

A Scoring System for Sales Tax Audit Selection

Thomas J. Winn, Jr.

**Audit Division Headquarters, Office of the State Comptroller of Public Accounts
Austin, Texas**

Introduction & Overview

Tax audits are investigative procedures which are conducted to ensure that taxpayers have paid the correct amount of tax due. The objectives of the audit program are to promote taxpayer compliance, and also to increase net tax revenue. It wouldn't be cost-effective for the Comptroller's Office to hire enough auditors to ensure complete audit coverage.

Many of the largest sales tax payers are audited about every four years. Combined, these *Priority I* taxpayers account for approximately the top two percent of all sales tax payers (and about 65 percent of total reported tax). The audit selection problem at the Comptroller's Office is to choose taxpayers from among the remaining 98 percent of the taxpayers, the *Priority II* accounts.

The purpose of this report is to describe a scoring system which furnishes some useful comparative information which is used in the selection process. The system utilizes statistically-based procedures for estimating the potential audit productivity of all sales tax payers under consideration.

All of the computer programs are written in the SAS language. The SAS System is used for accessing data from various types of data stores, for data manipulation, for statistical analysis, and for preparing reports. At the present time, there are sixteen computer programs in the system. The system is completely re-developed semi-annually.

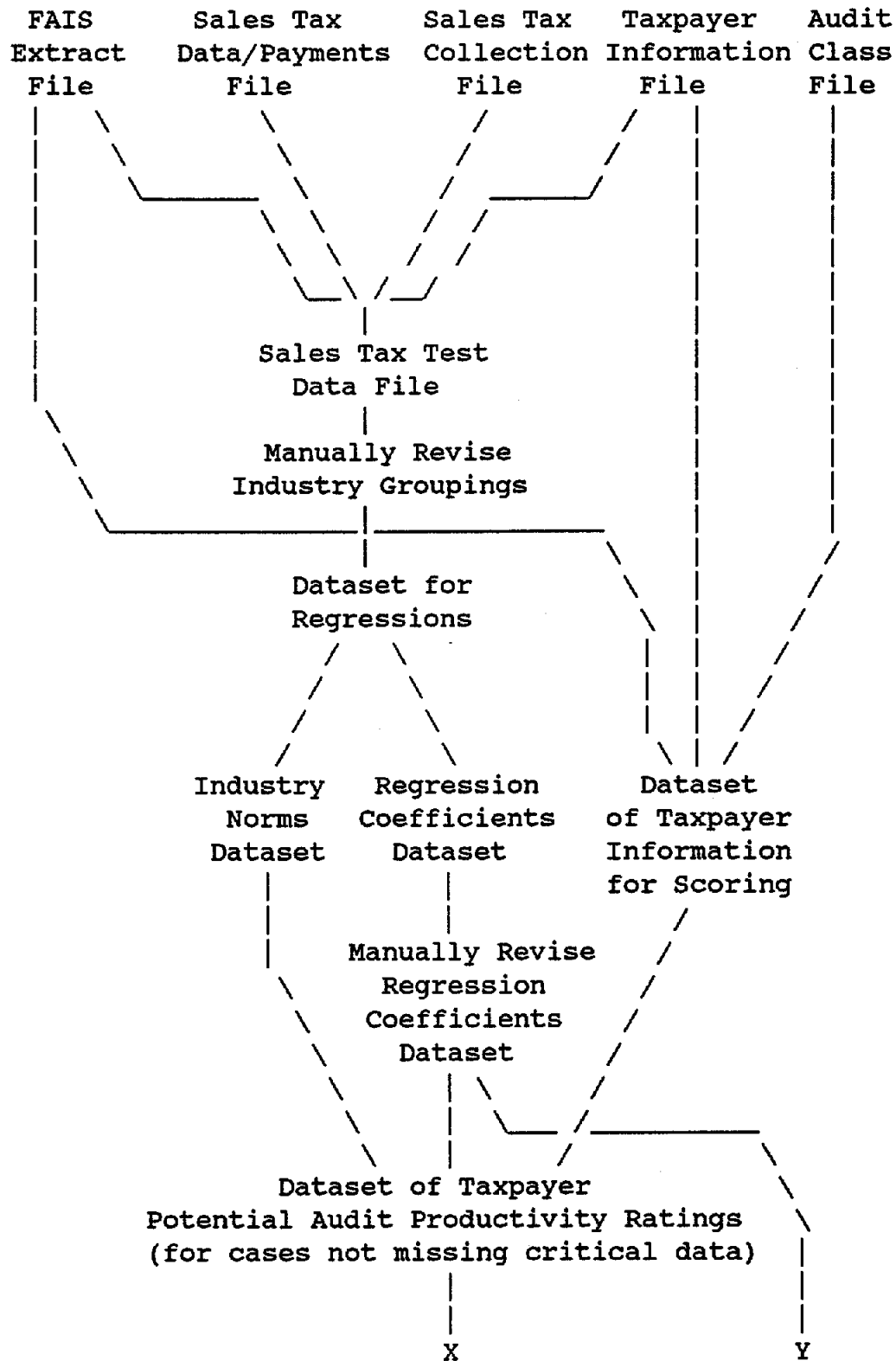
An Outline of the Sales Tax Audit Select System

