FRAUDULENT RECEIPT OF UNEMPLOYMENT INSURANCE BENEFITS: CHARACTERISTICS OF THOSE WHO COMMITTED FRAUD AND A PREDICTION PROFILE

PREPARED FOR

ARIZONA UNEMPLOYMENT INSURANCE TASK FORCE
UNEMPLOYMENT INSURANCE BUREAU
ARIZONA DEPARTMENT OF ECONOMIC SECURITY

BY

Robert D. St. Louis
Manager, Research and Reports Section
Arizona Unemployment Insurance Bureau

and

Paul L. Burgess and Jerry L. Kingston
Associate Professors of Economics
Arizona State University

JUNE, 1978

PREFACE

This report is one of a series of special reports prepared for the Unemployment Insurance Task Force. The Task Force was created July 12, 1976 through Interdivisional Directive No. 76-5 by Dr. Arlyn Larson, then Deputy Director of the Arizona Department of Economic Security. Because of the tremendous program and financial pressures that the DES Unemployment Insurance System had experienced during the prior two-year period, an in-depth analysis of many aspects of the UI program was to be conducted by the Task Force. Following discussions with employer and employee groups, members of the Arizona Legislature, the public at large and the regional staff of the Employment and Training Administration, it was determined that emphasis was to be placed on the formulation of recommendations for changes in both employment security law and policy, and on the development of a number of research studies designed to provide factual background or appropriate analysis of current policy problems.

The Unemployment Insurance Task Force is composed of the following members: Mr. Joseph Anderson (Task Force Project Officer and Cost Model Coordinator); Mr. Henry Haas (Chief, Unemployment Insurance Bureau); Mr. Harvey Finger (Chief, Appeals Bureau); Mr. Charles Vance (Contributions Section Manager); Mr. Tom Vaughn (Benefits Section Manager); and Dr. Robert D. St. Louis (Manager, Research and Reports Section). Dr. St. Louis and Drs. Paul L. Burgess and Jerry L. Kingston (Associate Professors of Economics, Arizona State University) served as research directors for the Task Force.

Special thanks are due Anna Christy (EDP Project Manager) for her outstanding assistance in determining the appropriate data elements for this study and for producing the data base for the study well before it reasonably could have been expected. Joseph Sloane of the Unemployment Insurance Bureau did an outstanding job of producing the many computer runs needed for the report. Appreciation also is expressed to Polly Green and members of the Task Force for their comments on the draft report; especially valuable were a number of points raised by Tom Vaughn.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

The results of an analysis of persons known to have received fraudulent overpayments during benefit years that were completed during the period FY 1973-1977 are summarized in this report.[1] The overall purpose of this pilot study was to determine whether it would be feasible to develop a profile for identifying claimants likely to fraudulently collect benefits. It is important to emphasize at the outset that the profile developed has not been operationally implemented. First, it was necessary to determine whether such a profile could be developed. As explained below, this initial attempt appears to have been successful. The analysis clearly indicates that the propensity to commit (known) fraud was not randomly distributed among the claimant population. Thus, it appears possible to systematically identify claimants who, for whatever reason, appear to be much more inclined to commit fraud then the typical claimant. It is important to emphasize strongly that the analysis is based solely on those who committed fraud and were detected. The first section provides a comparison of the characteristics of those who did/did not receive at least one fraudulent overpayment during the study period. The next section provides a brief explanation of discriminant analysis--the statistical technique utilized to develop a fraud profile screening device. The final sections provide some of the results obtained from utilizing this profile.

CHARACTERISTICS COMPARISONS: FRAUDS VS. NONFRAUDS

The characteristics of those who did/did not receive any fraudulent overpayment during the study period are compared in Tables 1-12. The findings for each characteristic are summarized below. Before turning to these individual characteristics, however, it should be noted that an estimated total of 5,304 claimants received fraudulent overpayments during this period. These claimants represented 1.6 percent of all claimants who were determined to be monetarily eligible for benefits during the five-year period (for what might be referred to as an overall fraud rate of 1.6%).

Sex

The fraud rate was much higher among men than among women (2.0% vs. 0.9%).[2] Although men represented only 67 percent of all claimants during this period, they accounted for 82 percent of all (detected) fraudulent overpayments.

Age

Interestingly, the fraud rate declined considerably for older vs. younger beneficiaries. For those under 25 years of age, 2.1 percent had fraudulent overpayments, compared with 1.6 percent for those 25-54 years of age and only 1.0 percent for beneficiaries 55 years or older.[3] (Evidently, the aging process increases either honesty or knowledge of how to avoid detection.)

Ethnic Category

The fraud rates found for each ethnic group that could be analyzed were nearly identical. The slightly higher rate of 1.7 percent for whites does not differ significantly (absolutely or in statistical terms) from the rates of 1.6 percent found for both whites with Spanish surnames and those in the "other" category.[4]

Occupational Category

The incidence of fraud differed considerably among the six occupational categories that could be analyzed (see Table 4).[5] The lowest rates were recorded for workers in processing/machine trades/bench work occupations (1.1%) and for professional/technical/managerial workers (1.2%). In contrast, the highest fraud rates were those found for structural workers (2.4%) and those in the miscellaneous/farming category (2.0%); in fact, the structural workers category accounted for nearly two-fifths of all detected frauds.

High Quarter and Base Period Earnings

The results for claimants, classified by earnings, also indicate that fraud incidence systematically varied among claimants who had different levels of previous earnings (see Tables 5 and 6). For high quarter earnings, those with earnings of less than $1000 had the lowest fraud rate for any earnings category of 1.1 percent, whereas those with earnings above $2,500 had the highest fraud rate of 1.9 percent. Interestingly, however, the results for base period earnings show that the highest fraud rate for these earnings categories was 2.6 percent and that this rate was for those with annual earnings of only $3,001-$4,000. For all other base period earnings categories, the incidence of fraud varied in a narrow range of 1.3-1.6 percent. For the interested reader, fraud rates also are shown for the base period/high quarter earnings ratio in Table 7.

Weekly Benefit Amount

Fraud rates for claimants, classified by the weekly benefit amount for which they qualified, varied from a low of 1.1 percent for those who qualified for a payment of less than $35 to a high of 1.9 percent for those eligible for weekly benefits of at least $66 (see Table 8). For the remaining WBA categories of $35-50 and $51-65, the fraud rates recorded were 1.7 percent and 1.5 percent, respectively. The results presented in Table 9 show that the incidence of fraud was considerably higher for those with weekly benefits that amounted to 61 percent or more of their average weekly earnings in the entire base period than for those who had a smaller percentage of their average weekly earnings in the base period replaced by UI benefits.

Potential Duration, Regular Benefits

The results shown in Table 10 indicate that the incidence of fraud tended to decline somewhat as the potential duration of benefits increased. The fraud rates recorded were 1.9 percent for those with less than 20 weeks of potential duration, 1.8 percent for those in the 20-25 weeks category and 1.5 percent for those who qualified for the maximum of 26 weeks.

Benefit Year Data

The previous variables all are known at the time claimants file for benefits. Thus, the previous variables could be used to develop a profile to screen claimants for fraud likelihood at the time they initially file to establish benefit eligibility. The two variables considered in this section--spells of unemployment in the benefit year and weeks with (reported) deductible earnings--obviously are not known until claimants have completed benefit years. Thus, these variables might be potentially useful in determining the likelihood that fraud has occurred once benefit years have been completed.

