NBER WORKING PAPER SERIES

THE RELATIONSHIP BETWEEN STATE AND FEDERAL TAX AUDITS

James Alm
Brian Erard
Jonathan S. Feinstein

Working Paper No. 5134

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 1995

This paper is part of NBER's research program in Public Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #5134
May 1995

THE RELATIONSHIP BETWEEN STATE AND FEDERAL TAX AUDITS

ABSTRACT

In this paper we present an econometric analysis of state and federal tax audits. We first present results from a survey of state tax administrators. The survey results indicate that most state tax audit programs are small and rely extensively on information provided by the IRS, although some programs are large and sophisticated. We then present results from a detailed econometric analysis of Oregon state and federal tax returns and tax audits for tax year 1987. Our analysis generates three main conclusions. First, Oregon state and IRS selection criteria are similar, but not identical, suggesting that both tax agencies might benefit from greater sharing of information, especially in some audit classes. Second, Oregon state and IRS audit assessments are strongly positively correlated, as expected. Third, we estimate the shadow values associated with providing additional audit resources to the Oregon Department of Revenue and the IRS in various audit classes, and find that for the IRS the shadow values range from two to five dollars, while for Oregon the values range from one to three dollars.

James Alm
Department of Economics
Campus Box 256
Room 112, Economics Building
University of Colorado
Boulder, CO 80309-0256

Brian Erard
Department of Economics
1125 Colonel By Drive
Carleton University
Ottawa, Ontario K1S 5B6
CANADA

Jonathan S. Feinstein
Yale School of Management
Box 20-8200
New Haven, CT 06520-8200
and NBER

1. Introduction

In this paper we present an analysis of state and federal individual income tax enforcement programs. We develop an econometric model of state and federal tax audit selection decisions and audit assessments and then present a detailed empirical analysis of Oregon state and federal tax audits. We investigate the degree to which the federal and state tax authorities employ similar audit selection criteria, the correlation between state and federal noncompliance, and the allocation of state audit resources between independent audits and "piggyback" audits based on federal enforcement efforts.

The majority of studies in the empirical academic literature on tax compliance have investigated compliance with central government tax obligations. We believe the study of state tax compliance and enforcement and the relationship between state and federal compliance and enforcement is important for several reasons. In the United States the magnitude of noncompliance with both state and local taxes is probably at least as large, as a percentage of total legal obligations, as the magnitude of noncompliance with federal tax obligations. In many cases a household's decision about how much of its state tax liability to pay may be closely related to its decision about how much of its federal tax liability to pay, especially when the state and federal tax bases are similarly defined, as they are in many states. As a result a state tax authority may find that an effective enforcement strategy is to "piggyback" on federal enforcement efforts, following up on federal audit cases for which a large amount of noncompliance is detected. The fact that state noncompliance is likely to be highly correlated with federal noncompliance has some important implications. First, it influences how state tax authorities allocate their limited tax enforcement budget between independent audits and piggyback audits. Second, it raises a host of policy questions about the proper balance and relationship between state and federal tax enforcement programs, questions that are especially relevant in an era in which the size of the federal government may be reduced and the role of state governments in providing basic goods and services may increase. Our analysis is intended to address all of these issues, as well as other related topics.

Since little information has been published about state tax enforcement programs, we decided to conduct a survey of the fifty states to learn more about their audit programs (a copy of this survey is available upon request).[1] To date we have received responses from thirty-two states.[2] Table 1, based on these survey responses, provides some information about state enforcement programs and compares these programs to Internal Revenue Service (IRS) enforcement efforts. As indicated by the figures in the table, state enforcement levels are quite small in comparison with federal enforcement levels, especially in regards to the individual income tax. Thus, state budgets for enforcement and tax administration are smaller than average IRS state-level budgets, state audit rates are generally much smaller than the federal audit rate, and the magnitude of assessments (the total of additional taxes, interest, and penalties) generated by independent state audits is much smaller than that generated by federal audits. Our focus in this paper is on the relationship between state and federal tax enforcement efforts. Our survey results indicate that the states rely extensively on information provided by the Internal Revenue Service (IRS) through its revenue agent reports (RARs) on federal audits and its CP2000 information returns matching program. In particular, as shown in table 1, on average states conduct more "piggyback" audits based on federal information than independent audits, and the total magnitude of assessments generated by piggyback audits is larger than that generated by independent audits. Other results from the survey (not reported in table 1) indicate that, although the states obtain much information from the IRS, they provide relatively little information in return. The states also seem to follow somewhat different audit selection procedures than the IRS, relying less on computer algorithms or statistical methods for the selection of returns for audit, and instead often choosing returns based on previously productive accounts, random selection, specific tax items or filer characteristics, or a comparison of information on state returns with information from other sources. A striking result of the survey is the degree of variation in state income tax enforcement efforts. Some states undertake no independent audit efforts, even though they have significant income tax programs, while others, such as Oregon (upon whom we concentrate our empirical analysis in this paper), have quite ambitious independent enforcement programs.

The survey results indicate that states rely extensively on federal enforcement efforts and that state and federal audit selection procedures may be similar, but are not identical. The results also suggest that there may be much room for changing and improving state audit programs. Our analysis in this paper is designed to investigate these findings in greater detail.

Our behavioral model accounts for state and federal tax audit decisions and assessments, including state piggyback audits, in a common framework. The model consists of two periods. In the first period the federal and state tax authorities simultaneously and independently select cases for audit and make revenue assessments. In the second period the state authority has the option of performing a "piggyback audit" on any case for which. in the first period, the federal authority performed an audit and the state did not. We assume that each authority selects a case for audit whenever its expectation of the revenue to be earned from conducting the audit exceeds the shadow cost of audit resources. We also assume that each authority observes a private signal of the revenue associated with a case prior to making its audit decision, and we allow the signals of the two authorities to be correlated. Our empirical framework includes a careful specification of the revenue assessment distribution faced by each authority, and it contains a rich stochastic structure that allows for several different kinds of correlations between federal and state assessments.