There are seven steps in the system:

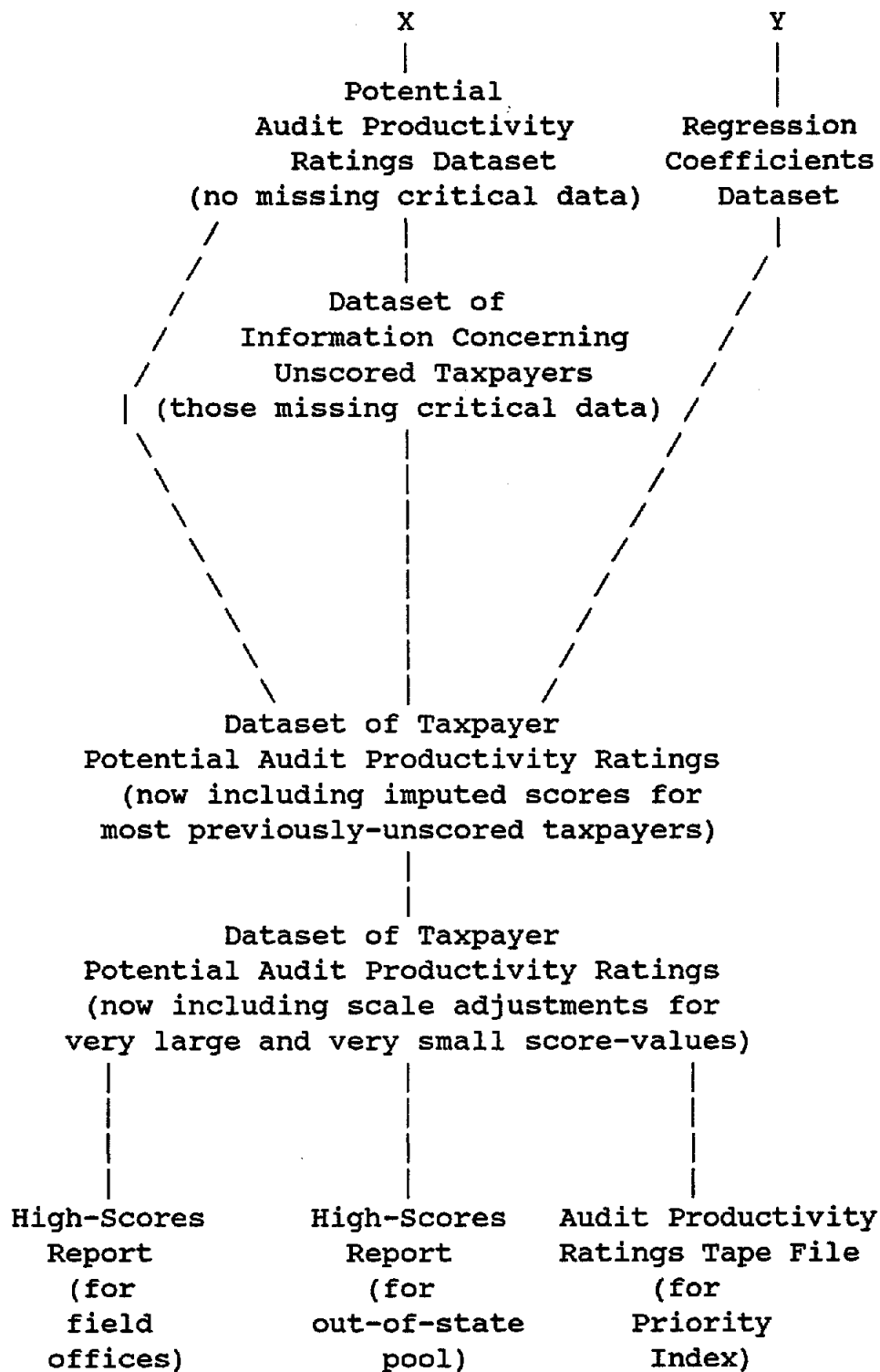
1. Gather Taxpayer Data Corresponding to Recent Audits,
2. Establish Industrial Groupings,
3. Establish Industrial Norms,
4. Estimate Audit Productivity Models,
5. Gather Data for Ratings,
6. Estimate Audit Productivity Ratings,
7. Prepare Reports and Tape File.

