to illustrate these consistency tests using the Automobile Insurers Bureau (AIB) fraud indicator data and the overall suspicion scores developed from (1) the

[18] These types of consistency checks apply to any other similar data scheme for which PRIDIT analysis is feasible and a separate, but related, ranking or classification is present (such as when exogenous classifications exist, as with insurance investigators or adjusters).

TABLE 4
AIB Fraud Indicator and Suspicion Score Data

	Spearman Rank Correlations				
Pearson Score Correlations	PRIDIT	Adjuster Natural Score	Investigator Natural Score	Adjuster Regression Score	Investigator Regression Score
PRIDIT	1.00	0.60	0.49	0.78	0.64
Adjuster Natural Score	0.56	1.00	0.52	0.81	0.58
Investigator Natural Score	0.48	0.51	1.00	0.52	0.75
Adjuster Regression Score	0.73	0.81	0.49	1.00	0.69
Investigator Regression Score	0.65	0.54	0.75	0.63	1.00

adjusters' assessment, (2) the investigators' assessment, (3) two ten-variable regressions of adjusters' and investigators' scores on the fraud indicator data,[19] and (4) the overall PRIDIT weights, scores, and (fraud/nonfraud) classes.[20] Table 4 displays the Pearson score correlations below the diagonal and Spearman rank correlations above the diagonal.

As supervised learning models with the same data, the Adjuster and Investigator Natural Integer Scores and Regression Scores are highly correlated, as they should be, under both measures: 0.81 and 0.75, respectively. In addition, in spite of the PRIDIT methodology being an unsupervised methodology, the PRIDIT/adjuster regression scores also appear as highly consistent, 0.73–0.78, since the same coders provide both subjective fraud indicator assessments and overall suspicion levels and both regression and PRIDIT smooth out inconsistencies. Note that the comparison of the PRIDIT and Adjuster Natural Integer Scores is much less consistent (0.56–0.65) without the smoothing effect of the regression procedure.[21] Still, this correlation range is very encouraging given that the PRIDIT method is automated and does not use the accumulated wisdom, gained from experience, of trained claims adjusters. However, in reality PRIDIT does use the experience of the insurance professionals in constructing the appropriate indicator variables, and this is the source of its strength.

Meanwhile, the PRIDIT/investigator scores are somewhat less consistent (0.65–0.64, comparing regression scores or 0.48–0.49, comparing natural scores). The comparison of adjuster and investigator regression score ranking consistency (0.69) is quite similar to the PRIDIT/investigator regression ranking comparison (0.64). Note also that the correlation exhibited between PRIDIT scores and adjusters' scores are of the same magnitude as the correlation between the investigators' and the adjusters' scores, but

[19] See Weisberg and Derrig, 1998, Tables 3 and 5.

[20] The PRIDIT weights are shown in Appendix B. An Excel Workbook is available from the authors containing the data and calculations for these comparisons.

[21] The use of natural integer scoring produces lumpy rankings with many ties by virtue of the discretized zero-to-ten scores.

TABLE 5
AIB Fraud Indicator and Suspicion Score Data Consistency

	Spearman Rank Correlations				
Pearson Score Correlations	PRIDIT	Adjuster Natural Score	Investigator Natural Score	Adjuster Regression Score	Investigator Regression Score
PRIDIT	FULLC	MODC	LOWC	FULLC	MODC
Adjuster Natural Score	MODC	FULLC	MODC	FULLC	MODC
Investigator Natural Score	LOWC	MODC	FULLC	MODC	FULLC
Adjuster Regression Score	MODC	FULLC	LOWC	FULLC	MODC
Investigator Regression Score	MODC	MODC	FULLC	MODC	FULLC

without the same level of labor costs. Given the lower level of correlation between the adjusters' and investigators' natural scores (0.56–0.60), this correlation reflects the intrinsic difficulty of identifying fraud, even for trained and experienced professionals. In this light, PRIDIT does well.

In the spirit of the fraud suspicion categories (none, low, moderate, high) used in the Weisberg and Derrig studies (1991–1998), we adopt a four-category consistency scale of (absolute) quartile correlation values: no consistency (0–0.24), low consistency (0.25–0.49), moderate consistency (0.50–0.74), and full consistency (0.75–1.00). With these conventions, Table 4 is transformed into Table 5, showing consistency levels. We leave what "adequate" consistency is to the reader's judgment.[22]

Table 6 displays the four 2-way tables comparing fraud/nonfraud classes derived by the negative/positive assignment using PRIDIT weights, with fraud/nonfraud classes defined by splitting the ten-point adjuster or investigator claim file suspiciousness score into two levels at the score 4. Highly (7–10) and moderately (4–6) suspicious claims define the fraud class, while the no-suspicion (zero) and low-suspicion (1–3) claims define the nonfraud class. Consistency is measured here by the cross-product, or odds ratio, α (Bishop et al., 1977; Rudas, 1998).[23] A large ratio indicates consistency (dependence), while zero indicates no consistency (independence). Confidence intervals for α at the 95 percent level are also shown (Upton, 1978). In all cases the cross-product ratio is significant at the level 0.05 (that is, the confidence interval excludes the value $\alpha = 1$).

The weights $\hat{W}^{(\infty)}$ that emerge from the PRIDIT procedure determine the relative importance of each indicator in determining the internally consistent PRIDIT classes.

[22] We suggest correlations of 0.50 and above (moderate to full consistency).

[23] The cross-product, or odds ratio, for a 2 × 2 table of counts or probabilities (m_{ij}) equals $m_{11}m_{22}/m_{12}m_{21}$, assuming all cells are nonzero. Several additional statistics measuring similarity, such as Yule's Q, chi-squared, and Guttman's Lamda could also be considered (Upton, 1978).