We estimate our model using a dataset that combines information on federal and state audit programs in Oregon. Oregon is a medium-sized state, containing approximately one percent of all U.S. households. It is a good state in which to study individual income tax compliance and enforcement, both because it collects more than two-thirds of its total tax revenues from the individual income tax (versus only about one-third on average for other states) and because it has a very active tax enforcement program. Our data includes detailed federal and Oregon state tax return information for 43,500 Oregon filers for tax year 1987, as well as audit results for the 4,400 filers in the sample whose 1987 federal returns were selected for an IRS audit and the 2,800 filers whose 1987 state returns were selected for either an independent audit or a piggyback audit by the Oregon Department of Revenue (ODR).

We report estimates for three separate audit classes: a business class, a farm class, and a nonbusiness, nonfarm class. Four aspects of our results are particularly interesting. First, we find that state and federal assessments are strongly positively correlated, as expected. Second, we find that the IRS and ODR audit selection criteria overlap, but only partially. In the business and farm classes, each authority seems to rely heavily on its private signal in deciding whether or not to select a case for audit. The private signals of the two agencies are highly correlated within these classes, indicating a substantial overlap between federal and state information. However, each agency's signal appears to contain information about the other agency's revenue assessment that is unknown to the other agency, a finding which suggests that the federal and state tax authorities could improve their audit selection procedures in these classes by exchanging more information. In the nonbusiness, nonfarm class the authorities seem to rely less on their private signals and more on filers' reports of certain tax return line-items in making their audit selection decisions, so information sharing seems less important. Third, using our results we are able to estimate the shadow value associated with providing additional audit resources to each tax authority. We estimate that the shadow value associated with providing the IRS with an additional dollar of audit resources is approximately five dollars for the business class, two and one-half dollars for the farm class, and four dollars for the nonbusiness, nonfarm class. We estimate that the shadow value associated with providing the ODR with an additional dollar of audit resources is approximately one, two, and three dollars for the business, farm, and nonbusiness, nonfarm classes, respectively. For ODR piggyback audits, we estimate that the shadow value of additional resources is between two and three dollars for the business and nonbusiness, nonfarm classes; however, we are unable to reliably estimate the value for the farm class. Our results indicate that the IRS might be able to increase its audit revenues by reallocating some of its audit resources from farm audits to business audits. In contrast, it would appear that the ODR could increase its audit revenues by shifting some of its resources out of business audits and into nonbusiness, nonfarm audits. The piggyback audit results indicate that the ODR might also benefit by performing more piggyback audits and less independent audits within the business and farm classes. Fourth, we report a number of interesting findings from a detailed examination of Oregon's audit programs, which suggest that the state could increase revenues by making greater use of IRS information and by increasing the number of audits of non-resident filers.

Our work in this paper is related to several other studies of compliance and enforcement. The econometric specification we develop builds upon the theoretical analysis of the tax compliance game presented in Erard and Feinstein (1994a). In addition, the data employed in this study have been used in prior research by Erard and Feinstein (1994b, 1995) on federal tax reports and audit selection decisions. Our work is also related to studies of taxpayer and auditor behavior by Alm, Bahl, and Murray (1993), Beron, Tauchen and Witte (1991), and Dubin, Graetz, and Wilde (1990).

2. Modeling State and Federal Audit Interactions

In this section we present our framework for analyzing tax audit decisions and assessments. We divide our presentation into two parts. First, we present a model of auditing by a single tax authority. We then extend our model to the case in which there are two separate tax authorities, which we label federal and state. These models form the basis for our empirical analysis of Oregon and IRS tax audit decisions and assessments for tax year 1987, which is presented in section 3.

2.1 Model for a Single Tax Authority

Consider a tax authority charged with collecting the taxes owed by each member of a community of individuals or households. Although many members of the community pay their full tax liability voluntarily, others do not. The authority cannot costlessly observe each member's true liability and determine whether the member is fully in compliance. Instead, the authority must use audits to detect noncompliance. In this subsection we present a simple model of the authority's audit selection decisions and assessments. We divide our discussion into four parts. First, we specify an audit assessment distribution; next, we describe the tax authority's calculation of the expected assessment to be earned from performing an audit; then, we define audit costs and derive an audit selection criterion; finally, we present the likelihood function associated with our model.

Audit Assessments

Consider first the specification of audit revenue assessments, which is the most complex part of our model. For a taxpayer who has been audited, we define R as the amount of additional taxes, interest, and penalties that the taxpayer is assessed. For a taxpayer who has not been audited, we define R as the amount that would have been assessed if an audit had taken place. The assessment R may be positive, in which case the taxpayer owes R dollars to the government; zero; or negative, in which case the individual has overpaid his taxes.

In our data, described much more fully below, approximately 70% of all federal audits result in a positive assessment, 23% result in no assessment, and 7% result in a negative assessment. Further, the mean positive federal assessment is more than twice as large as the mean negative assessment, and the variance of positive assessments is much larger than the variance of negative assessments, in part because there are a small number of very large positive assessments. The statistics for state audit assessments are similar, though not identical. Our specification of the distribution of R reflects these facts, in two main ways. First, in contrast to several previous studies of tax compliance, including Clotfelter (1983) and Feinstein (1991), we distinguish negative assessments from zero assessments, in order to more precisely model the assessment distribution.[4] Second, we specify a log-normal distribution for positive assessments, in order to fit the long right-hand tail of very large positive assessments recorded in our data.

We specify the distribution associated with R in terms of a two-step process consisting of three equations. The first step distinguishes positive assessments from non-positive (zero or negative) assessments. We define the latent variable P^* as

$$P^* = \beta_1 x_1 + u^*, \tag{1}$$

and assume that the assessment is positive if P^* is greater than zero, but otherwise the assessment is either zero or negative. In expression (1), x_1 are characteristics of the individual or household under consideration, β_1 is a parameter vector, and u^* is a stochastic disturbance. The second step involves one of two expressions, depending on the sign of P^*. If P^* is greater than zero, the assessment R is positive and is defined by

$$R = exp\{\beta_2 x_2 + \epsilon\}, \tag{2}$$

where x_2 are characteristics of the individual or household, β_2 is a parameter vector, and ϵ is a stochastic disturbance. Alternatively, if P^* is less than or equal to zero, then the assessment R is either negative or zero, according to a tobit specification given by

$$R = \begin{cases} -\alpha + u & \text{if } u < \alpha \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

where α is a constant and u is a stochastic disturbance. We discuss the distributions associated with the stochastic disturbances w, ϵ, and u below.