The results for spells of unemployment are presented in Table 11, and the pattern revealed is extremely interesting. The incidence of fraud increased sharply for those in each higher spells of unemployment category. For those with only one spell of unemployment, the fraud rate was 0.9 percent. This rate rose to 1.6 percent for those with two spells, to 2.7 percent for those with three spells and to 3.0 percent for those with four or more spells. The basis of this pattern is not entirely clear. Perhaps the temptation of not reporting earnings when reemployed rises for the frequently unemployed or perhaps they feel they can "beat the system" as they gain experience with it. In any case, the pattern is a rather striking one, and even might be useful in detecting fraud before benefit years are completed.

An interesting pattern also emerged for the results for weeks reported with deductible earnings (see Table 12). The fraud rates are nearly the same for persons with no deductible earnings reported (1.5%) and for those with three or more weeks of earnings reported (1.4%). For those with only one or two weeks of reported earnings, however, the incidence of fraud was 2.0 percent.

DISCRIMINANT ANALYSIS AS A TECHNIQUE FOR DEVELOPING A FRAUD PROFILE

Some brief background on discriminant analysis is provided in this section[6] to set the stage for the empirical estimation of the fraud profiles in the next section. The basic purpose of discriminant analysis in the present context is to develop a profile to predict the propensity for fraud of any particular claimant. In the subsequent empirical analysis, two different profiles are estimated--one based just on data available at the time initial benefit eligibility is determined and one that also depends on information generated during the benefit year. Thus, to the extent the profiles prove useful, one could be used to indicate persons that should be carefully monitored during the benefit year, whereas the other could be used to indicate the most effective way to allocate resources for auditing those who already have completed benefit years.

It is possible through discriminant analysis to develop a profile that places persons (with some probability) into any one of several mutually exclusive categories. For the fraud profile of interest here, however, only two categories are relevant--frauds and nonfrauds. It should again be noted that only detected frauds are included in the fraud category. The more undetected frauds who are left for purposes of analysis with the nonfraud group, the more difficult it would be to develop an accurate profile. Although the magnitude of this problem is unknown, it surely reduced the effectiveness that could have been achieved if all undetected frauds actually had been detected. To the extent one or more reasonably accurate profiles could be developed, however, the problem of undetected frauds should diminish through time as more are detected; this, in turn, should make it possible to develop even more accurate profiles and to further enhance detection efforts.

The essential feature of discriminant analysis in the present context is to classify UI claimants, on the basis of their individual and labor market characteristics, into one of two mutually exclusive and exhaustive groups--detected frauds vs. all other claimants. The choice of a decision criterion for this classification of individual claimants into these two categories is the basic problem to be solved.[7] In Figures 1 and 2, discriminant scores are some function of the various personal and labor market characteristics of the claimants analyzed. Ideally, a completely accurate classification criterion would allow identification of each person in each of the two groups (at the time of initial claims filing), and the result would be no overlap between the two distributions, as illustrated in Figure 1. In fact, however, a decision criterion that produces classification with 100 percent accuracy is generally not possible. Given equal misclassification costs and equal prior probability of misclassification (not in fact the case here), the objective in selecting the decision criterion essentially would be to minimize the extent of overlap of the two distributions and thus to minimize the number of misclassifications. An hypothetical example of an imperfect decision criterion is illustrated by point A in Figure 2; in this example, all nonfrauds to the right of point A in the nonfraud distribution are misclassified as frauds, and all frauds to the left of point A in the fraud distribution are misclassified as nonfrauds.

Since some misclassification (as in Figure 2) generally would be unavoidable, the selection of the decision criterion requires estimation of the "costs" of misclassification for each category. In this context, if