Table 6
AIB Fraud Indicator and Suspicious Scores Classes

		Fraud/Nonfraud Classifications		
PRIDIT		Fraud	Nonfraud	All
Adjuster Natural Score	Fraud	33	13	46
	Nonfraud	29	52	81
	All	62	65	127
(α = **4.6**)[**2.1,10.0**]				
Adjuster Regression Score	Fraud	30	5	35
	Nonfraud	32	60	92
	All	62	65	127
(α = **11.3**)[**4.0,31.8**]				
Investigator Natural Score	Fraud	46	19	65
	Nonfraud	16	46	62
	All	62	65	127
(α = **7.0**)[**3.2,15.2**]				
Investigator Regression Score	Fraud	49	19	68
	Nonfraud	13	46	59
	All	62	65	127
(α = **9.1**)[**4.1,20.6**]				

Table 7
AIB Fraud and Suspicion Score Data—Top Ten Fraud Indicators by Weight

PRIDIT	Adjuster Regression Score	Investigator Regression Score
ACC3	ACC1	ACC11
ACC4	ACC9	CLT4
ACC15	ACC10	CLT7
CLT11	ACC19	CLT11
INJ1	CLT11	INJ1
INJ2	INS6	INJ3
INJ5	INJ2	INJ8
INJ6	INJ9	INJ11
INS8	TRT1	TRT1
TRT1	LW6	TRT9

Similarly, the regression coefficients provide higher weights for the most important indicators, given the related overall suspicion score. Table 7 shows the top ten fraud indicators from the three models: PRIDIT, Adjuster Regression, and Investigator Regression, when all 65 indicators are considered.

While only two indicators (CLT11 and TRT1) appear in the top ten of both the adjuster and investigator models, four out of the ten indicators picked by PRIDIT are also considered in the top ten by either the adjusters or the investigators (PRIDIT analysis shares these same two indicators in common with adjusters and investigators but has one additional indicator in common with each). In addition, the dispersion across indicator categories highlights both the importance of the multiple categories of indicators and the multicollinearity of the indicators. Moreover, note that the derivation of the regression weights for the Adjusters' Regression Model required hiring and paying senior claims managers (adjusters) to go through each of the 127 claim files to secure an overall claim file suspicion judgment, while PRIDIT did not use this accumulated wisdom (and expense). A similar comment applies to the Insurance Investigators' Regression Model.

In addition, while adjusters and investigators both find about 10 percent fraud in their work, they focus on a different 10 percent, with only a small overlap. This is because adjusters are looking for variables that assist in adjusting claims (reducing payments), while investigators are more focused on variables that indicate legal determinants of fraud. That is, they have different perspectives. The PRIDIT method uses it all.

CONSISTENCY TESTS FOR TIME VARIATION

It is important for any fraud detection system to recognize when significant changes occur over time. The groups to be detected can change as a result of the deterrent value of prior detection efforts. The indicator patterns may change due to the use of more or less highly trained claim adjusters. Unless the relationship of the indicator patterns and suspicion scores, however determined, is stationary (unchanging with time), scoring models estimated with prior data may suffer deteriorating accuracy. Again, this is another reason why we cannot use the standard "test group" methodologies. Fraud behavior is constantly evolving as perpetrators learn and adjust. The PRIDIT methodology supports a testing procedure to evaluate changes over time.

As an example, suppose regression scoring models were in use at the time the auto injury data above were collected. The consistency tests above could be used to test whether the new adjuster regression models are consistent with the old models and with the new indicator data. Old PRIDIT fraud/nonfraud claims could be compared to new PRIDIT claims as well. Table 8 shows an "update" to Table 4 that tests the relationship of two "old" models, ten- and 20-variable regression models,[24] to the current adjusters' raw data and suspicion score regression model.

Note that consistency between the two "old" regression models built on the same data is extremely high (0.96–0.97). The consistency between the old model ranks and either new model rank is quite high (0.79–0.81), about the same consistency level relating the natural score ranks to the (old) regression ranks (0.81). The PRIDIT/regression rank comparison does deteriorate (0.78–0.69), but less so when compared to the more complex 20-variable regression model ranks (0.78–0.73). Moreover, the level of correlation (0.69–0.73) remains fairly constant over time, which, since the PRIDIT method can be automated and is thus less expensive, is encouraging concerning the use of the unsupervised learning PRIDIT methodology.

[24] The models were estimated on nearly identical fraud indicators applied to data four accident years removed (1989 versus 1993) from the 127 claim data. For this illustration, the regression models derived using the 1993 data are treated as the old models.

TABLE 8
AIB Fraud Indicator and Suspicion Score Data

	Spearman Rank Correlations				
Pearson Score Correlations	PRIDIT	Adjuster Natural Score	Adjuster Regression Score	10-Variable Regression Score	20-Variable Regression Score
PRIDIT	1.00	0.60	0.78	0.69	0.73
AdjusterNatural Score	0.56	1.00	0.81	0.63	0.64
Adjuster Regression Score	0.73	0.81	1.00	0.81	0.79
10-Variable Regression Score	0.67	0.57	0.70	1.00	0.96
20-Variable Regression Score	0.72	0.59	0.75	0.97	1.00

TABLE 9
AIB Fraud Indicator and Suspicion Score Data Consistency

	Spearman Rank Correlations				
Pearson Score Correlations	PRIDIT	Adjuster Natural Score	Adjuster Regression Score	10-Variable Regression Score	20-Variable Regression Score
PRIDIT	FULLC	MODC	FULLC	MODC	MODC
Adjuster Natural Score	MODC	FULLC	FULLC	MODC	MODC
Adjuster Regression Score	MODC	FULLC	FULLC	FULLC	FULLC
10-Variable Regression Score	MODC	MODC	MODC	FULLC	FULLC
20-Variable Regression Score	MODC	MODC	FULLC	FULLC	FULLC

Table 9 shows the levels of consistency according to our categorized criteria in Table 5. The criteria indicate the need to "update" may be worthwhile, but the gain is marginal at best. Similarly, the comparison between PRIDIT fraud/nonfraud classes and "old" and "new" classes can be formalized. Table 10 provides an update to Table 6.

These data show substantial deterioration (odds ratios change of 5.5/11.3) in the alignment between the indicator patterns (PRIDIT) and the ten-variable regression relationship. The comparison with a more complex 20-variable regression shows a more consistent (9.2/5.5) result. Neither change represents a (95%) significant change.[25] A similar comparison between the old and new PRIDIT scores would reveal indicator

[25] In practice, sample sizes would be chosen to be much larger than the 127-claim example shown here. Since the variation in the cross-product is a function of the sum of the reciprocals, 95 percent intervals for practical settings would be much smaller than shown here.