It is important to recognize that our specification of the assessment distribution is neither derived from nor meant to be interpreted as a structural model of reporting behavior.[5] Instead, the model reflects our view of the way in which a tax authority is likely to evaluate the assessment distribution. We believe the authority is likely first to evaluate the probability that the taxpayer has underpaid his taxes, using equation (1) and including in x_1 individual or household characteristics that affect the probability of an underreport. We believe the authority is likely next to evaluate the magnitude of an underreport, conditional on an underreport occurring, using equation (2). In equation (2), x_2 includes individual or household characteristics that influence the extent of underreporting, and the exponential parameterization captures the long right-hand tail, reflecting the small probability of a very large positive assessment. We note that x_1 and x_2 may contain some common elements, but are unlikely to be identical. We believe the authority is likely to consider last the possibility that, conditional on no underreport, there is a negative assessment, evaluating both the probability and likely extent of such a negative assessment by means of equation (3). We doubt that most tax authorities develop a careful model of negative assessments, partly because such assessments are relatively infrequent and of small magnitude, and partly because it is not obvious what individual or household variables are likely to be associated with overpayments. Hence we do not include any explanatory variables in equation (3).

Note that the assessment R is not equivalent to the difference between the taxpayer's legal tax obligation and his tax payment, for four reasons. First, the tax examiner may not detect all of the taxpayer's underpayment or overpayment. Second, the examiner may tend to exaggerate the size of an underpayment, in an effort to obtain greater enforcement revenue for the tax agency. Third, the assessment may include interest and penalty charges for detected underpayments. Fourth, the examiner and the taxpayer may negotiate over the size of the assessment, in which case the final outcome will depend on the relative bargaining strengths of the two parties.[6]

Calculation of the Expected Assessment

As the next step in the description of our model, consider the tax authority's calculation of the expected value of the assessment associated with a particular taxpayer, a calculation that plays a central role in the authority's audit selection decision. We assume that the authority knows the form of equations (1), (2), and (3), including the forms of the distributions from which the stochastic disturbances are drawn, and knows the values of all parameters that enter into these three equations. We also assume that the authority observes the explanatory variables x_1 and x_2 prior to making its audit selection decision, but does not observe the values of the stochastic disturbances w, ϵ, and u, and therefore does not observe the actual assessment R. Lastly, and importantly, we assume that prior to making its audit selection decision the authority is able to observe the value of a signal, denoted η, that provides information about the assessment R. We integrate the signal into our revenue assessment model by assuming that η is correlated with each of w and ϵ, and therefore provides information about both the probability of a positive assessment and, conditional on a positive assessment occurring, the likely magnitude of the assessment. We do not allow for the possibility that η is correlated with u, since we doubt that the authority is likely to observe information about the likelihood or magnitude of a negative assessment. Intuitively, we expect that the greater is η, the greater will be the tax authority's calculation of the expected value of the assessment to be earned from performing an audit. Although η is observable to the tax authority, we assume that it is not recorded in the data available for analysis, and therefore must be treated as a stochastic disturbance in the econometric specification. The fact that η is a stochastic disturbance is important for the structure of both our econometric model and the associated likelihood function; we discuss the role of η in the model and the likelihood function below.

Having introduced the signal η, we can now specify distributions for w, ϵ, η, and u. We assume that w, ϵ, and η are jointly drawn from a trivariate normal distribution, and that u is independently drawn from a separate normal distribution. We impose several restrictions on these distributions. The first restriction is that the unconditional mean of each disturbance is zero. The second restriction is that the standard error of w (σ_w) is equal to one, a normalization that is required for identification for the same reason that the standard error in a probit model is set equal to one; namely, because only the sign of P^* affects the assessment R. The third restriction, similar to the second, is that the standard error of the signal η (σ_η) is equal to one, a normalization that is required for identification because, in the likelihood function, η is associated with the tax authority's decision about whether or not to conduct an audit, a binary choice that is modeled in a manner analogous to a probit model. The final restriction is that (conditional on η) u and ϵ are independent of one another: for the trivariate normal distribution, this restriction is equivalent to the condition that $\rho_{u\epsilon} = \rho_{\eta u}\rho_{\eta\epsilon}$, where ρ_{ab} is the correlation between random variables a and b. We impose this final restriction primarily to ease the computational burden associated with estimating the model.

After these restrictions have been imposed, there are four remaining parameters to be estimated. Two of the parameters are standard errors: σ_ϵ, the standard error associated with positive assessments; and σ_u, the standard error associated with negative assessments. The other two parameters are correlations: $\rho_{\eta u}$, which measures the information contained in the signal about the probability of a positive assessment; and $\rho_{\eta\epsilon}$, which measures the information contained in the signal about the likely magnitude of a positive assessment, conditional on a positive assessment occurring.[7] Note that although ϵ is normally distributed, the distribution of positive assessments is log-normal, due to the exponential form of equation (2).

We let $E(R|\eta)$ denote the expected value of the audit assessment, conditional on the value of the signal η. Using well known properties of the normal and log-normal distributions, $E(R|\eta)$ can be expressed as

$$(4) \quad \Phi\left[\frac{\beta_1 x_1 + \rho_{\eta u}\eta}{\sqrt{1-\rho_{\eta u}^2}}\right] exp\left\{\beta_2 x_2 + \rho_{\eta\epsilon}\sigma_\epsilon\eta + \frac{\sigma_\epsilon^2(1-\rho_{\eta\epsilon}^2)}{2}\right\} - \left[1-\Phi\left(\frac{\beta_1 x_1 + \rho_{\eta u}\eta}{\sqrt{1-\rho_{\eta u}^2}}\right)\right]\left[\alpha\Phi\left(\frac{\alpha}{\sigma_u}\right)+\sigma_u\phi\left(\frac{\alpha}{\sigma_u}\right)\right].$$

where $\phi(\bullet)$ and $\Phi(\bullet)$ are, respectively, the standard normal probability and cumulative density functions. The first term in equation (4) is the probability of a positive assessment multiplied by the expectation of the magnitude of the assessment, conditional on a positive assessment occurring. The second term is the probability of a nonpositive assessment multiplied by the expectation of the magnitude of a negative assessment, conditional on a nonpositive assessment occurring and taking into account the probability of a zero assessment.