Here is a schematic diagram of the current system:

SALES TAX AUDIT SELECTION SYSTEM



SALES TAX AUDIT SELECTION SYSTEM (continued)



The Components of the Sales Tax Audit Select System

Preliminary Preparations.

Before the actual audit selection process begins, a few preliminaries must be taken care of. First, recent historical data regarding audit hours and administrative expenditures pertaining to auditing for audit field offices, audit headquarters, data processing, enforcement, administrative law judges, and the legal division need to be gathered and analyzed, to produce an estimate for administrative costs per audit hour. Administrative costs per audit hour is used for indexing the potential audit productivity estimates which will be produced by the system for each sales taxpayer.

Next, values of the Implicit Price Deflator for Gross Domestic Product are obtained, for the previous five calendar years and for the current calendar year. The deflators are used for deriving inflation factors which, in turn, are used for adjusting dollar amounts for audits completed in prior years to amounts which are comparable to current-year values. Finally, programmers in the Comptroller's Applications Systems Division are contacted to create the Class File. The Class File contains information about taxpayers who are subject to audit selection. It includes data about the taxpayers' characteristics, taxes, four years of sales and franchise tax information, as well as information for any prior audits.

Step 1 - Gather Data Corresponding to Recent Audits

A program is run which selects and sorts data from the Audit Field Assignment Inventory System Extract File. For purposes of selection, the only audits which are included are those which were completed less than six years, and more than six months, from the month in which the program is run,

Another program builds a dataset containing information concerning tax return data records from the three Sales Tax Data/Payments VSAM Files.

A dataset is built containing information from selected collection records in the purged and current versions of the Sales Tax Collection File. The dataset contains selected information about certain types of collection records. The records which are included correspond to "problem" accounts, with which the agency has experienced some difficulty in receiving amounts determined to be due to the State.

Results from the three previously-mentioned programs are combined with some general taxpayer information, and an industrial groupings dataset, to produce the Sales Tax Test Data File. The Sales Tax Test Data File contains information

regarding the taxpayers' SIC code, outlets, duration-in-business, and reported dollar amounts, as well as such audit information as hours, results, period, deficiency amounts, and important dates. An inflation adjustment is included, to make the dollar amounts for different time periods comparable. Depending on the audit completion date, amounts for previous years are inflated to current-calendar-year dollars.

The purpose of the Sales Tax Test Data File is to gather together the data for variables which are potentially the most important determinants of audit productivity. These data are needed for estimating multiple regression models for each industrial grouping which will be used for determining the audit productivity ratings.

Step 2 - Establish Industrial Groupings.

The next step in the sales tax audit selection system is to establish the industrial groupings which are used for analyzing the taxpayer population. This is accomplished by means of a recursive procedure. Different combinations of SIC codes are used, until the printed report -- which displays the number of audits by SIC for each industrial grouping -- satisfies the desired criteria.

The industrial groupings should be chosen so that the members of each grouping are involved in similar business activities, their types of taxable transactions should be similar, and there should have been numerous (more than just a few) previous audits performed for each grouping.

The fundamental idea is that each industrial grouping should be a fairly homogeneous collective, which includes enough audits so as to be deserving of further consideration. Currently, there are 74 industry groupings.