TABLE 10
AIB Fraud Indicator and Suspicious Score Classes

Fraud/Nonfraud Classifications				
PRIDIT		Fraud	Nonfraud	All
Adjuster Natural Score	Fraud	33	13	46
	Nonfraud	29	52	81
	All	62	65	127
(α = **4.6**)[**2.1**,**10.0**]				
Adjuster Regression Score	Fraud	30	5	35
	Nonfraud	32	60	92
	All	62	65	127
(α = **11.3**)[**4.0**,**31.8**]				
10-Variable Regression Score	Fraud	27	8	35
	Nonfraud	35	57	92
	All	62	65	127
(α = **5.5**)[**2.2**,**13.4**]				
20-Variable Regression Score	Fraud	35	8	43
	Nonfraud	27	57	84
	All	62	65	127
(α = **9.2**)[**3.8**,**22.6**]				

pattern changes. Again, the observer or decision maker must determine whether the detected changes warrant changing a fraud detector process in place.

CONCLUSION

We view the insurance fraud detection problem as a data gathering and data analysis problem. Claim adjusters and investigators gather different types of information (data) during the settlement of a claim. Historically, some characteristics or features of the claim have been labeled as fraud indicators. While traditional fraud indicators are binary-valued (true or false), other response items may naturally have a small number of categories (age group) or have a large number of potential responses (dollar size of loss) that can be segregated into a small number of bins.[26]

This article introduces the principal component analysis of RIDITs technique (PRIDIT) for use when the fraud indicators are ordered categorical variables. Our approach makes use of the matrix structure of the indicator responses and their correlations. The approach provides (1) scoring weights for the indicators that allow both an assessment of individual variable "worth" on a [−1,1] scale and a weighted summative suspicion score, (2) a partition of the sample of claim files examined into positive (nonfraud) and negative (fraud) consistent with the pattern of indicator responses, and (3) a ranking by overall score of the individual claim files that can be used to

[26] Many fraud indicators contain linguistic variables (old, low-valued car) that may be more accurately reported as fuzzy numbers, but that is beyond the scope of this article.

test consistency with scoring methods developed on other (prior) data sets or with the PRIDIT patterns of older response data. Estimation of the weights is shown to be equivalent to the first principal component of an $n \times n$ matrix where n is the number of fraud indicators. An advantage of this method over hiring fraud examiners or claims adjusters to individually examine and classify each individual claim file is that the PRIDIT procedure exposed is automated and hence more economical (due to labor costs) while being consistent with the labor-intensive manual examination method.

All the methods introduced here are illustrated in detail using a data set of automobile injury claims occurring in Massachusetts. Explicit correlation and contingency table criteria are explored to characterize consistency among response data sets and between responses and scoring models. Additional consistency measures could be formulated by analogy to sensitivity and specificity characteristics of 2×2 classification tables. Receiver operating characteristic (ROC) plots can be used to quantify the sensitivity/specificity tradeoff (Weisberg and Derrig, 1998). Finally, the natural setting of fraud indicators as linguistic variables may offer the opportunity to apply the many fuzzy set theoretic data analysis models, such as k -means fuzzy clusters or neural network methods such as self-organizing feature maps.[27] Further analyses using these methods are possible as well. For example, certain indicators arise (nonsubjectively) at the time of the accident or shortly thereafter, while others follow from investigative analysis later on. A useful sequel to this article is to use PRIDIT techniques on the early-arriving indicators to determine whether to proceed forward to collect the full range of indicators. In addition, PRIDIT scores can be correlated with (or built into a model to predict) subsequent outcome variables such as dollars saved on claims, likelihood of successful prosecution, and so on.

APPENDIX A

Discussion and Development of a Measure of Variable Discriminatory Power

Before discussion of the discriminatory power of our methodology, we begin with some background. Even without assuming any specific probabilistic structure for the fraud/nonfraud groups, some very interesting consequences of this scoring system are worth noting. First, the RIDIT score for the categorical response value (B_{ti}) is linearly related to the RIDIT scoring method introduced by Bross (1958), now commonly used in epidemiology (see Bertrand et al., 1980; Bross, 1960; Brown et al., 1975; Forthofer and Koch, 1973; Kantor et al., 1968; Selvin, 1977; Williams and Grizzle, 1972; Wynder et al., 1960). In fact, if R_{ti} is Bross's original RIDIT score for category i on variable t, then $B_{ti} = 2R_{ti} - 1$. Thus, heuristic justifications given by several authors for the use of RIDIT scoring (for example, Bross, 1958; Selvin, 1977) carry over directly to our situation, as well as the theoretical and empirical attributes described by Brockett (1981) and Golden and Brockett (1987).

The relationship between relative ranks, Wilcoxon test statistics, and RIDIT scores also carries over directly (see Selvin, 1977). In particular, one can show, assuming two groups, such as fraud and nonfraud, that the average RIDIT score, $\bar{B}_t^{(1)}$, for a fraud group member on variable t is $2W_1^*/N_1 - 2$, where W_1^* is the Wilcoxon rank sum

[27] Derrig and Ostaszewski (1995) and Brockett et al. (1998) provide examples of these approaches. Zimmerman (1999) provides a large collection of methods and applications.

statistic for comparing the fraud group against the nonfraud group responses to variable t (see Beder and Heim, 1990; Selvin, 1977).

Accordingly, our RIDIT scoring method yields a measure of "variable discriminatory power," assuming the group members are known (the supervised learning context). In fact, this scoring method and related rank scoring methods often do better than natural integer scoring for linear discriminant analysis of non-normal data (see Broffit et al., 1976; Conover and Iman, 1978; Golden and Brockett, 1987; Randles et al., 1978).[28] We emphasize, however, that in the automobile bodily injury fraud detection situation considered herein, we do not know group membership and hence Wilcoxon is not applicable (although PRIDIT analysis as introduced in this article is). Still, the intuitive appeal of the connection and the correspondence with known statistical methodologies under the additional assumptions of known group membership give greater confidence in the results in the situation of this article where group membership is unknown.