Audit Selection Criterion

Consider now the tax authority's audit selection criterion. Our specification of this criterion is based upon the theoretical analysis of the tax compliance game presented in Erard and Feinstein (1994a). They show that a revenue-maximizing tax authority which has a fixed audit budget, is risk-neutral, and cannot precommit to its audit rule will select a return for audit whenever the expected revenue to be earned from performing the audit exceeds λc, where c is the audit cost and λ is a Lagrange multiplier associated with the budget constraint. The multiplier λ is an important policy parameter because it provides a measure of the increase in tax revenues that can be achieved by increasing the tax authority's audit budget. In particular, if λ exceeds one the government can increase its total revenue by raising the tax authority's audit budget. Therefore, a revenue-maximizing government would want to provide the tax agency with sufficient resources to make λ equal to one. This revenue-maximizing policy might not, however, be an optimal policy from a social welfare perspective. The revenue raised from additional audit resources is merely a transfer from noncompliant taxpayers to the government, whereas the audit resources employed to effect this transfer represent a genuine resource cost. The welfare-maximizing value of λ, therefore, may be well in excess of one.

Adapting the criterion of Erard and Feinstein (1994a) to our context, we conclude that the tax authority will choose to audit a taxpayer whenever the expected revenue conditional on the observed signal, $E(R|\eta)$, is equal to or greater than the shadow cost of an audit, λc. In our econometric analysis, we treat λc as a single parameter; however, we have separate information about audit costs, so we are able to deduce an estimate of λ by dividing our estimate of λc by a rough estimate of c.

We assume that the expected value of the revenue assessment is nondecreasing in η, which implies that there exists some threshold signal, η^*, such that for all $\eta > \eta^*$, $E(R|\eta) \geq \lambda c$, while for $\eta < \eta^*$, $E(R|\eta) \leq \lambda c$.[8] It then follows that the authority will choose to audit a taxpayer if and only if the signal η is equal to or larger than the threshold value

η^*. The threshold value for the signal is determined implicitly by the equation

$$E(R|\eta^*) = \lambda c. \tag{5}$$

Likelihood Function

As the final step in the presentation of our model of a single tax authority. we present the likelihood function associated with the model. Each observation refers to a particular individual or household and falls into one of four categories: no audit; audit and positive assessment: audit and zero assessment: audit and negative assessment. The likelihood associated with no audit is simply

$$L_1 = \Phi(\eta^*). \tag{6}$$

or the probability that the signal is below the threshold value η^*. Note that η^* is a function of the characteristics x_1 and x_2. and therefore varies across individuals and households. The likelihood associated with an audit and a positive assessment in the amount R is

(7)
$$L_2 = \frac{1}{R\sigma_\epsilon}\varphi\left(\frac{lnR - \beta_2 x_2}{\sigma_\epsilon}\right) BN\left[-\frac{\eta^* - \rho_{\eta\epsilon}\left(\frac{lnR - \beta_2 x_2}{\sigma_\epsilon}\right)}{\sqrt{1-\rho_{\eta\epsilon}^2}}, \frac{\beta_1 x_1 + \rho_{\eta\epsilon}\rho_{\eta w}\left(\frac{lnR - \beta_2 x_2}{\sigma_\epsilon}\right)}{\sqrt{1-\rho_{\eta w}^2\rho_{\eta\epsilon}^2}}, \frac{\rho_{\eta w}\sqrt{1-\rho_{\eta\epsilon}^2}}{\sqrt{1-\rho_{\eta u}^2\rho_{\eta\epsilon}^2}}\right].$$

where R is the audit assessment and $BN[\bullet, \bullet, \rho]$ represents the standard bivariate normal cumulative distribution function with correlation ρ. The likelihood associated with an audit and a zero assessment is

$$L_3 = BN\left[-\eta^*, -\beta_1 x_1, -\rho_{\eta u}\right] \Phi\left(-\frac{\alpha}{\sigma_u}\right). \tag{8}$$

Finally, the likelihood associated with an audit and a negative assessment is

$$L_4 = BN\left[-\eta^*, -\beta_1 x_1, -\rho_{\eta u}\right] \frac{1}{\sigma_u} \phi\left(\frac{R+\alpha}{\sigma_u}\right). \tag{9}$$

where again R is the assessment, which in this case is negative.

2.2 State and Federal Model

In the United States and many other countries most individuals and households are obligated to pay taxes to more than one political jurisdiction. When this is the case and the tax authority associated with each jurisdiction conducts tax audits, many questions arise concerning issues that are important both for the understanding of tax enforcement systems and for tax policy formulation. To what extent do the authorities employ similar selection criteria? To what extent do they coordinate their selection processes? To what extent do they share audit results and other information? When they both audit the same individual or household, do they detect the same noncompliant behavior? Finally, is the marginal value of an additional dollar of audit resources approximately the same for the different authorities?

In this subsection we present a model that addresses these and related questions. The model includes a large collection of individuals (or households), each of whom is obligated to pay taxes to two jurisdictions, labeled federal and state. We assume that each jurisdiction has a tax authority, and each authority conducts audits. In addition, we assume that the state tax authority can use results from federal audits to perform piggyback audits. We divide our presentation into two parts. We first describe the conceptual structure of the model, which is based on the framework presented in the previous subsection. We then derive the likelihood function associated with the model.