Data pertaining to audits which are believed to be typical in each industrial grouping are selected from the Sales Tax Test Data File. Regional datasets also are constructed, which are useful for analyzing audit information on the basis of both industry grouping and geographical location. Overall audit productivity measures (the audit deficiency rate, and average dollars per audit hour) for each grouping in each economic region are calculated.

A single program performs this analysis, generates the reports, and creates datasets for each of the six economic regions of Texas, plus a "region" comprised of all out-of-state taxpayers, as well as for the entire state. These data are needed for estimating multiple regression models which will be used for determining the audit productivity ratings. Coding for the regions is streamlined by using the capabilities of the SAS Macro Language.

Step 3 - Establish Industrial Norms.

After the industry groupings are determined, the next step is to establish norms for audit productivity in each industry grouping, for comparative purposes.

A SAS program reads information regarding individual taxpayer audits. Each audit is characterized as being either productive or non-productive, depending on the comparative relationship between the dollars per audit hour and an estimate of the current-year average expense per audit hour (the previous year's total expenditures attributable in some manner to audit, times a factor to account for inflation in the level of prices from the previous year to the current year, divided by the previous year's total audit hours).

For each industry grouping, the program calculates the ratios of net taxable sales to gross sales, and use-taxable purchases to the sum of gross sales and use-taxable purchases, for productive and non-productive audits. The means of the two ratios are calculated for both productive and non-productive audits, and a t-test is performed (in a DATA step) to test the hypothesis that the mean ratio pairs are equal for the productive and the non-productive audits. If both ratios are significantly different, then the program identifies the best ratio. The program calculates a typical range for each ratio, as well as the average dollars per audit hour, and the deficiency rate, in each industry grouping.

A "permanent" SAS dataset on tape is created, which is used in Audit Select Step 6, as a preparation for printing the High-Scores Report.

Step 4 - Estimate Audit Productivity Models

The fourth step in the audit selection system for sales tax is to identify the variables and to estimate the parameters of statistically-optimal models for deficiency dollars per audit-hour for each of the industrial groupings, as well as for those taxpayers whose SIC codes are missing.

Each of the audit productivity models generated for the industry groupings identify a small number of variables which have the greatest combined explanatory power with regard to potential audit productivity. Different weighted combinations chosen from 13 variables are selected for each industry grouping, based on statistics. The variables which are considered for inclusion in the regression models are:

- * Number of years the business has been open,
- * Square of the number of years in business,

- * Percent change in number of outlets in the preceding 18 months,
- * Number of different SIC codes for outlets,
- * Number of open records before generation of audit,
- * Annualized audit period gross sales (inflation-adjusted),
- * Annualized audit period deductions (adjusted for inflation)),
- * Annualized audit period use-taxable purchases (inflation-adjusted),
- * Annualized amount subject to sales tax (adjusted for inflation),
- * Percent of gross sales that is net taxable sales,
- * Percent of gross sales that is taxable purchases,
- * Taxable purchases as percent of gross sales plus purchases,
- * Indicator variable for corporations.

PROC MEANS is used to calculate some simple descriptive statistics, which are used to exclude outliers (extreme values more than 1.96 standard deviations of the distribution of the sample averages from the mean -- about 2 1/2 percent off of each "tail") from the test data. Then, PROC REG is invoked, to perform stepwise regression analysis for modeling audit productivity in each of the industry groupings. PROC REG with SELECTION=STEPWISE implements a procedure which begins with no variables in the particular model under consideration, adds variables to the model on a one-by-one basis until all of the variables in the model produce significant F statistics, and at each iteration drops any variable which may have been included previously in the model but which produce F statistics which are so low as to indicate that variable's contribution to the model is not sufficient to justify its continued inclusion, relative to the other variables in the model. The significance levels for entry and retention in the models are specified beforehand, usually 15 or 20 percent. PROC REG writes the regression coefficients to an TYPE=EST data set on disk, which must be carefully checked for suitability and completeness.

It sometimes happens that none of the data for certain industry groupings will satisfy the specified significance level criterion for inclusion in a regression model. Any industry grouping for which regression models are not estimated by the stepwise method, or for which the stepwise models are unsatisfactory, are fitted separately and then are added to the Regressions Coefficients dataset, using PROC FSEDIT. The "Statistics for Entry" table produced by PROC REG with the MODEL statement STEPWISE and DETAILS options is helpful in identifying regressor variables for any missing industry groupings. Parameters can be appended to the Regression Coefficients dataset by a DATA step, or manually using PROC FSEDIT.