We now proceed to the problem of determining variable discriminatory power and discrimination between two groups when *we do not know* the criterion variable (that is, when the sample is "unclassified" with respect to group characteristic of likelihood of fraud versus nonfraud). Assume that N claim files are available containing ordinal valued categorical variables. Two groups are represented (the fraudulent claims and the nonfraudulent claims); however, we have no knowledge a priori concerning which individual claims belong to which group.

First, assume (valid on our application because of the construction and selection of fraud indicator variables) that the individual variables are constructed in such a manner that a univariate stochastic dominance relationship exists between the two groups on each variable (fraudulent claims tend to fall into lower categorical classifications than the nonfraudulent claims). Also assume that a single overall dimension exists on which a combination of the individual variables discriminates (this dimension may be called the fraud suspiciousness dimension). One can justify this assumption because of the way the variables were constructed using the expertise of the insurance fraud professionals.[29] A quick examination of the fraud indicator variables listed in

[28] Accordingly, if one is forced to use standard statistical methods, we would advocate using RIDIT scoring as a precursor to applying the statistical method in the situation wherein one encounters rank-ordered categorical but not interval-level data in a supervised learning context (that is, actual training samples are available from the fraud and nonfraud groups). One could, for example, extend regression or probit or logit analysis as has been used previously for fraud detection modeling to the situation wherein rank-ordered non-interval categorical fraud variables are used (see Golden and Brockett, 1987).

[29] There has been a long history of using these variables by experienced claims adjusters and investigators. Moreover, in those situations in which one has a categorical but non-ordinal variable, one can often rearrange the variable categories in such a way that the newly rearranged variable satisfies the stochastic dominance assumption. This can be done as follows: First run the PRIDIT analysis without this variable and then rank-order the resulting claim files in accordance with overall score. Then take the upper and lower quartiles of the ordered files and calculate a probability for each categorical response for each quartile. Finally, rearrange the categories in such a manner than the stochastic dominance relation holds. Since order means nothing in non-ordinal categorical data, this transforms the variable in a manner appropriate for the assumptions of the model to be valid.

Appendix B shows that these variables do indeed satisfy the above assumption for the bodily injury fraud detection problem under consideration in this article.

We shall call the group with a higher propensity to respond toward the lower end of the set of categorical classifications for a particular fraud indicator variable the fraud group and label it as group 1, while we call the group with a higher propensity to respond toward the upper end of the set of categorical classifications the nonfraud group, or group 2 (although precisely which claim belongs to which group is not assumed to be known at the onset of the analysis). The vector of observed response proportions for the k_t categories of variable t using all N claim files, p_t, may be modeled as a sample from a mixture from these two groups:

$$p_t = \theta\pi_t^{(1)} + (1-\theta)\pi_t^{(2)}, \tag{A1}$$

where for $q = 1$ and 2, $\pi_t^{(q)} = (\pi_{t1}^{(q)}, \ldots, \pi_{tkt}^{(q)})$ is the multinomial response probability vector for group q. Here $\pi_{tj}^{(1)}$ is the proportion of the fraud, or group 1, claims that fall into category j on fraud indicator variable t; $\pi_{tj}^{(2)}$ is the proportion of the nonfraud, or group 2, claims that fall into category j on fraud indicator variable t; and $\theta = N_1/N$ is the proportion of claims that belong to the fraud group. We will prove that the expected (suspicion) score for a claim file belonging to the nonfraud group is $(\theta - 1)A_t$ and the expected score for a fraud group claim file is θA_t, where

$$A_t = \sum_{i=1}^{k_t-1}\sum_{j>i}\left\{\pi_{ti}^{(1)}\pi_{tj}^{(2)} - \pi_{ti}^{(2)}\pi_{tj}^{(1)}\right\} \tag{A2}$$

is a measure of the individual fraud indicator variable t's discriminatory ability. Of course, in many applications (including those in this article), $\mathrm{N}_1,\theta,\pi_t^{(1)}$ and $\pi_t^{(2)}$ are all unknowns, so we shall have to find an alternative way to estimate A_t. We shall subsequently show that the first principal component of the matrix constructed from the respondents' RIDIT scores can be used to estimate A_t and hence complete the analysis.

Although $\pi_t^{(1)}$, $\pi_t^{(2)}$ and N_1 are unknown (and hence A_t and θ are not directly computable), the important quantity A_t deserves further discussion. As mentioned, the value of A_t indicates the discriminatory power of fraud indicator variable t in that $|\mathrm{A}_t| \leq 1$ with $\mathrm{A}_t = 1$ if and only if the variable discriminates perfectly between the fraud and nonfraud groups in the sense that for some category j, $\sum_{i=1}^{j-1}\pi_{ti}^{(1)} = 1$ and $\sum_{i=j}^{k_t}\pi_{ti}^{(2)} = 1$. If the variable still discriminates perfectly, but in the opposite direction, then $\mathrm{A}_t = -1$. If the variable does not discriminate at all in the sense that $\pi_{ti}^{(1)} = \pi_{ti}^{(2)}$ for $i = 1, 2, \ldots, k_t$, then $\mathrm{A}_t = 0$.

Since for each variable, we have a stochastic ordering of the cumulative distribution function of the two fraud/nonfraud groups (that is, a first-order stochastic dominance relationship), the size of A_t indicates ordinal discriminatory power and measures the expected between-group difference on variable t. Note also that the term inside the brackets of the summation defining A_t is the standard cross-product measure of association for 2×2 contingency tables (see Stuart and Ord, 1991). Thus, A_t is a new measure of association for a $2 \times k_t$ contingency table with ranked response categories which generalizes the classical 2 × 2 measure to our situation of general rank-ordered categorical variables.