Model Structure

Consider a particular individual or household. Our model of the state and federal tax authorities' decisions about whether or not to audit the individual or household consists of two periods. In the first period each authority decides whether or not to audit the individual or household, and, if it conducts an audit, makes a revenue assessment. The authorities' period one decisions are made simultaneously and are independent of one another. If in period one either the state authority has conducted an audit or the federal authority has not conducted an audit, then period two is not applicable and the model terminates at the end of period one. Otherwise, there is a second period, during which the state authority learns the federal period one audit results and then decides whether or not to perform a piggyback audit. The piggyback audit consists of two stages: first, the federal audit results and the taxpayer's state return are used to determine the additional revenue owed to the state; second, the individual is notified of the assessment. The piggyback audit is much less expensive than a period one audit, since it involves neither direct face-to-face contact with the individual nor a careful investigation of the individual's tax records.[9]

We use the framework presented in subsection 2.1 to model both the federal audit assessment distribution and the federal tax authority's period one audit selection decision. In particular, we let R_f denote the federal assessment that either is generated by a federal audit or would have been generated if the individual or household had been subjected to a federal audit. Similarly, we let x_{1f} and x_{2f} denote characteristics of the individual or household that affect the federal assessment and enter into equations (1) and (2); w_f, ϵ_f and u_f denote the stochastic disturbances that are associated with the federal revenue assessment distribution and enter into equations (1), (2), and (3); and η_f denote the signal observed by the federal tax authority. Finally, we define the threshold value for the federal signal η_f^* as the value of η_f for which $E(R_f|\eta_f^*) = \lambda_f c_f$, where λ_f is the multiplier associated with the federal audit budget constraint and c_f is the cost of a federal audit. We assume that $E(R_f|\eta_f)$ is greater than or equal to $\lambda_f c_f$ if and only if $\eta_f \geq \eta_f^*$. It then follows that the federal tax authority performs an audit whenever η_f is equal to or greater than η_f^*. As should be clear from our description, the federal tax authority observes only its own signal η_f prior to making its audit decision; it does not observe the analogous state signal η_s, which is introduced below.

Our specification of the state period one independent audit assessment distribution also is based on the framework presented in the subsection 2.1; however our specifications of the state piggyback assessment distribution and the state tax authority's audit selection procedure are somewhat different. We let R_s denote the state period one assessment: x_{1s} and x_{2s} denote characteristics of the individual or household that affect the period one assessment and enter into the state versions of equations (1) and (2); w_s, ϵ_s and u_s denote stochastic disturbances that enter into the state versions of equations (1), (2), and (3); and η_s denote the signal observed by the state tax authority. We define the cost of a state independent audit to be $\lambda_s c_s$, where λ_s is the multiplier associated with the state's independent audit budget constraint and c_s is the cost of an independent state audit. We expect c_s to be smaller than c_f, since state audits are normally shorter and simpler than federal audits. If the federal and state governments were able to share budgets and revenues, we might expect λ_s to be approximately equal to λ_f, since, if one λ value were larger than the other, audit resources could be transferred to the authority with the larger λ value, increasing total government revenues. Since governments do not share budgets and revenues to this extent, however, we expect λ_s to be somewhat smaller than λ_f, because at least in most cases the state tax rate is substantially below the federal rate.

We let R_p denote the assessment that either is generated by a period two piggyback audit or would have been generated if the state had chosen to conduct such an audit. Since the piggyback audit is based directly on the period one federal assessment, R_p is likely to depend upon R_f. However, R_p may not be exactly proportional to R_f, due to differences in tax progressivity, differences in the tax treatment of certain issues between the federal government and the state, differences in penalty and interest charges, and possible administrative errors. We define

$$ln(R_p + K) = h(R_f, x_p; \beta_p) + \epsilon_p. \tag{10}$$

In equation (10), $h(\bullet)$ is a parametric function that depends on the federal revenue assessment R_f and a vector of explanatory variables x_p, which control for differences between the federal and state tax bases, tax rate schedules, and credit structures. We assume that the state tax authority knows the functional form of h, inclusive of the values of the parameter vector β_p and the "displacement parameter" K, which accounts for the possibility that R_p is negative. The term ϵ_p is a stochastic disturbance. The state does not have direct knowledge of ϵ_p: rather it observes a signal η_p of its likely value at the beginning of period two, prior to deciding whether to perform a piggyback audit. We assume that a piggyback audit has a shadow cost of $\lambda_p c_p$. We expect c_p to be substantially below both c_s and c_f. If the state tax authority is allocating its audit resources efficiently, λ_s should be approximately equal to λ_p, since, if one λ value were much larger than the other, the state could increase its revenues by shifting audit resources from the audit program associated with the smaller λ value to the audit program associated with the larger λ value.[10]

We consider mainly a *nonstrategic* model of the state tax authority's audit selection process. This model is based on the assumption that the state authority performs a period one audit whenever the expected revenue assessment exceeds the audit cost, without taking into consideration the possibility that it may be able to perform a piggyback audit in period two if the federal tax authority performs a period one audit and the state authority does not. In contrast, a *strategic* model would assume that the state does take into account the potential for a period two piggyback audit when making its period one audit decision. We believe the nonstrategic model is descriptive of actual state audit selection decisions, but that states might be able to increase their audit revenues by adopting the audit selection rule generated by the strategic model. In the remainder of this section we describe the nonstrategic model in more detail and derive the likelihood function associated with it. In section 3 we present results from estimation of the nonstrategic model, and in the appendix we provide a more detailed description of the strategic model.