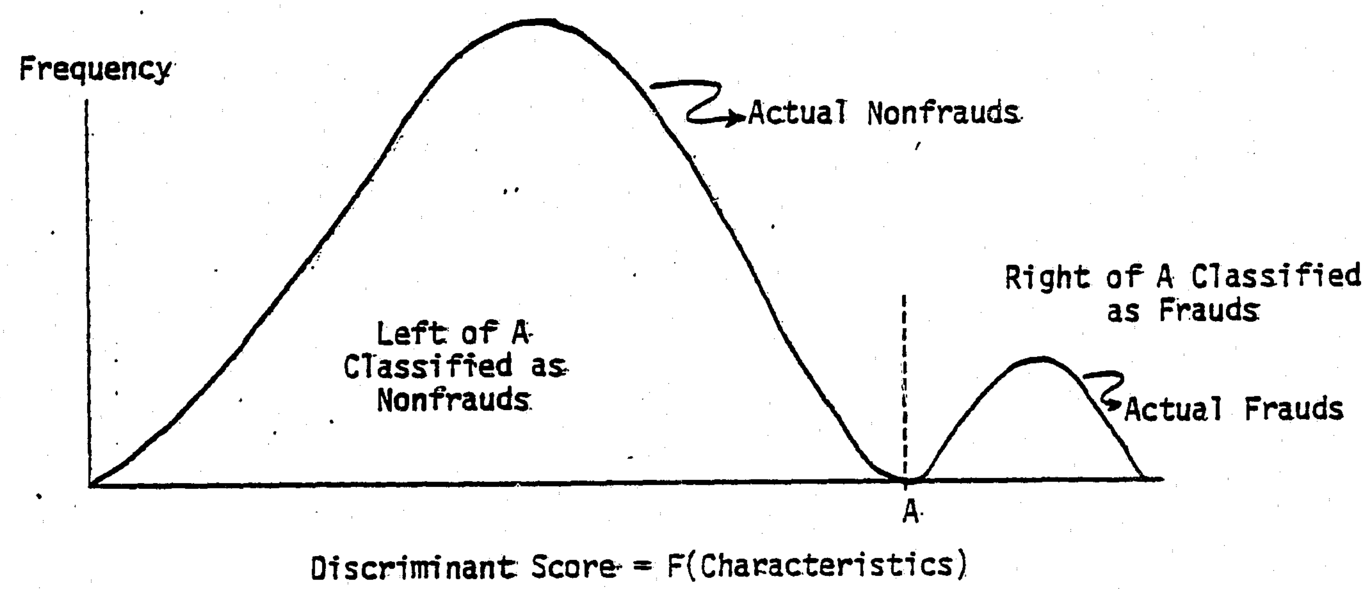

FIGURE 1

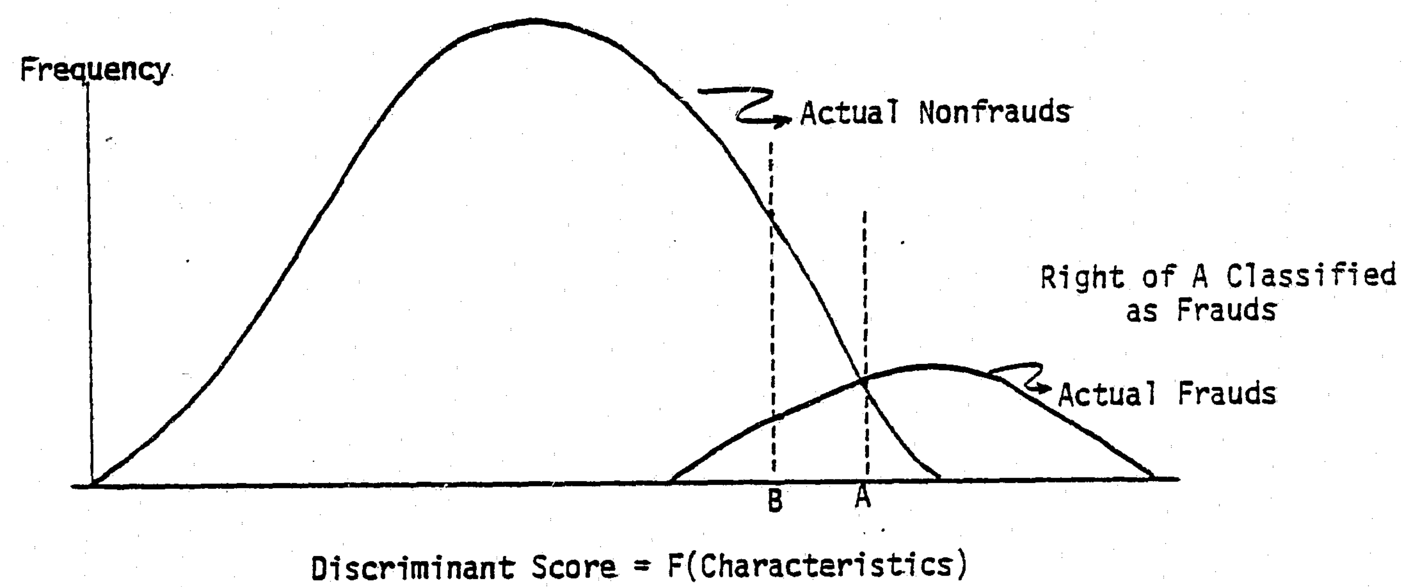

FIGURE 2

the costs of misclassifying frauds and nonfrauds were equal, an individual claimant would be assigned to that category in which he had the highest posterior probability (e.g., above 50% in the present two-category case). In some instances, however, misclassification costs are not equal, and an adjustment in the decision criterion must be made to account for the unequal costs. In the example shown in Figure 2, if it were assumed that the cost of misclassifying a fraud as a nonfraud were about four times the opposite error, the appropriate adjustment of the decision criterion would require a movement from point A to about point B, so that fewer frauds would be misclassified as nonfrauds, even though more nonfrauds would be misclassified as frauds. This decision criterion equates the expected value of the costs of misclassification of a person into either category. The possibility of directly incorporating assumptions as to misclassification costs into the decision criterion represents a great strength of discriminant analysis, particularly compared with regression analysis in which we are aware of no technique for introducing unequal misclassification costs.[8]

Although it is possible to estimate the relative importance of individual influences on the classification power of discriminant functions, the emphasis in this paper is on the overall capability of the discriminant functions to correctly classify individual claimants as frauds vs. nonfrauds.[9] One indicator of this capability is the classification (or "confusion") matrix. An hypothetical classification matrix for the fraud vs. nonfraud categories is provided below to illustrate the way in which the overall classification power of a discriminant function is indicated by a classification matrix.

Actual Status \ Predicted Status	Fraud	Nonfraud	% Correct By Category
Fraud	40	60	40
Nonfraud	10	90	90

In this example, a total of 200 claimants (100 of whom actually belong in each category) have been classified. Of the 100 claimants who actually committed fraud, this hypothetical discriminant function correctly identified 40 percent as frauds, and the remaining 60 percent were classified incorrectly. In contrast, of the 100 nonfrauds, 90 percent were correctly classified and the remaining 10 percent were classified incorrectly as frauds. Overall, 65 percent (130/200) of the total sample of 200 claimants were classified correctly by the hypothetical discriminant function estimated. As is apparent from this example, the classification matrix provides a convenient and direct indication of the discriminatory power of any particular discriminant function for any number of categories, and these matrices are presented in the next section to summarize the results of the analysis.[10]

FRAUD PROFILE AT THE TIME OF INITIAL FILING: EMPIRICAL RESULTS

One version of the fraud profile developed is based only on data available at the time claimants initially file for benefits. The purpose was to determine whether it would be feasible to identify at the start of their benefit years those likely to fraudulently collect benefits during the year. The following were included in the discriminant function estimated: 1) sex; 2) age; 3) ethnic category; 4) occupation; 5) high quarter earnings; 6) base period earnings; 7) the high quarter/base period earnings ratio; 8) the weekly benefit amount; and 9) the weekly benefit amount as a percent of base period weekly earnings.

A difficulty encountered in estimating a discriminant function for populations of vastly different sizes, as is the case here, is that one frequency distribution can be completely contained within the other, as shown below in Figure 3. When this occurs the number of classification errors is minimized by putting everyone into the larger population. Over 98 percent of the claimants would be correctly classified if everyone were categorized as a nonfraud by using A as the decision rule (because nonfrauds accounted for 98.4 percent of all claimants). Given the situation confronted in this study (as illustrated in Figure 3), fewer than 98.4 percent of the claimants would be correctly classified if anyone were classified as a fraud. This is the case because the fraud distribution is completely contained within the nonfraud distribution. Thus, choosing any decision rule to the left of A (which classifies all claimants as nonfrauds) would increase the number of classification errors, as can be illustrated with the choice of C as the decision rule. Given C as the rule, the area CEG would represent detected frauds who now would be accurately classified as frauds; more than offsetting this gain, however, would be the

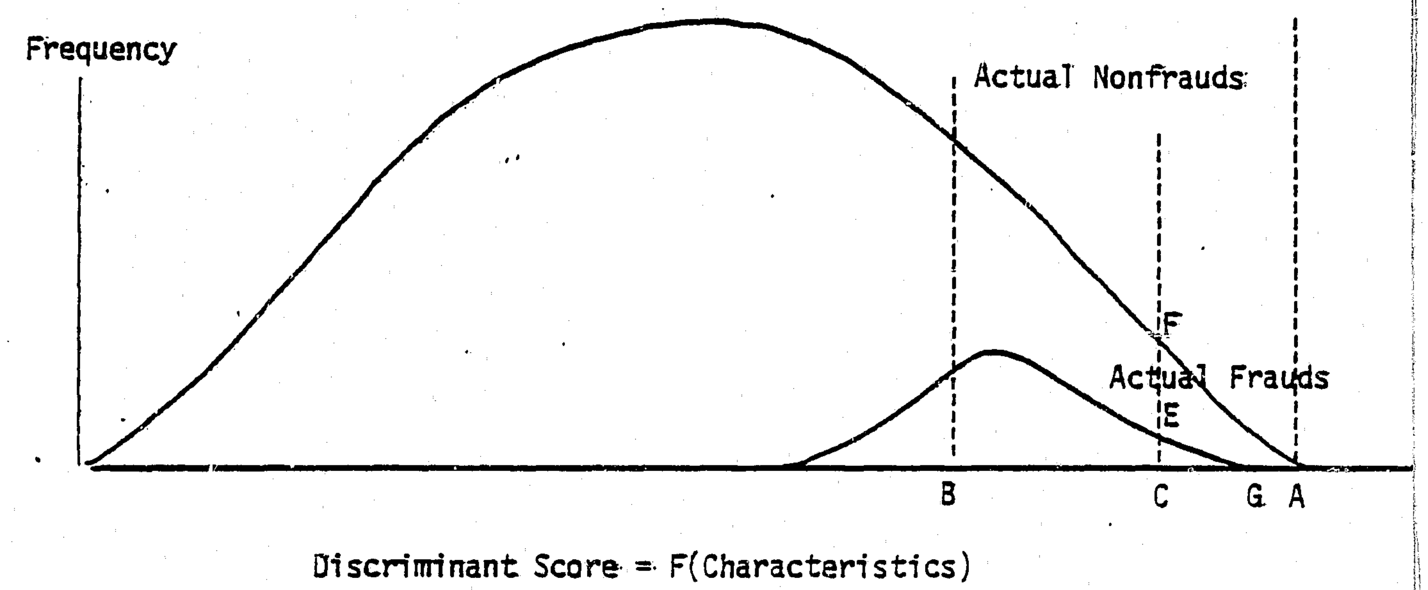

FIGURE 3

nonfrauds incorrectly classified as frauds (represented by the area CFA). In short, the number of errors is minimized in a situation such as this by classifying everyone into the larger group (in this case nonfrauds).