When the true classification of each claim (fraud/nonfraud) is unknown, one cannot calculate directly the contingency table measure A_t mentioned in the body of the article. However, the relationship between relative ranks, Wilcoxon test statistics, and RIDIT scores does provide a connection between our scoring method and a measure of variable discriminatory worth when group membership is known that is related to A_t. Since our method is also available when group membership is unknown, this connection provides intuition and credibility for the results in the frequent situation wherein group membership is not known. This connection may be developed as follows: Let Y_{tj} denote the relative rank position of respondent j on the ranked response categorical variable t relative to all N claim files containing a categorization for variable t, and let I_A denote the indicator function of the set A. The relative rank of a particular claim's score may be represented as

$$Y_{tj} = \frac{N}{2} \sum_{i=1}^{k_t} B_{ti} I\,[category\ i\ given] + \frac{N}{2}. \tag{A3}$$

Now, if N_q claim files originate from group q, $q = 1{,}2$, then (using the relationship between average relative ranks of two groups in a combined sample versus the Wilcoxon rank sum statistic) the average score for a group 1 claim file on t is $\bar{B}_t^{(1)} = 2W_1^*/N_1 - 2$ where W_1^* is the Wilcoxon rank sum statistic for comparing group 1 and 2 classifications to t, allowing for ties (see Beder and Heim, 1990; Selvin, 1977).

Accordingly, if we had knowledge of group membership (as occurs, for example, in discriminant analysis, supervised neural networks, or logistic regression), then our RIDIT scoring technique, B_{ti}, immediately yields a measure $\bar{B}^{(1)}$ of "variable discriminatory power," namely, the Wilcoxon statistic. This encouraging result motivates us to continue onward to the more difficult situation wherein group membership is unknown. To this end, we show how to use results from principal component analysis and factor analysis to obtain a consistent *estimate* of A_t and, hence, obtain a measure of an individual variable's value for discriminating fraudulent claims.[30]

An important assumption we make is that we are dealing with a first-order probability model; that is, if for a particular collection of claims in either the fraud or nonfraud group, one wishes to know what proportion of files are expected to be simultaneously classified into category i on variable t and into category j on variable s, one needs only know the marginal probabilities of category i on variable t and category j on variable s. This first-order model assumption statistically translates into the property of conditional lack of correlation of variable classifications, given group membership (fraud or nonfraud). Remember that since we are only assuming attributes of the conditional

[30] Note that in PRIDIT analysis, the tasks of determining variable discriminatory ability, scoring of variables, and obtaining overall fraud suspicion scores are not done separately as in most other fraud analyses. Rather, our analysis shows that the principal component analysis of RIDITs technique introduced here provides an inner consistency between these aspects of fraud suspicion analysis, which is also related to measures of external validity.

distributions, this assumption of a first-order model does *not* imply lack of correlation between variable classifications for the total sample of claim files. Indeed, in practice there is *much* dependence among variables (which is why they are useful for fraud detection). However, it is sufficient that this dependence drops significantly *within* the individual subgroups. Note also that this conditional lack of correlation assumption is also common to latent trait models, where it is called "local independence" (see Hambleton and Cook, 1977). The latent trait in our analysis is suspiciousness of the claim.

We turn now to the situation in which the group classifications are unknown. As before we have $p_t = \theta\pi_t^{(1)} + (1-\theta)\pi_t^{(2)}$ for fraud group $q = 1$ and nonfraud group $q = 2$ where $\theta = N_1/N$ is the proportion of claim files belonging to group 1. The following lemma relates the theoretical (and unobserved) quantities $\theta, \pi_t^{(1)}$, and $\pi_t t^{(2)}$ to the previous scoring system and determines a measure of predictor variable discriminatory power (A_t).

Lemma 1: *Assume that a claim file produces a score B_{ti} on predictive fraud indicator variable t. Then the expected score for a claim file in the low suspicion group* 1 *is $(\theta-1)A_t$, and the expected score for a high suspicion group* 2 *claim file is θA_t where A_t is given by* (3).

Proof of Lemma 1: Only prove the result for the high fraud suspicion group 1 because a similar calculation will yield the low or nonfraud suspicion group result. Select a respondent at random from the fraud group 1 and let $I_i = 1$ if the ith categorical classification level was assigned to that claim on that particular variable t. Then the desired expected value is

$$\sum_{i=1}^{kt} \pi_{ti}^{(1)} E[B_{ti}|I_i = 1]$$
$$= \sum_{i=1}^{kt} \pi_{ti}^{(1)} \left[\sum_{j<i} \frac{(N_1-1)\pi_{tj}^{(1)} + N_2\pi_{tj}^{(2)}}{N} - \sum_{j>i} \frac{(N_1-1)\pi_{tj}^{(1)} + N_2\pi_{tj}^{(2)}}{N} \right]. \tag{A4}$$

Using the identities $N_1 = N - N_2 = N\theta$ and $\sum_{i=1}^{k_t-1}\sum_{j<i} x_i y_j = \sum_{i=1}^{k_t-1}\sum_{j>i} x_j y_i$, we find after elementary computation this expected value is

$$-\frac{N_2}{N}\sum_{i=1}^{k_t-1}\sum_{j>i}\left\{\pi_{ti}^{(1)}\pi_{tj}^{(2)} - \pi_{ti}^{(2)}\pi_{tj}^{(1)}\right\} = (\theta-1)A_t. \tag{A5}$$

Now prove the results of the theorem stated in the text. Let B_{qt}^* denote the score of a randomly selected claim file from suspicion level group q on variable t, and let $\mathbf{G} = E(\mathbf{F}'\mathbf{F})$.