We define the threshold value for the state signal η_s^* as the value of η_s for which $E(R_s|\eta_s^*) = \lambda_s c_s$, and we assume that $E(R_s|\eta_s)$ is greater than or equal to $\lambda_s c_s$ if and only if $\eta_s \geq \eta_s^*$. For the nonstrategic model, it then follows that the state chooses to conduct a period one audit if and only if the signal η_s is equal to or greater than η_s^*. Note that the state observes only the signal η_s in period one, and has no knowledge of either the federal signal η_f or the federal authority's decision about whether or not to audit the individual. If the state authority does not perform a period one audit but the federal tax authority does perform such an audit, then in the second period the state authority must decide whether or not to perform a piggyback audit. Define η_p^* to be the value of the signal η_p for which $E(R_p|\eta_p^*, R_f) = \lambda_p c_p$, and assume that for all values of R_f, $E(R_p|\eta_p, R_f)$ is equal to or greater than $\lambda_p c_p$ if and only if $\eta_p \geq \eta_p^*$. It then follows that the state authority will choose to perform a piggyback audit if and only if the signal η_p is equal to or greater than η_p^*. Note that, for a given value of the signal η_p, $E(R_p|\eta_p, R_f)$ is nondecreasing in R_f, so the threshold value η_p^* is a nonincreasing function of R_f, and the probability of a piggyback audit is a nondecreasing function of the federal assessment R_f.

As the final step in the description of our model of state and federal tax audit decisions, consider the distribution of the stochastic disturbances in the model. We assume that $(w_f, w_s, \epsilon_f, \epsilon_s, \eta_f, \eta_s)$ are drawn from a multivariate normal distribution, and we impose the following conditions on this distribution. Following the specification presented in the subsection 2.1, we normalize each of σ_{η_f}, σ_{η_s}, σ_{w_s}, and σ_{w_f} to one. We then make the following assumptions: (conditional on η_f) ϵ_f and w_f are independent; (conditional on η_s) ϵ_s and w_s are independent; (conditional on η_s) w_s is independent of both η_f and ϵ_f; and (conditional on η_f) w_f is independent of both η_s and ϵ_s. The latter two assumptions are made primarily to ease the computational burden associated with estimating the model. The distribution then depends upon a total of eleven parameters. Six of the parameters are familiar from the model of the previous subsection: the two standard deviations, σ_{ϵ_f} and σ_{ϵ_s}; and the four correlations, $\rho_{\eta_f w_f}$, $\rho_{\eta_s w_s}$, $\rho_{\eta_f \epsilon_f}$, and $\rho_{\eta_s \epsilon_s}$. The remaining five parameters are new, and they represent features of the relationship between the federal and state audit assessment processes. These five parameters are: (1) $\rho_{\eta_f \eta_s}$, the correlation between the two signals; (2) $\rho_{\epsilon_f \epsilon_s}$, the correlation between the two ϵ disturbances; (3) $\rho_{w_f w_s}$, the correlation between the two w disturbances; (4) $\rho_{\eta_f \epsilon_s}$, the correlation between the federal signal η_f and the disturbance ϵ_s in the state revenue model; and (5) $\rho_{\eta_s \epsilon_f}$, the correlation between the state signal η_s and the disturbance ϵ_f in the federal revenue model. We discuss the interpretation of these parameters below.[11]

We assume that the disturbances u_f and u_s are drawn from a bivariate normal distribution with parameters σ_{u_f}, σ_{u_s}, and $\rho_{u_f u_s}$, and that each of u_f and u_s is independent of the other stochastic disturbances in the model. We also assume that η_p and ϵ_p are drawn from a bivariate normal distribution, and we impose as a normalization that the standard deviation of η_p is one, leaving two free parameters in this distribution, σ_{ϵ_p} and $\rho_{\eta_p \epsilon_p}$. In addition, we assume that each of η_p and ϵ_p is independent of the other stochastic disturbances in the model. Finally, we note from equation (10) that the assumption that ϵ_p

is normal implies that R_p is distributed according to the displaced log-normal distribution. with displacement parameter K.

Our model provides a rich structure for analyzing the relationship between federal and state tax audit decisions and assessments. A comparison of the variables included in x_{1f} and x_{2f} with those included in x_{1s} and x_{2s}. and of the parameter vectors β_{1f} and β_{2f} with the vectors β_{1s} and β_{2s}. can provide information about the extent to which the variables that influence federal audit decisions and assessments are similar to those that influence state decisions and assessments. If the two sets of variables and parameters turn out to be significantly different. further research will be necessary to determine whether the differences are due to differences in tax law. differences in reporting behavior. or differences in audit selection and assessment procedures. If there are substantial differences resulting from different audit and assessment procedures. our models may help administrators to improve their procedures. The parameter $\rho_{\eta_f \eta_s}$ measures the correlation between the signals observed by the federal and state authorities. If this correlation is positive. it will serve as an indication that the two authorities have access to similar sources of information. and draw upon similar experiences. in evaluating assessment distributions. Conversely. if the correlation is negative. we might infer that the two authorities tend to draw upon different sources of information and experiences in formulating their audit policies. In either case. so long as the two signals are not perfectly correlated. the potential will exist for both agencies to improve their audit policies by sharing the information contained in their signals. Comparisons of $\rho_{\eta_f u_f}$ to $\rho_{\eta_s u_s}$ and of $\rho_{\eta_f \epsilon_f}$ to $\rho_{\eta_s \epsilon_s}$ will indicate whether or not the signal observed by the federal authority is more informative than the signal observed by the state authority. a comparison which provides an interesting measure of the relative capabilities of the two authorities. The cross correlations $\rho_{\eta_f \epsilon_s}$ and $\rho_{\eta_s \epsilon_f}$ measure the degree to which the signal observed by one authority provides information about the likely magnitude of a positive assessment for the other authority. conditional on a positive assessment occurring. To understand the empirical phenomena to which these cross correlations relate. consider $\rho_{\eta_f \epsilon_s}$. If this correlation is larger (more positive) than $\rho_{\eta_f \eta_s} \rho_{\eta_s \epsilon_s}$. then conditional on η_s the likely magnitude of a positive state assessment is larger, the larger is η_f. Since the federal authority is more likely to audit the larger is η_f, a large positive value for $\rho_{\eta_f \epsilon_s}$ implies that. conditional on the state making a positive assessment. the magnitude of the assessment is expected to be larger when the federal authority has chosen to perform an audit than when the federal authority has chosen not to perform an audit. More importantly from the viewpoint of policy, these cross correlations provide an indication of the value of information sharing between the two tax authorities. For example, so long as the value of $\rho_{\eta_f \epsilon_s}$ is different from the value of $\rho_{\eta_f \eta_s} \rho_{\eta_s \epsilon_s}$, the federal government will be able to help the state improve its audit selection procedures by sharing the information contained in the federal signal.[12] Finally, $\rho_{w_f w_s}$, $\rho_{\epsilon_f \epsilon_s}$, and $\rho_{u_f u_s}$ measure the extent to which federal and state assessments are correlated. We expect each of these correlations to be positive, but we are particularly interested in their magnitudes, especially that of $\rho_{\epsilon_f \epsilon_s}$. Of course, we do not expect federal and state assessments to be identical, or even proportional, due to differences in tax law and in the specific items of noncompliance detected during the respective audits.[13] However, a finding of large differences between state and federal assessments may serve as an indication that revenues can be increased by pooling audit results for cases subjected to independent audit by both agencies.