If the costs of misclassification were not equal, however, the optimal or cost minimizing decision may not minimize the number of errors. In particular, if the cost of misclassifying a fraud as a nonfraud is three times as great as the cost of misclassifying a nonfraud as a fraud, the optimal decision point in Figure 3 becomes approximately B rather than A. It is possible to specify the optimal decision point for all combinations or ratios of misclassification costs, but the difficulty of doing so led to a somewhat different approach for this study. The goal was to identify smaller groups that would have a much larger proportion of frauds than the 1.6 percent found for all claimants. If discriminant analysis could be used to accomplish this purpose, then it would be possible to monitor groups with higher fraud propensities than the population as a whole. As noted in the above section on the characteristics of frauds vs. nonfrauds, several of the individual characteristics of the two groups did differ. For example, male claimants are more fraud prone than female claimants. Although interesting, such information is not particularly useful from an operational standpoint because to investigate all males for fraud would be impractical. Thus, the advantage of discriminant analysis is that multivariate control for characteristics hopefully can identify smaller, fraud-prone groups that could be investigated. The way these smaller groups were identified in this analysis was on the basis of their discriminant scores. The higher the discriminant score for a particular claimant (on a scale of -3 to +3), the stronger the indication that claimant belonged in the fraud group. The following table summarizes the results found for all claimants with discriminant scores of 1.0 or more:

Discriminant Score	Number of Claimants With Given Score	Actual Frauds As Percent of Identified Group	Identified Group As Percent of All Claimants
1.0 or more	49,629	3.6%	15.6%
1.5 or more	21,645	4.6%	6.8%
2.0 or more	8,133	5.9%	2.6%
2.5 or more	2,020	8.3%	0.6%

As is evident from the above data, it was not possible to identify any groups that were predominated by frauds. It was possible, however, to identify groups with fraud propensities far greater than the 1.6 percent rate for claimants as a whole. Even the rather large group with discriminant scores of only 1.0 or more had a fraud rate of 3.6 percent, or more than double the fraud rate for claimants as a whole; this group amounted to about one-sixth of all claimants who were monetarily eligible for benefits during this five-year period. As groups with successively higher discriminant scores were identified, group fraud rates increased to 4.6 percent, 5.9 percent and 8.3 percent, respectively. Depending on available resources, any of these groups could be selected for a careful audit during their benefit years. Moreover, groups with high detected fraud rates almost certainly contain undetected frauds. Thus, it would appear that focusing investigative resources on these groups very likely would result in identifying higher percentages of fraudulent claims than those shown in the summary table.

FRAUD PROFILE AFTER THE COMPLETION OF BENEFIT YEARS: EMPIRICAL RESULTS

This section presents the results for the fraud profile developed on the basis of the data utilized in the prior section plus data available once claimants complete their benefit years. Specifically, the nine items listed in the previous section plus number of unemployment spells and weeks with deductible earnings reported by claimants were the factors utilized in estimating the discriminant function for this section. The results are summarized below:

Discriminant Score	Number of Claimants With Given Score	Actual Frauds As Percent of Identified Group	Identified Group As Percent of All Claimants
0.5 or more	96,620	3.4%	30.4%
1.0 or more	53,774	3.7%	16.9%
1.5 or more	22,837	5.8%	7.2%
2.0 or more	7,459	8.9%	2.4%
2.5 or more	1,663	11.5%	0.5%

These results indicate that the addition of spells of unemployment and weeks with deductible earnings substantially increases the ability to discriminate between frauds and nonfrauds. Especially impressive is the increased size of the fraud-prone groups that were identified. In this case, for example, it was possible to identify a large group with a fraud rate of 5.8 percent (3.6 times the rate for all claimants); this group accounted for 7.2 percent of all claimants who were determined to be monetarily eligible for benefits during this period. If fewer investigative resources were available, then those with discriminant scores of 2.0 or more (2.4% of all claimants) could be investigated; this group had a fraud rate of 8.9 percent. Even more selective investigation would focus resources on only about one half of one percent of all claimants. This group had a fraud rate of 11.5 percent, more than 8 times the incidence of fraud found for all claimants. It should be emphasized that the results in this section were obtained without any of the information generated by the crossmatch program (e.g., quarterly wages reported by covered employers for claimants during their already completed benefit years); these quarterly wage data were not available for this study.

Utilization of this information almost certainly would increase substantially the predictive power of the profiles developed for this pilot study. One of two approaches could be taken in utilizing the information from the crossmatch program. That information on quarterly earnings, combined with information on weeks when benefits were received, could be used as an input variable for a discriminant function, or discriminant functions could be estimated for persons classified by earnings received during a quarter of benefit payments. In any case, experimentation along these lines is needed before operationally implementing discriminant analysis for fraud detection.

FRAUD PROFILE: SUMMARY OF EMPIRICAL RESULTS AND CONCLUSIONS

The work conducted for this study indicates that discriminant analysis represents an extremely promising technique for identifying claimants who fraudulently collect benefits. It was possible to develop screening profiles with only data available at the time of initial claims filing or with that data plus information generated during the benefit year. The latter approach proved to be the more effective one, both in terms of identifying larger groups and groups with higher fraud rates. The fraud rate for all claimants during this five-year period was 1.6 percent. Yet through discriminant analysis it was possible to identify a fairly large group of claimants (2.4% of the total) with a fraud rate of 8.9 percent and a smaller group (0.5% of the total) with a fraud rate of 11.5 percent. The policy significance of this ability to identify fraud-prone groups of claimants is obvious--expensive and limited resources can be effectively focused on groups with much higher than average fraud rates. To efficiently allocate fraud detection resources (strictly from a benefit/cost viewpoint, ignoring any other desirable effects of increased fraud detection), additional dollars would be spent for fraud detection up to the point where detection costs just equalled the value of benefits recovered at the margin. Moreover, it seems likely that intensive investigation of these fraud-prone groups would uncover cases of fraud that previously have gone undetected. This effect could cumulate as further detection likely would enhance the ability to identify accurately fraud-prone claimants.

The further research needed to refine the approach developed here consists of four main thrusts. First, the possibility of further improving the profiles developed should be explored. For example, a number of additional variables--such as claimant's residence (urban vs. rural) or whether claims were reopened after overpaid weeks--could be added to determine their effects on the predictive power of the profile. Also, it seems very likely that some improvement could be made by separately estimating profiles for subgroups of workers (such as men and women or young and old workers). The second extension would be to apply the profiles estimated to a new group of claimants (for which data soon will be available), and assess the accuracy of the results obtained. The third extension would be to utilize the information generated by the crossmatch program, as discussed in the prior section, to increase the predictive power of the fraud profiles developed; this extension very likely would result in a dramatic increase in the predictive power of the fraud profiles reported here. Finally, the results found here are so encouraging that the approach should be experimentally developed with data for other states. This extension to other states logically would involve two different steps. For states in which employers must report wages for each employee, as is the case in Arizona, the model developed with the information utilized in this report plus that generated by the crossmatch system could serve as the prototype for other states. For wage-request states, the prototype would be the model developed in this paper. The extension to these latter states appears particularly important since they presently have no crossmatching system for detecting unreported earnings.