Lemma 2: *Let $\mathbf{G} = \mathbf{M} + \mathbf{U}$ where $\mathbf{M} = N_1N_2(A_iA_j)/N$ and $\mathbf{U}$ is diagonal with $U_{tt} = N\sigma_{1t}^2 + N_2\sigma_{2t}^2$, where σ_{qt}^t = variance (B_{qt}^*). (Note that $\mathbf{G}$ has a factor analytic structure with U_{tt} being the uniqueness component of the variance for variable t).*

Proof of Lemma 2: We have $\mathbf{G} = (g_{st})$ with

$$g_{st} = E\left[\sum_{i=1}^{N} f_{is}f_{it}\right] = E\left[\sum_{i=1}^{N_1} f_{is}f_{it} + \sum_{i=N_1+1}^{N} f_{is}f_{it}\right], \tag{A6}$$

where we arranged the first N_1 claim files so that they belong to group 1. Having a first-order model yields $g_{st} = N_1E[f_{is}f_{it}|i\text{th claim file in the fraud group 1}] + N_2E[f_{is}f_{it}|i\text{th claim file in the nonfraud group 2}]$. If $s \neq t$, use conditional independence and Lemma 1 to obtain $\frac{N_1N_2A_tA_s}{N}$. If $s = t$, add and subtract $\frac{N_1N_2A_t^2}{N}$ to obtain

$$g_{tt} = \frac{N_1N_2A_t^2}{N} + \left[N_1E(B_{2t}^{*2}) - \frac{N_1N_2^2A_t^2}{N}\right] + \left[N_2E(B_{2t}^{*2}) - \frac{N_2N_1^2A_t^2}{N}\right]. \tag{A7}$$

It follows that $\mathrm{G} = \mathrm{M} + \mathrm{U}$ as claimed.

Lemma 3: *The characteristic polynomial of* $\mathbf{G} = E(\mathbf{F}'\mathbf{F})$ *is*

$$\begin{aligned} f(\mu) &= \left\{\prod_{t=1}^{m}(U_{tt} - \mu) + \sum \frac{N_1N_2}{N}A_t^2 \prod_{s,t}(U_{ss} - \mu)\right\} \\ &= \prod_{t=1}^{m}(U_{tt} - \mu)\left\{1 + \sum \frac{N_1N_2}{N}A_t^2/(U_{tt} - \mu)\right\} \end{aligned} \tag{A8}$$

The largest eigenvalue μ_1 *of* $\mathbf{G}$ *is positive and larger than max* U_{tt}. *The second largest eigenvalue is between the two largest values of* U_{tt}.

Proof of Lemma 3: The equation for $f(\mu)$ follows from elementary operations on G (see Graybill, 1969).

To obtain the location of the eigenvalues of G, consider first the case in which all the U_{tt} are distinct and arrange them in order $\mathrm{U}_1 < \mathrm{U}_2 < \ldots < \mathrm{U}_m$. From the formula for $f(\mu)$ the algebraic sign of $f(\mu)$ is clearly plus if $\mu \leq U_1$ and $f(U_k)$ has the sign $(-1)^{k+1}$. Since $f(U_t)$ alternates in sign, f has at least one root between U_t and $\mathrm{U}_{t+1}, t = 1, \ldots, m-1$. We have accounted for $(m-1)$ of the m eigenvalues, and the remaining eigenvalues must be larger than U_m. If all the U_{tt} are not distinct, say, k of them are equal to U_{jj}, then $f(\mu)$ has a root of multiplicity k at U_{jj} and again we have accounted for $(m-1)$ roots to the left of max U_{tt}.

We now may prove the following.

Theorem 1. *The sequences of predictor variable weights* $\{\mathbf{W}^{(n)}\}$ *and overall summative claim file suspicion scores* $\{\mathbf{S}^{(n)}\}$ *converge. Moreover, the limiting predictor variable weight* $\hat{W}^{(\infty)}$ *is the first principal component of* $\mathbf{F}'\mathbf{F}$, *which is a consistent estimate of the principal component* $\mathbf{W}^{(\infty)}$ *of* $E(\mathbf{F}'\mathbf{F})$, *the* tth *component of which is explicitly*

$$W^{(\infty)}{}_t = \frac{A_t}{(\mu_1 - U_{tt})\sqrt{\sum_{s=1}^{m} A_s^2/(\mu_1 - U_{ss})^2}}, \tag{A9}$$

where μ_1 is the largest eigenvalue of $E(\mathbf{F}'\mathbf{F})$ and $U_{tt} = N_1\sigma_{1t}^2 + N_2\sigma_{2t}^2$ is the "uniqueness component of variance" in a single-factor analytic model.

Proof of Theorem 1: First show that $\lim_{(n)} \mathbf{W}^{(n)}$ and $\lim_{(n)} \mathbf{S}^{(n)}$ exist. Note that $\mathbf{W}^{(n)} = \mathbf{W}^{(n)}\mathbf{F}'\mathbf{S}^{(n-1)}/\|\mathbf{F}'\mathbf{S}^{(n-1)}\| = (\mathbf{F}'\mathbf{F})^{n}\mathbf{W}^{(0)}/\|(\mathbf{F}'\mathbf{F})^{n}\mathbf{W}^{(0)}\|$. By elementary linear algebra, it follows that $\lim \mathbf{W}^{(n)}$ exists and is the projection of $\mathbf{W}^{(0)}$ onto the eigenspace generated by the largest eigenvalue of $F'F$. To calculate the first principal component of $E(F'F)$, use Lemma 3 and the matrix $E(\mathbf{F}'\mathbf{F}) - \mu\mathbf{I}$. Note that if V_1 is the first principal component, then it must satisfy the equations

$$V_{1i} = \frac{(\mu_1 - U_{22})A_i}{(\mu_1 - U_{ii})A_2} V_{12} \quad \text{for } i \neq 2. \tag{A10}$$

This together with the fact that $f(\mu_i) = 0$ implies $\sum_{1=1}^{m} \frac{N_1 N_2 A_i^2}{(U_{ii}-\mu_1)} + 1 = 0$ shows that with $W_t^{(\infty)}$ given as in the theorem we have $E(\mathbf{F}'\mathbf{F})W^{(\infty)} = \mu_1 W^{(\infty)}$, proving the theorem.