Likelihood Function

The likelihood function associated with our model is rather complex. There are four qualitatively distinct outcomes possible for the federal tax authority in the model: (1) no audit; (2) audit, positive assessment; (3) audit, no assessment; and (4) audit, negative assessment. There are five qualitatively distinct outcomes possible for the state authority: (1) no audit; (2) independent audit, positive assessment; (3) independent audit, no assessment; (4) independent audit, negative assessment; and (5) no independent audit, piggyback audit. Since the form of the likelihood function depends on the outcomes for both authorities, there are therefore twenty potential cases to be considered. However, because the state authority can perform a piggyback audit only when the federal authority has performed an audit in period one, only nineteen of these cases are relevant. We do not present the likelihoods associated with all nineteen cases. Rather, we present the likelihoods for six representative cases, and leave the remaining cases to be worked out by the reader, if interested.

Case 1: federal audit and positive assessment, state audit and positive assessment.

$$
(11) \quad \begin{aligned} L_1 = & \frac{1}{R_f R_s \sigma_{\epsilon_f} \sigma_{\epsilon_s}} BN\left(\frac{lnR_f - \beta_{2f}x_{2f}}{\sigma_{\epsilon_f}}, \frac{lnR_s - \beta_{2s}x_{2s}}{\sigma_{\epsilon_s}}, \rho_{\epsilon_f\epsilon_s}\right) \\ & \int_{\eta_f^*}^{\infty}\int_{\eta_s^*}^{\infty}\int_{-\beta_{1f}x_{1f}}^{\infty}\int_{-\beta_{1s}x_{1s}}^{\infty} f(\eta_f, \eta_s, w_f, w_s | R_f, R_s)\, dw_s\, dw_f\, d\eta_s\, d\eta_f, \end{aligned}
$$

where $BN(\bullet, \bullet, \rho)$ represents the standard bivariate normal distribution function with correlation coefficient ρ and $f(\bullet)$ is a multivariate conditional normal density function, derivable from the distribution functions defined above for the stochastic disturbances.

Case 2: federal audit and positive assessment, state audit and zero assessment.

$$
(12) \quad \begin{aligned} L_2 = & \frac{1}{R_f \sigma_{\epsilon_f}} \Phi\left(\frac{-\alpha_s}{\sigma_{u_s}}\right) \phi\left(\frac{lnR_f - \beta_{2f}x_{2f}}{\sigma_{\epsilon_f}}\right) \\ & \int_{\eta_f^*}^{\infty}\int_{\eta_s^*}^{\infty}\int_{-\beta_{1f}x_{1f}}^{\infty}\int_{-\infty}^{-\beta_{1s}x_{1s}} f(\eta_f, \eta_s, w_f, w_s | R_f)\, dw_s\, dw_f\, d\eta_s\, d\eta_f. \end{aligned}
$$

where $\phi(\bullet)$ and $\Phi(\bullet)$ are, respectively, the standard normal probability and cumulative density functions.

Case 3: federal audit and positive assessment, state audit and negative assessment.

$$
(13) \quad \begin{aligned} L_3 = & \frac{1}{R_f \sigma_{\epsilon_f} \sigma_{u_s}} \phi\left(\frac{R_s + \alpha_s}{\sigma_{u_s}}\right) \phi\left(\frac{lnR_f - \beta_{2f}x_{2f}}{\sigma_{\epsilon_f}}\right) \\ & \int_{\eta_f^*}^{\infty}\int_{\eta_s^*}^{\infty}\int_{-\beta_{1f}x_{1f}}^{\infty}\int_{-\infty}^{-\beta_{1s}x_{1s}} f(\eta_f, \eta_s, w_f, w_s | R_f)\, dw_s\, dw_f\, d\eta_s\, d\eta_f. \end{aligned}
$$

Case 4: federal audit and positive assessment, no state independent audit, state piggyback.

(14)

$$L_4 = \frac{1}{R_f \sigma_{\epsilon_f} (R_p + K)\sigma_{\epsilon_p}} \phi\left(\frac{\ln(R_p+K) - h\left(R_f, x_p, \beta_p\right)}{\sigma_{\epsilon_p}}\right) \phi\left(\frac{\ln R_f - \beta_{2f} x_{2f}}{\sigma_{\epsilon_f}}\right)$$

$$\Phi\left(\frac{-\eta_p^* + \rho_{\eta_p \epsilon_p}\left(\frac{\ln(R_p+K) - h\left(R_f, x_p, \beta_p\right)}{\sigma_{\epsilon_p}}\right)}{\sqrt{1-\rho^2_{\eta_p \epsilon_p}}}\right) \int_{\eta_f^*}^{\infty} \int_{-\infty}^{\eta_s^*} \int_{-\beta_{1f} x_{1f}}^{\infty} f(\eta_f, \eta_s, w_f | R_f)\, dw_f\, d\eta_s\, d\eta_f .$$

Case 5: federal audit and positive assessment, no independent state audit, no state piggyback.

(15)

$$L_5 = \frac{1}{R_f \sigma_{\epsilon_f}} \phi\left(\frac{\ln R_f - \beta_{2f} x_{2f}}{\sigma_{\epsilon_f}}\right) \Phi(\eta_p^*) \int_{\eta_f^*}^{\infty} \int_{-\infty}^{\eta_s^*} \int_{-\beta_{1f} x_{1f}}^{\infty} f(\eta_f, \eta_s, w_f | R_f)\, dw_f\, d\eta_s\, d\eta_f .$$

Case 6: no federal audit, no state audit.