In the future, multiple regression models for potential audit productivity will be constructed for each industrial grouping, and in each economic region.

Step 5 - Gather Data for Ratings.

After the data for performing the regressions are assembled, and the regression coefficients are determined, the next step in the process of calculating the potential audit productivity ratings is to gather the data needed for making the actual ratings, or for printing the audit select reports. An important difference between Step 5 and Step 1 is that the Sales Tax Test Data File contains pertinent data regarding taxpayers *who have been audited recently*, whereas Step 5 brings in similar data for *all* taxpayers who are subject to sales tax audit selection.

Step 6 - Estimate Audit Productivity Ratings.

After the regression coefficients for the audit productivity models are determined, and the explanatory data for all taxpayers to be evaluated are gathered, the next step in developing the sales tax audit selection system is to estimate the potential audit productivity for all sales tax payers to be scored. The "heart" of the program is an application of PROC SCORE, which multiplies values from the regression coefficients dataset times corresponding values from the dataset containing taxpayer data, and then produces linear combinations of these multiplicative products. The ability of PROC SCORE to do BY-group processing enables the program to evaluate each taxpayer account according to its particular industry grouping.

The ratings are indexed according to a per-audit-hour measure which reflects administrative costs associated with tax audits. Therefore, a score of 100 would be interpreted as a probable "break-even" situation -- one might expect that the audit deficiency resulting from the audit would just equal the total cost of performing the audit, plus a few other audit-related expenses. Scores greater than 100 could be expected to result in a net gain to the State. Scores less than 100 would mean that administrative costs probably would exceed the audit deficiency assessment.

The scoring program also brings in the audit productivity norms from Audit Select Step 3, and uses them to determine whether or not a particular ratio, which has proven itself to be an important indicator of audit productivity, is outside of the range of values which might be expected of the ratio for "typical" businesses in the industry grouping.

It is not uncommon for the data pertaining to an individual taxpayer in a particular industry grouping to be missing one or more of the values of variables which were determined in Step 4 to be important determinants of audit productivity. In this event, the scoring program would not be able to calculate a rating. In some

instances, it may be plausible to impute values for certain of these missing critical variables, and to estimate scores. This is currently being done in a separate program.

It is possible for the regression equations to produce some anomalous results -- that is, certain values of the independent variables can generate either very small (large negative) or very large scores. Because of the existence of a practical limit for storage and printing purposes of the width of score values, a DATA step is used to adjust the scale of very large or very small ratings accordingly.

Step 7 - Prepare Reports and Tape File.

A report-writer program generates the "High-Scores Report" for the field audit offices, and for the out-of-state pool. This report lists several taxpayers which received high scores in the potential audit productivity ratings analysis. Only those Priority II taxpayers with the highest ratings in the various industries are printed for each audit office.

Along with the High Scores Report, the field offices also receive a listing of "confidence measures", to help them with their audit select decision-making. For each audit select industrial grouping, these represent a subjective assessment of the quality of the audit select scores as general predictors of potential audit productivity. The confidence measures are derived from goodness-of-fit statistics.

As a final stage, a SAS program is used to write the audit select scores out to an OS dataset on tape for use by the Applications Systems Division in creating the Priority Index.

Evaluation of the Sales Tax Audit Selection System

Periodically, ratings from the sales tax audit selection scoring system are compared with actual audit results. It is disappointing to note that, in general, there is no significant correlation between the audit select scores assigned by the system and the actual deficiency dollars per audit hour. However, for some industries, the audit select scores do seem to be better predictors of the level of audit productivity than for other industry groupings. An analysis of variance revealed that deficiency dollars per audit hour is significantly larger for higher scoring taxpayers than for the lower-scoring taxpayers. A non-parametric Kruskal-Wallis test yielded a comparable result. Annual average deficiency dollars per audit hour has increased each year since the system was initiated. Moreover, a survey of the field office audit managers and audit selectors

confirmed that these users of the system's results find the information generated by the system to be useful.

Summary

Using the SAS System, a separate regression model is fitted for each major industrial grouping. These models are used to evaluate the potential audit productivity of each taxpayer, which are indexed according to a measure which reflects administrative costs associated with tax audits. Audit field office managers and audit selectors report that they are well-satisfied with the information generated by the system.