FOOTNOTES

[1]Specifically, the study is based on claimants who established benefit years as regular, Arizona UI claimants. Transitional claims from prior benefit years were excluded but interstate liable claims were included. The claimants included represent all who were determined to be monetarily eligible for benefits and had benefit years established during fiscal years 1972-1976; these benefit years were completed during fiscal years 1973-1977. All numbers in the report are estimated totals for all claimants who met the above criteria, based on a randomly selected sample.

[2]William Papier reports that fraud rates also were substantially higher for men than for women in Ohio. See William Papier, "Unemployment Benefit Fraud in Ohio," unpublished paper, June 1977.

[3]Papier, ibid., did not find such a pattern for age (in fact, the oldest claimants had the highest fraud rates of any age group in Ohio).

[4]This pattern differs from that found by Papier (op. cit., p. 7) who reports that the fraud rate was substantially higher among nonwhites than among whites.

[5]One-digit DOT categories were utilized. Occupations were combined in instances where too few cases existed for reliable statistical analysis.

[6]This section draws heavily on Paul L. Burgess and Jerry L. Kingston, "Applications of Multiple Linear Discriminant Analysis to the Labor Market Experience of UI Claimants," paper prepared for the Unemployment Insurance Service, Manpower Administration, U.S. Department of Labor, June, 1974, 24 pp.

[7]Essentially, the assumptions of multiple linear discriminant analysis are as follows: (1) the number of variables used for classification purposes must be greater than or equal to the number of populations minus one; (2) the number of sample observations must exceed the number of variables used for classification purposes; (3) the costs of misclassifying an individual into any one of the other populations are known or can be estimated; (4) the probability that an individual selected at random belongs to a specific population is known for each population; (5) the populations are multivariate normal; and (6) the populations have identical variance-covariance matrices. For a more complete yet highly intuitive discussion of discriminant techniques, see Wendy Lee Graham, "The Labor Force Decisions of Married Female Teachers: A Discriminant Analysis Approach," Review of Economics and Statistics (August, 1973). For a more technical and detailed treatment see William F. Massey, "Discriminant Analysis of Audience Characteristics," Journal of Advertising Research (March, 1965); Ethel S. Gilbert, "On Discrimination Using Qualitative Variables," American Statistical Association Journal (December, 1968); and William G. Cochran, "Some Classification Problems with Multivariate Data," Biometrics (March, 1961).

FOOTNOTES (continued)

[8]The example in the text is for a two-category situation, but the technique is applicable where the number of categories is larger. In fact, another advantage of discriminant analysis over regression analysis is that a large number of mutually exclusive categories (which represent either equal or unequal population sizes) can be analyzed easily with the former but not with the latter.

[9]Coefficients are estimated for each variable used to classify claimants into the relevant categories, and these coefficients can be tested for statistical significance. As noted in the text, however, the emphasis is on the overall power of the discriminant functions estimated so that coefficients for individual variables are not presented in this paper. Moreover, because the fraud profiles developed will be further refined and may be implemented on an operational basis, the relative importance of different characteristics in classifying persons as frauds should not be divulged.

[10]It should be emphasized that the "proper" way to "test" the power of the discriminant functions estimated would be to test the functions with data not utilized to develop them. In this paper, the classification power of the estimated discriminant functions is indicated by presenting the classification matrices for the same data utilized to estimate the functions; this tends to bias the results toward more accuracy than likely would be achieved if different data were utilized to "test" the accuracy of the estimated functions.

TABLE 1

CROSS TABULATION OF SEX BY NUMBER
OF FRAUDULENT OVERPAYMENTS DETECTED*

BENEFIT YEARS ESTABLISHED FY 1972-1976

Number Fraud Overpayments	Sex: Male	Sex: Female	Row Total[a] (Percent)
None	214,099 66.9	105,772 33.1	319,871 98.4
1 or more	4,360 82.2	944 17.8	5,304 1.6
Column Total	218,459 67.2	106,716 32.8	325,175 100.0
(Fraud Rate)[b]	(2.0)	(0.9)	

Chi Square[c] = 551**

*Based on claimants who established benefit years as regular, Arizona UI claimants. Transitional claims from prior benefit years excluded but interstate liable claims included. The claimants included represent all who were determined to be monetarily eligible for benefits and had benefit years established during fiscal years 1972-1976; these benefit years were completed during fiscal years 1973-1977. All numbers are estimated totals for all claimants who met the above criteria, based on a randomly selected sample.

[a]Each cell of the cross tabulation contains two numbers. The upper number represents the estimated cases that fell in the cell. The lower number is the percent of all cases in the entire row accounted for by the estimated cases in the cell. The statistical reliability is extremely high for percentages based on 2,000 or more claimants. For percentages based on 1,000-1,999 claimants, the estimates should be viewed as very good but not exact ones. For percentages based on 500-999 (less than 500) claimants, the estimates should be viewed only as good (rough) approximations of the true totals.

[b]The fraud rate is defined as the percentage of claimants that had one or more detected fraudulent overpayments.

[c]The Chi Square statistic indicates whether the two classification variables in the cross tabulation (number of overpayments and the personal characteristic included in the cross tabulation) are related or not. The larger the value of the statistic (other things equal), the more confident one can be that the two classification variables are related. In particular, a double asterisk (**) is used to denote those statistics that strongly indicate the two classification variables are related; in these instances, it can be asserted that the two variables are related, at a confidence level of .99.

TABLE 2

CROSS TABULATION OF AGE BY NUMBER
OF FRAUDULENT OVERPAYMENTS DETECTED*

Number Fraud Overpayments	Age: Less Than 25 yrs	Age: 25-54 Yrs	Age: 55 Yrs & Over	Row Total[a] (Percent)
None	58,763 18.4	217,992 68.1	43,116 13.5	319,871 98.4
1 or more	1,236 23.3	3,617 68.2	451 8.5	5,304 1.6
Column Total	59,999 18.5	221,609 68.2	43,567 13.4	325,175 100.0
(Fraud Rate)[b]	(2.1)	(1.6)	(1.0)	

Chi Square[c] = 165**

*Based on claimants who established benefit years as regular, Arizona UI claimants. Transitional claims from prior benefit years excluded but interstate liable claims included. The claimants included represent all who were determined to be monetarily eligible for benefits and had benefit years established during fiscal years 1972-1976; these benefit years were completed during fiscal years 1973-1977. All numbers are estimated totals for all claimants who met the above criteria, based on a randomly selected sample.

[a]Each cell of the cross tabulation contains two numbers. The upper number represents the estimated cases that fell in the cell. The lower number is the percent of all cases in the entire row accounted for by the estimated cases in the cell. The statistical reliability is extremely high for percentages based on 2,000 or more claimants. For percentages based on 1,000-1,999 claimants, the estimates should be viewed as very good but not exact ones. For percentages based on 500-999 (less than 500) claimants, the estimates should be viewed only as good (rough) approximations of the true totals.

[b]The fraud rate is defined as the percentage of claimants that had one or more detected fraudulent overpayments.