APPENDIX B

Suspicion of Fraud Indicator Variables From the AIB Data

		Weights	
		By Category	Overall
A. Accident Characteristics			
ACC1	No report by police officer at scene	−0.13	0.33
ACC2	No witnesses to accident	0.22	0.15
ACC3	Rear-end collision	0.42	0.40
ACC4	Single vehicle accident	−0.63	−0.43
ACC5	Controlled intersection collision	0.43	0.08
ACC6	Claimant was in parked vehicle	−0.02	0.14
ACC7	Two drivers were related or friends	−0.21	0.06
ACC8	Late-night accident	0.01	0.01
ACC9	No plausible explanation for accident	−0.26	0.02
ACC10	Claimant in an old, low-value vehicle	0.45	0.18
ACC11	Rental vehicle involved in accident	−0.18	−0.01
ACC12	No tow from scene despite severely damaged car	0.11	0.20
ACC13	Site investigation raised questions	0.00	0.00
ACC14	Property damage was inconsistent with accident	0.31	0.24
ACC15	Very minor-impact collision	0.54	0.57
ACC16	Claimant vehicle stopped short	0.52	0.11
ACC17	Claimant vehicle made unexpected maneuver	0.38	0.15
ACC18	Insured/claimant versions differ	0.24	0.14
ACC19	Insured felt set up, denied fault	0.43	0.17

APPENDIX B CONTINUED

		Weights	
		By Category	Overall
B. BI Claimant Characteristics			
CLT1	Retained an attorney very quickly	0.23	0.16
CLT2	Had a history of previous claims	0.33	0.28
CLT3	Gave address as hotel or P.O. Box	0.00	0.00
CLT4	Was an out-of-state resident	−0.22	−0.16
CLT5	Retained a "high-volume" attorney (list obtained from known table)	−0.03	0.24
CLT6	Was difficult to contact/uncooperative	0.69	0.14
CLT7	Was one of three or more claimants in vehicle	0.18	−0.02
CLT8	Was resident of high-claim town (obtained from known list)	0.38	0.35
CLT9	Avoided use of telephone or mail	0.64	0.15
CLT10	Was unemployed	0.57	0.12
CLT11	Appeared to be "claims-wise"	0.12	0.48
C. BI Insured Driver Characteristics			
INS1	Had a history of previous claims	0.74	0.12
INS2	Gave address as hotel or P.O. Box	0.00	0.00
INS3	Readily accepted fault for accident	0.05	0.13
INS4	Was acquainted with other vehicle occupants	0.66	−0.05
INS5	Was not willing to provide a sworn statement	0.00	0.00
INS6	Was difficult to contact/uncooperative	0.37	0.03
INS7	Accident occurred soon after policy effective date	0.31	0.01
INS8	Appeared to be "claims-wise"	0.30	0.43
D. Injury Characteristics			
INJ1	Injury consisted of strain/sprain only	0.77	0.72
INJ2	No objective evidence of injury	0.75	0.61
INJ3	Police report showed no injury or pain	0.51	0.18
INJ4	Claimant refused to appear for an independent medical examination	0.10	0.16
INJ5	No emergency treatment was given for the injury	0.70	0.42
INJ6	Non-emergency treatment was delayed	0.53	0.42
INJ7	First non-emergency treatment was by a chiropractor	0.17	0.35
INJ8	Activity check cast doubt on injury	−0.07	−0.10
INJ9	Injuries were inconsistent with police report	0.36	0.07
INJ10	Independent medical examination suggests injury was unrelated to accident	0.29	0.02

APPENDIX B CONTINUED

		Weights	
		By Category	Overall
INJ11	Unusual injury for this auto accident	0.14	0.10
INJ12	Evidence of an alternative cause of injury	0.02	0.26
E. Treatment Characteristics			
TRT1	Large number of visits to a chiropractor	0.3	0.65
TRT2	Chiropractor provided 3 or more modalities on most visits	0.19	0.37
TRT3	Large number of visits to a physical therapist	0.53	0.13
TRT4	MRI or CT scan but no inpatient hospital charges	0.38	0.05
TRT5	Use of “high-volume” medical provider (from a known list)	0.02	0.36
TRT6	Significant gaps in course of treatment	0.70	0.17
TRT7	Treatment was unusually prolonged (more than 6 months)	0.82	0.18
TRT8	Independent medical examiner questioned extent of treatment	0.37	0.18
TRT9	Medical audit raised questions about charges	−0.13	0.28
F. Lost Wage Characteristics			
LW1	Claimant worked for self or family member	−0.11	−0.05
LW2	Employer wage differs from claimed wage loss	0.48	−0.01
LW3	Claimant recently started employment	0.90	−0.18
LW4	Employer unknown/hard to reach	0.00	0.00
LW5	Lost wages statement looked unofficial	0.77	−0.13
LW6	Long disability for a minor injury	0.25	0.08

REFERENCES

Artís, M., M. Ayuso, and M. Guillén, 1999, Modeling Different Types of Automobile Insurance Fraud Behavior in the Spanish Market, *Insurance*: *Mathematics and Economics*, 24: 67-81.

Ayuso, M., 1999, El Fraude en el Seguro del Automóvil: Cómo Detectarlo, *Cuaderno de la Fundación*, No. 47, Barcelona, Spain.

Beder, J. H., and R. C. Heim, 1990, On the Use of RIDIT Analysis, *Psychometrika*, 55: 603-616.

Belhadji, E. B., G. Dionne, and F. Tarkhani, 2000, A Model for the Detection of Insurance Fraud, *The Geneva Papers on Risk and Insurance*, 25: 517-538.

Benyouseff, A., J. L. Cutler, A. Levine, P. Mansourian, T. Phan-Tan, and R. Baylett, 1974, Health Effects of Rural-Urban Migration, *Social Science and Medicine*, 8: 243-254.

Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland, 1977, *Discrete Multivariate Analysis, Theory and Practice* (Cambridge, Mass.: MIT Press).

Brockett, P., 1981, A Note on Numerical Assignment of Scores to Ranked Categorical Data, *Journal of Mathematical Sociology*, 8: 91-101.

Brockett, P. L., R. A. Derrig, and X. Xia, 1998, Using Kohonen's Self-Organizing Feature Map to Uncover Automobile Bodily Injury Claims Fraud, *Journal of Risk and Insurance*, 65: 245-274.

Brockett, P. L., and L. L. Golden, 1992, A Comment on "Using Rank Values as an Interval Scale" by Dowling and Midgley, *Psychology and Marketing*, 9: 255-261.

Brockett, P. L., and A. Levine, 1977, On a Characterization of RIDITs, *Annals of Statistics*, 5: 1245-1248.