[c]The Chi Square statistic indicates whether the two classification variables in the cross tabulation (number of overpayments and the personal characteristic included in the cross tabulation) are related or not. The larger the value of the statistic (other things equal), the more confident one can be that the two classification variables are related. In particular, a double asterisk (**) is used to denote those statistics that strongly indicate the two classification variables are related; in these instances, it can be asserted that the two variables are related, at a confidence level of .99.

TABLE 3

CROSS TABULATION OF ETHNIC CATEGORY
BY NUMBER OF FRAUDULENT OVERPAYMENTS DETECTED*

Number Fraud Overpayments	Ethnic Category: White	Ethnic Category: White Spanish Surname	Ethnic Category: Other	Row Total[a] (Percent)
None	225,899 70.6	46,480 14.5	47,492 14.8	319,871 98.4
1 or more	3,792 71.5	748 14.1	764 14.4	5,304 1.6
Column Total	229,691 70.6	47,228 14.5	48,256 14.8	325,175 100.0
(Fraud Rate)[b]	(1.7)	(1.6)	(1.6)	

Chi Square[c] = 1.9

*Based on claimants who established benefit years as regular, Arizona UI claimants. Transitional claims from prior benefit years excluded but interstate liable claims included. The claimants included represent all who were determined to be monetarily eligible for benefits and had benefit years established during fiscal years 1972-1976; these benefit years were completed during fiscal years 1973-1977. All numbers are estimated totals for all claimants who met the above criteria, based on a randomly selected sample.

[a]Each cell of the cross tabulation contains two numbers. The upper number represents the estimated cases that fell in the cell. The lower number is the percent of all cases in the entire row accounted for by the estimated cases in the cell. The statistical reliability is extremely high for percentages based on 2,000 or more claimants. For percentages based on 1,000-1,999 claimants, the estimates should be viewed as very good but not exact ones. For percentages based on 500-999 (less than 500) claimants, the estimates should be viewed only as good (rough) approximations of the true totals.

[b]The fraud rate is defined as the percentage of claimants that had one or more detected fraudulent overpayments.

[c]The Chi Square statistic indicates whether the two classification variables in the cross tabulation (number of overpayments and the personal characteristic included in the cross tabulation) are related or not. The larger the value of the statistic (other things equal), the more confident one can be that the two classification variables are related. In particular, a double asterisk (**) is used to denote those statistics that strongly indicate the two classification variables are related; in these instances, it can be asserted that the two variables are related, at a confidence level of .99.

TABLE 4

CROSS TABULATION OF OCCUPATIONAL CATEGORY
BY NUMBER OF FRAUDULENT OVERPAYMENTS DETECTED*

Number of Fraud Overpayments	Occupational Category: Prof./Tech./Mgrl	Cler./Sales	Service	Misc/Farm.	Process/Machine/Bench	Structural	Row Total[a] (Percent)
None	70,245 22.0	59,630 18.6	27,339 8.5	33,929 10.6	48,384 15.1	80,344 25.1	319,871 98.4
1 or more	836 15.8	826 15.6	453 8.5	699 13.2	515 9.7	1,975 37.2	5,304 1.6
Column Total	71,081 21.9	60,456 18.6	27,792 8.5	34,628 10.6	48,899 15.0	82,319 25.3	325,175 100.0
(Fraud Rate)[b]	(1.2)	(1.4)	(1.6)	(2.0)	(1.1)	(2.4)	

Chi Square[c] = 555**

*Based on claimants who established benefit years as regular, Arizona UI claimants. Transitional claims from prior benefit years excluded but interstate liable claims included. The claimants included represent all who were determined to be monetarily eligible for benefits and had benefit years established during fiscal years 1972-1976; these benefit years were completed during fiscal years 1973-1977. All numbers are estimated totals for all claimants who met the above criteria, based on a randomly selected sample.

[a]Each cell of the cross tabulation contains two numbers. The upper number represents the estimated cases that fell in the cell. The lower number is the percent of all cases in the entire row accounted for by the estimated cases in the cell. The statistical reliability is extremely high for percentages based on 2,000 or more claimants. For percentages based on 1,000-1,999 claimants, the estimates should be viewed as very good but not exact ones. For percentages based on 500-999 (less than 500) claimants, the estimates should be viewed only as good (rough) approximations of the true totals.

[b]The fraud rate is defined as the percentage of claimants that had one or more detected fraudulent overpayments.

[c]The Chi Square statistic indicates whether the two classification variables in the cross tabulation (number of overpayments and the personal characteristic included in the cross tabulation) are related or not. The larger the value of the statistic (other things equal), the more confident one can be that the two classification variables are related. In particular, a double asterisk (**) is used to denote those statistics that strongly indicate the two classification variables are related; in these instances, it can be asserted that the two variables are related, at a confidence level of .99.

TABLE 5

CROSS TABULATION OF UI HIGH QUARTER EARNINGS BY NUMBER OF FRAUDULENT OVERPAYMENTS DETECTED*

Number of Fraud Overpayments	High Quarter Earnings Less Than $1000	$1000-1500	$1501-2000	$2001-2500	$2501 +	Row Total[a] (Percent)
None	47,819 14.9	67,135 21.0	59,554 18.6	39,134 12.2	106,229 33.2	319,871 98.4
1 or more	530 10.0	1,211 22.8	919 17.3	539 10.2	2,105 39.7	5,304 1.6
Column Total	48,349 14.9	68,346 21.0	60,473 18.6	39,673 12.2	108,334 33.3	325,175 100.0
(Fraud Rate)[b]	(1.1)	(1.8)	(1.5)	(1.4)	(1.9)	

Chi Square[c] = 183**

*Based on claimants who established benefit years as regular, Arizona UI claimants. Transitional claims from prior benefit years excluded but inter-state liable claims included. The claimants included represent all who were determined to be monetarily eligible for benefits and had benefit years established during fiscal years 1972-1976; these benefit years were completed during fiscal years 1973-1977. All numbers are estimated totals for all claimants who met the above criteria, based on a randomly selected sample.

[a]Each cell of the cross tabulation contains two numbers. The upper number represents the estimated cases that fell in the cell. The lower number is the percent of all cases in the entire row accounted for by the estimated cases in the cell. The statistical reliability is extremely high for percentages based on 2,000 or more claimants. For percentages based on 1,000-1,999 claimants, the estimates should be viewed as very good but not exact ones. For percentages based on 500-999 (less than 500) claimants, the estimates should be viewed only as good (rough) approximations of the true totals.

[b]The fraud rate is defined as the percentage of claimants that had one or more detected fraudulent overpayments.

[c]The Chi Square statistic indicates whether the two classification variables in the cross tabulation (number of overpayments and the personal characteristic included in the cross tabulation) are related or not. The larger the value of the statistic (other things equal), the more confident one can be that the two classification variables are related. In particular, a double asterisk (**) is used to denote those statistics that strongly indicate the two classification variables are related; in these instances, it can be asserted that the two variables are related, at a confidence level of .99.

TABLE 6

CROSS TABULATION OF ARIZONA UI BASE PERIOD EARNINGS BY NUMBER OF FRAUDULENT OVERPAYMENTS DETECTED*

Number of Fraud Overpayments	Base Period Earnings Less Than $2000	$2000-3000	$3001-4000	$4001-6000	$ 6001-10000	$10,000+	Row Total[a] (Percent)
None	36,886 11.5	37,937 11.9	36,860 11.5	65,883 20.6	79,235 24.8	63,070 19.7	319,871 98.4
1 or more	482 9.1	567 10.7	975 18.4	974 18.4	1,320 24.9	986 18.6	5,304 1.6
Column Total	37,368 11.5	38,504 11.8	37,835 11.6	66,857 20.6	80,555 24.8	64,056 19.7	325,175 100.0
(Fraud Rate)[b]	(1.3)	(1.5)	(2.6)	(1.5)	(1.6)	(1.5)	

Chi Square[c] = 260**

*Based on claimants who established benefit years as regular. Arizona UI claimants. Transitional claims from prior benefit years excluded but inter-state liable claims included. The claimants included represent all who were determined to be monetarily eligible for benefits and had benefit years established during fiscal years 1972-1976; these benefit years were completed during fiscal years 1973-1977. All numbers are estimated totals for all claimants who met the above criteria, based on a randomly selected sample.

[a]Each cell of the cross tabulation contains two numbers. The upper number represents the estimated cases that fell in the cell. The lower number is the percent of all cases in the entire row accounted for by the estimated cases in the cell. The statistical reliability is extremely high for percentages based on 2,000 or more claimants. For percentages based on 1,000-1,999 claimants, the estimates should be viewed as very good but not exact ones. For percentages based on 500-999 (less than 500) claimants, the estimates should be viewed only as good (rough) approximations of the true totals.

[b]The fraud rate is defined as the percentage of claimants that had one or more detected fraudulent overpayments.

[c]The Chi Square statistic indicates whether the two classification variables in the cross tabulation (number of overpayments and the personal characteristic included in the cross tabulation) are related or not. The larger the value of the statistic (other things equal), the more confident one can be that the two classification variables are related. In particular, a double asterisk (**) is used to denote those statistics that strongly indicate the two classification variables are related; in these instances, it can be asserted that the two variables are related, at a confidence level of .99.

TABLE 7

CROSS TABULATION OF THE RATIO OF ARIZONA BASE PERIOD TO HIGH QUARTER EARNINGS BY NUMBER OF FRAUDULENT OVERPAYMENTS DETECTED*

Number of Fraud Overpayments	Earnings Ratio: Less Than 2	2.0-2.5	2.6-3.0	3.1+	Row Total[a] (Percent)
None	40,507 13.0	64,345 20.6	60,905 19.5	146,536 46.9	312,293 98.4
1 or more	707 14.1	1,202 23.9	1,201 23.9	1,910 38.0	5,020 1.6
Column Total	41,214 13.0	65,547 20.7	62,106 19.6	148,446 46.8	317,313 100.0
(Fraud Rate)[b]	(1.7)	(1.8)	(1.9)	(1.3)	

Chi Square[c] = 163**

Number of Missing Observations[d] = 7862

*Based on claimants who established benefit years as regular, Arizona UI claimants. Transitional claims from prior benefit years excluded but interstate liable claims included. The claimants included represent all who were determined to be monetarily eligible for benefits and had benefit years established during fiscal years 1972-1976; these benefit years were completed during fiscal years 1973-1977. All numbers are estimated totals for all claimants who met the above criteria, based on a randomly selected sample.

[a]Each cell of the cross tabulation contains two numbers. The upper number represents the estimated cases that fell in the cell. The lower number is the percent of all cases in the entire row accounted for by the estimated cases in the cell. The statistical reliability is extremely high for percentages based on 2,000 or more claimants. For percentages based on 1,000-1,999 claimants, the estimates should be viewed as very good but not exact ones. For percentages based on 500-999 (less than 500) claimants, the estimates should be viewed only as good (rough) approximations of the true totals.

[b]The fraud rate is defined as the percentage of claimants that had one or more detected fraudulent overpayments.

[c]The Chi Square statistic indicates whether the two classification variables in the cross tabulation (number of overpayments and the personal characteristic included in the cross tabulation) are related or not. The larger the value of the statistic (other things equal), the more confident one can be that the two classification variables are related. In particular, a double asterisk (**) is used to denote those statistics that strongly indicate the two classification variables are related; in these instances, it can be asserted that the two variables are related, at a confidence level of .99.

[d]Missing observations result because some claimants had zero high quarter earnings (because of combined wages). The earnings ratio can not be calculated for such claimants.

TABLE 8

CROSS TABULATION OF THE UI WEEKLY BENEFIT AMOUNT BY NUMBER OF FRAUDULENT OVERPAYMENTS DETECTED*

Number of Fraud Overpayments	Weekly Benefit Amount: Less Than $35	$35-50	$51-65	$66-85	Row Total[a] (Percent)
None	32,058 10.0	52,296 16.3	118,102 36.9	117,415 36.7	319,871 98.4
1 or more	366 6.9	913 17.2	1,786 33.7	2,239 42.2	5,304 1.6
Column Total	32,424 10.0	53,209 16.4	119,888 36.9	119,654 36.8	325,175 100.0
(Fraud Rate)[b]	(1.1)	(1.7)	(1.5)	(1.9)	

Chi Square[c] = 111**

*Based on claimants who established benefit years as regular, Arizona UI claimants. Transitional claims from prior benefit years excluded but interstate liable claims included. The claimants included represent all who were determined to be monetarily eligible for benefits and had benefit years established during fiscal years 1972-1976; these benefit years were completed during fiscal years 1973-1977. All numbers are estimated totals for all claimants who met the above criteria, based on a randomly selected sample.

[a]Each cell of the cross tabulation contains two numbers. The upper number represents the estimated cases that fell in the cell. The lower number is the percent of all cases in the entire row accounted for by the estimated cases in the cell. The statistical reliability is extremely high for percentages based on 2,000 or more claimants. For percentages based on 1,000-1,999 claimants, the estimates should be viewed as very good but not exact ones. For percentages based on 500-999 (less than 500) claimants, the estimates should be viewed only as good (rough) approximations of the true totals.

[b]The fraud rate is defined as the percentage of claimants that had one or more detected fraudulent overpayments.

[c]The Chi Square statistic indicates whether the two classification variables in the cross tabulation (number of overpayments and the personal characteristic included in the cross tabulation) are related or not. The larger the value of the statistic (other things equal), the more confident one can be that the two classification variables are related. In particular, a double asterisk (**) is used to denote those statistics that strongly indicate the two classification variables are related; in these instances, it can be asserted that the two variables are related, at a confidence level of .99.

TABLE 9

CROSS TABULATION OF THE UI WEEKLY BENEFIT AMOUNT AS A PERCENT OF ARIZONA AVERAGE WEEKLY EARNINGS IN THE BASE PERIOD BY NUMBER OF FRAUDULENT OVERPAYMENTS DETECTED*

Number of Fraud Overpayments	WBA as Percent of Weekly Earnings[a] Less Than 40%	40%-60%	61%-80%	Over 80%	Row Total[b] (Percent)
None	67,134 21.0	93,155 29.1	71,946 22.5	87,383 27.3	319,618 98.4
1 or more	982 18.5	1,305 24.6	1,490 28.1	1,527 28.8	5,304 1.6
Column Total	68,116 21.0	94,460 29.1	73,436 22.6	88,910 27.4	324,922 100.0
(Fraud Rate)[c]	(1.4)	(1.4)	(2.0)	(1.7)	

Chi Square[d] = 128**

Number of Missing Observations = 253.

*Based on claimants who established benefit years as regular, Arizona UI claimants. Transitional claims from prior benefit years excluded but interstate liable claims included. The claimants included represent all who were determined to be monetarily eligible for benefits and had benefit years established during fiscal years 1972-1976; these benefit years were completed during fiscal years 1973-1977. All numbers are estimated totals for all claimants who met the above criteria, based on a randomly selected sample.