Broffit, J. D., R. H. Randles, and R. V. Hogg, 1976, Distribution-Free Partial Discriminant Analysis, *Journal of the American Statistical Association*, 71: 934-939.

Bross, I., 1958, How to Use RIDIT Analysis, *Biometrics*, 14: 18-38.

Bross, I., 1960, How to Cut the Highway Toll in Half in the Next Ten Years, *Public Health Reports*, 75: 573-581.

Brown, S. M., S. Selvin, and W. Winkelstein Jr., 1975, The Association of Economic Status With the Occurrence of Lung Cancer, *Cancer*, 36: 1903-1911.

Canadian Coalition Against Insurance Fraud, 1997, Red Flags for Detecting Insurance Fraud, 1-33, Toronto.

Caron, L., and G. Dionne, 1999, Insurance Fraud Estimation: More Evidence From Quebec Automobile Insurance Industry, in: G. Dionne and C. Laberge-Nadeau, eds., *Automobile Insurance: Road Safety, New Drivers, Risks, Insurance Fraud and Regulation* (Boston, Mass.: Kluwer).

Clemens, W. V., 1956, An Analytical and Empirical Examination of Some Properties of Ipsative Measures, *Psychometric Monographs*, vol. 14 (Chicago, Ill.: University of Chicago Press).

Conover, J., and R. Iman, 1978, The Rank Transformation as a Method of Discrimination with Some Examples, Research Report SAND-0533, Sandia Laboratories.

Daultrey, S., 1976, *Principal Components Analysis*, Geo Abstracts, University of East Anglia, Norwich, England.

Day, N. E., 1969, Estimating the Components of a Mixture of Two Normal Distribution, *Biometrika*, 56: 463-474.

Derrig, R. A., and K. M. Ostaszewski, 1995, Fuzzy Techniques of Pattern Recognition in Risk and Claim Classification, *Journal of Risk and Insurance*, 62: 447-482.

Dionne, G., and E.-B. Belhadji, 1996, Evaluation de la Fraude à l'Assurance Automobile au Quebec, *Assurances*, 64: 365-394.

Forthofer, R. N., and G. Koch, 1973, An Analysis for Compounded Functions of Categorical Data, *Biometrics*, 29: 143-157.

Ganesalingam, S., and G. J. McLachlan, 1978, The Efficiency of a Linear Discriminant Function Based on Unclassified Initial Samples, *Biometrika*, 65: 658-622.

Gensch, D. H., and W. W. Recker, 1979, The Multinomial, Multiattribute LOGIT Choice Model, *Journal of Marketing Research*, 16: 124-133.

Golden, L. L., and P. L. Brockett, 1987, The Effects of Alternative Scoring Techniques on the Analysis of Rank Ordered Categorical Data, *Journal of Mathematical Sociology*, 12: 383-414.

Graybill, F., 1969, *Introduction to Matrices With Applications in Statistics* (Duxbury Press).

Hambleton, R. K., and L. L. Cook, 1977, Latent Trait Models and Their Use in the Analysis of Educational Test Data, *Journal of Educational Measurement*, 14: 75-96.

Kantor, S., W. Winkelstein Jr., D. L. Sackett, et al., 1966, A Method for Classifying Blood Pressure: An Empirical Approach to Reduction of Misclassification Due to Response Instability, *American Journal of Epidemiology*, 84: 510-523.

Levine, D. W., 1991, Classification and Estimation of Misclassification for Ordered Categorized Data, Ph.D. dissertation, University of California.

Medgyessey, P., 1978, *Decomposition of Superpositions of Density Functions and Discrete Distributions* (New York: Wiley).

Randles, R. H., J. D. Broffit, and R. V. Hogg, 1978, Discriminant Analysis Based on Ranks, *Journal of the American Statistical Association*, 73: 379-384.

Rubin, D., 1976, Inference and Missing Data, *Biometrika*, 63: 581-590.

Rudas, T., 1998, Odds Ratios in the Analysis of Contingency Tables, Sage University Paper 119, London.

Sellitz, C., et al., 1964, *Research Methods in Social Relations* (New York: Holt, Rinehart, and Winston).

Selvin, S., 1977, A Further Note on the Interpretation of RIDIT Analysis, *American Journal of Epidemiology*, 105: 16-20.

Stuart, A., and J. K. Ord, 1991, *Kendall's Advanced Theory of Statistics*, Vol. 2, (New York: Oxford University Press).

Upton, G. J., 1978, *The Analysis of Cross-Tabulated Data* (New York: Wiley).

Weisberg, H. I., and R. A. Derrig, 1991, Fraud and Automobile Insurance: A Report on Bodily Injury Liability Claims in Massachusetts, *Journal of Insurance Regulation*, 9: 497-541.

Weisberg, H. I., and R. A. Derrig, 1992, Massachusetts Automobile Bodily Injury Tort Reform, *Journal of Insurance Regulation*, 10: 384-440.

Weisberg, H. I., and R. A. Derrig, 1996, Coping with the Influx of Suspicious Strain and Sprain Claims, A Report on the AIB Study of 1993 Personal Injury Protection and Bodily Injury Liability Claims, Massachusetts DOI Fraudulent Claims Payment Filing.

Weisberg, H. I., and R. A. Derrig, 1998, Quantitative Methods for Detecting Fraudulent Automobile Bodily Injury Claims, *Risques*, 35: 75-101.

Weisberg, H. I., R. A. Derrig, and X. Chen, 1994, Behavioral Factors and Lotteries Under No-Fault With a Monetary Threshold: A Study of Massachusetts Automobile Claims, *Journal of Risk and Insurance*, 61: 245-275.

Williams, O. D., and J. E. Grizzle, 1972, Analysis of Contingency Tables Having Ordered Response Categories, *Journal of the American Statistical Association*, 67: 55-63.

Wynder, E. L., I. D. J. Bross, and T. Hiroyama, 1960, A Study of the Epidemiology of Cancer of the Breast, *Cancer*, 13: 559-601.

Zimmerman, H. J., 1999, *Practical Applications of Fuzzy Technologies* (Boston, Mass.: Kluwer).