[a]Weekly earnings are defined as base period earnings divided by 52.

[b]Each cell of the cross tabulation contains two numbers. The upper number represents the estimated cases that fell in the cell. The lower number is the percent of all cases in the entire row accounted for by the estimated cases in the cell. The statistical reliability is extremely high for percentages based on 2,000 or more claimants. For percentages based on 1,000-1,999 claimants, the estimates should be viewed as very good but not exact ones. For percentages based on 500-999 (less than 500) claimants, the estimates should be viewed only as good (rough) approximations of the true totals.

[c]The fraud rate is defined as the percentage of claimants that had one or more detected fraudulent overpayments.

[d]The Chi Square statistic indicates whether the two classification variables in the cross tabulation (number of overpayments and the personal characteristic included in the cross tabulation) are related or not. The larger the value of the statistic (other things equal), the more confident one can be that the two classification variables are related. In particular, a double asterisk (**) is used to denote those statistics that strongly indicate the two classification variables are related; in these instances, it can be asserted that the two variables are related at a confidence level of .99.

TABLE 10

CROSS TABULATION OF POTENTIAL DURATION OF REGULAR BENEFITS BY NUMBER OF FRAUDULENT OVERPAYMENTS DETECTED*

Number of Fraud Overpayments	Potential Duration Less than 20 weeks	20-25 weeks	26 wks	Row Total[a] (Percent)
None	61,849 19.3	56,921 17.8	201,101 62.9	319,871 98.4
1 or more	1,183 22.3	1,021 19.2	3,100 58.4	5,304 1.6
Column Total	63,032 19.4	57,942 17.8	204,201 62.8	325,175 100.0
(Fraud Rate)[b]	(1.9)	(1.8)	(1.5)	

Chi Square[c] = 46**

*Based on claimants who established benefit years as regular, Arizona UI claimants. Transitional claims from prior benefit years excluded but interstate liable claims included. The claimants included represent all who were determined to be monetarily eligible for benefits and had benefit years established during fiscal years 1972-1976; these benefit years were completed during fiscal years 1973-1977. All numbers are estimated totals for all claimants who met the above criteria, based on a randomly selected sample.

[a]Each cell of the cross tabulation contains two numbers. The upper number represents the estimated cases that fell in the cell. The lower number is the percent of all cases in the entire row accounted for by the estimated cases in the cell. The statistical reliability is extremely high for percentages based on 2,000 or more claimants. For percentages based on 1,000-1,999 claimants, the estimates should be viewed as very good but not exact ones. For percentages based on 500-999 (less than 500) claimants, the estimates should be viewed only as good (rough) approximations of the true totals.

[b]The fraud rate is defined as the percentage of claimants that had one or more detected fraudulent overpayments.

[c]The Chi Square statistic indicates whether the two classification variables in the cross tabulation (number of overpayments and the personal characteristic included in the cross tabulation) are related or not. The larger the value of the statistic (other things equal), the more confident one can be that the two classification variables are related. In particular, a double asterisk (**) is used to denote those statistics that strongly indicate the two classification variables are related; in these instances, it can be asserted that the two variables are related at a confidence level of .99.

TABLE 11

CROSS TABULATION OF SPELLS OF UNEMPLOYMENT BY NUMBER OF FRAUDULENT OVERPAYMENTS DETECTED*

Number of Fraud Overpayments	Spells of Unemployment					Row Total[a] (Percent)
	Zero	One	Two	Three	Four or more	
None	23,955 7.5	121,049 37.8	72,829 22.8	55,773 17.4	46,265 14.5	319,871 98.4
1 or more	-0- 0.0	1,121 21.1	1,203 22.7	1,526 28.8	1,454 27.4	5,304 1.6
Column Total	23,955 7.4	122,170 37.6	74,032 22.8	57,299 17.6	47,719 14.7	325,175 100.0
(Fraud Rate)[b]	(-)	(0.9)	(1.6)	(2.7)	(3.0)	

Chi Square[c] = 1761**

*Based on claimants who established benefit years as regular, Arizona UI claimants. Transitional claims from prior benefit years excluded but interstate liable claims included. The claimants included represent all who were determined to be monetarily eligible for benefits and had benefit years established during fiscal years 1972-1976; these benefit years were completed during fiscal years 1973-1977. All numbers are estimated totals for all claimants who met the above criteria, based on a randomly selected sample.

[a]Each cell of the cross tabulation contains two numbers. The upper number represents the estimated cases that fell in the cell. The lower number is the percent of all cases in the entire row accounted for by the estimated cases in the cell. The statistical reliability is extremely high for percentages based on 2,000 or more claimants. For percentages based on 1,000-1,999 claimants, the estimates should be viewed as very good but not exact ones. For percentages based on 500-999 (less than 500) claimants, the estimates should be viewed only as good (rough) approximations of the true totals.

[b]The fraud rate is defined as the percentage of claimants that had one or more detected fraudulent overpayments.

[c]The Chi Square statistic indicates whether the two classification variables in the cross tabulation (number of overpayments and the personal characteristic included in the cross tabulation) are related or not. The larger the value of the statistic (other things equal), the more confident one can be that the two classification variables are related. In particular, a double asterisk (**) is used to denote those statistics that strongly indicate the two classification variables are related; in these instances, it can be asserted that the two variables are related at a confidence level of .99.

TABLE 12

CROSS TABULATION OF WEEKS WITH REPORTED DEDUCTIBLE EARNINGS BY NUMBER OF FRAUDULENT OVERPAYMENTS DETECTED*

Number of Fraud Overpayments	Weeks With Deductible Earnings			Row Total[a] (Percent)
	Zero	1-2	3 or More	
None	196,424 61.4	94,799 29.6	28,648 9.0	319,871 98.4
1 or more	2,920 55.1	1,983 37.4	401 7.6	5,304 1.6
Column Total	199,344 61.3	96,782 29.8	29,049 8.9	325,175 100.0
(Fraud Rate)[b]	(1.5)	(2.0)	(1.4)	

Chi Square[c] = 151**

*Based on claimants who established benefit years as regular, Arizona UI claimants. Transitional claims from prior benefit years excluded but interstate liable claims included. The claimants included represent all who were determined to be monetarily eligible for benefits and had benefit years established during fiscal years 1972-1976; these benefit years were completed during fiscal years 1973-1977. All numbers are estimated totals for all claimants who met the above criteria, based on a randomly selected sample.

[a]Each cell of the cross tabulation contains two numbers. The upper number represents the estimated cases that fell in the cell. The lower number is the percent of all cases in the entire row accounted for by the estimated cases in the cell. The statistical reliability is extremely high for percentages based on 2,000 or more claimants. For percentages based on 1,000-1,999 claimants, the estimates should be viewed as very good but not exact ones. For percentages based on 500-999 (less than 500) claimants, the estimates should be viewed only as good (rough) approximations of the true totals.

[b]The fraud rate is defined as the percentage of claimants that had one or more detected fraudulent overpayments.

[c]The Chi Square statistic indicates whether the two classification variables in the cross tabulation (number of overpayments and the personal characteristic included in the cross tabulation) are related or not. The larger the value of the statistic (other things equal), the more confident one can be that the two classification variables are related. In particular, a double asterisk (**) is used to denote those statistics that strongly indicate the two classification variables are related; in these instances, it can be asserted that the two variables are related at a confidence level of .99.

END