(16)
$$L_6 = BN\left(\eta_f^*, \eta_s^*, \rho_{\eta_f \eta_s}\right).$$

The parameters of the model can be estimated jointly by maximizing the full likelihood function. However, we have chosen to estimate the model using a much simpler estimation strategy, which is based on the observation that the joint model we have outlined in this subsection nests the single tax authority model presented in subsection 2.1. In particular, we estimate the parameters of the model in four steps. First, we estimate the federal audit selection and assessment parameters corresponding to equations (1), (2), and (3) from the previous subsection; this estimation yields consistent estimates of β_{1f}, β_{2f}, σ_{ϵ_f}, $\rho_{\eta_f w_f}$, $\rho_{\eta_f \epsilon_f}$, σ_{u_f}, and $\lambda_f c_f$. Next, we estimate the state audit selection and assessment parameters corresponding to the state version of equations (1), (2), and (3); this estimation yields consistent estimates of β_{1s}, β_{2s}, σ_{ϵ_s}, $\rho_{\eta_s w_s}$, $\rho_{\eta_s \epsilon_s}$, σ_{u_s}, and $\lambda_s c_s$. Note that each of these first two steps involves estimating the model presented in subsection 2.1. Third, we estimate the piggyback audit selection and assessment equation (10) on the subset of cases for which the federal government performed a period one audit and the state did not; this estimation yields consistent estimates of β_p, $\rho_{\eta_p \epsilon_p}$, and σ_{ϵ_p}. Finally, we estimate the full likelihood function, maximizing over the six remaining parameters, $\rho_{\eta_f \eta_s}$, $\rho_{\eta_f \epsilon_s}$, $\rho_{\eta_s \epsilon_f}$, $\rho_{w_f w_s}$, $\rho_{u_f u_s}$, and $\rho_{\epsilon_f \epsilon_s}$. As discussed in section 3, we perform our estimation using a choice-based data sample. We therefore make an adjustment to the estimated standard errors associated with all parameter estimates to account for this feature of the data. The two critical assumptions that are necessary for our estimation procedure to yield consistent estimates are that $\rho_{\eta_f w_s} = \rho_{\eta_f \eta_s} \rho_{\eta_s w_s}$ and $\rho_{\eta_s w_f} = \rho_{\eta_s \eta_f} \rho_{\eta_f w_f}$. If, for example, the first of these equalities failed to hold, then $E(R_s|\eta_s, \eta_f)$ would no longer be equal to $E(R_s|\eta_s)$ and we would not be able to estimate the state audit selection and assessment process separately from the federal selection and assessment process.[14]

3. Empirical Results

In this section we describe the results from our empirical analysis of Oregon Department of Revenue (ODR) and Internal Revenue Service (IRS) tax audits for tax year 1987. We divide our discussion into three parts. First, we describe the data we used in our study. Second, we present and discuss results from the estimation of the model of state and federal tax audit decisions and assessments presented in the previous section. Third, we discuss a number of other interesting findings that have emerged from a detailed investigation of Oregon's individual income tax audit programs.

3.1 The Data

Our analysis is based on tax return and audit information that has been compiled from several different sources. Our information about federal and state 1987 tax returns comes from two sources, the ODR Personal Income Tax Return Extract File database and the IRS Individual Returns Transaction File (IRTF) database; research staff at ODR matched the taxpayer records from these two databases for us.[15] Our information about audits of federal and state 1987 tax returns also comes from two sources, the ODR Audit Casedata File database. which contains information about ODR audits, and the IRS Audit Information Management System database, which contains information about IRS audits. We obtained the IRS audit information through a match based on social security numbers that was performed by the IRS research staff. We merged the tax return data with the audit data to create the dataset used in our analysis.

Our dataset is a subset of the total population of Oregon filers for tax year 1987. In order to ensure that a large number of audit cases would be included in our sample. we heavily oversampled filers that were subjected to an audit. In addition. we sampled business and farm returns at a higher rate than other returns. Our sampling procedure was as follows. For returns that were subjected to any form of enforcement action by the ODR or the IRS. we selected every return that reported any business (federal schedule C) or farm (federal schedule F) income (positive or negative) and one-half of all those returns that reported neither business nor farm income. For returns that were not subjected to any enforcement action at either the state or federal level. we selected 6% of returns that reported business income. 25% of returns that reported farm income but no business income. and 1.75% of returns that reported neither business nor farm income.

For our econometric analysis. we have excluded returns for part-year residents and non-residents. and returns that failed to match with the IRS IRTF database, leaving 43.587 observations.[16] This total includes including approximately 4.500 returns that were selected for a federal audit. 1.700 returns that were selected for an independent state audit. and 1.200 returns that were selected for a state piggyback audit. Our data include extremely detailed tax return information. In particular. for each observation we have nearly every line-item of the federal 1040. 1040A. or 1040EZ form as well as federal Schedule A; selected line-item information from federal schedules C. D. E. and F; and nearly every line-item from the Oregon return, which may be either a Form 40F (full-year long-form) or a Form 40S (full-year short-form). The data identify whether the return was selected for a federal audit, and if the return was selected for a federal audit, the data indicate the auditor's assessment of the additional tax. interest. and penalties owed by the taxpayer at the time the case closed. Similarly, the data identify whether the return was selected for an Oregon audit, and if the return was selected for an Oregon audit, the data identify whether the audit was an independent or a piggyback audit and indicate the final assessment made at the time the case closed.

Tables 2 through 4 provide additional information about the data sample from which we selected returns for our econometric analysis. The frequency figures in table 2 are unweighted and provide information about numbers of observations in our sample, while dollar amounts are weighted to reflect population statistics. The dollar figures presented in each of the tables represent the additional tax, interest, and penalties that were assessed during the relevant audits. Table 2 indicates that there were slightly more than one million filers in Oregon in 1987, 63,000 of whom were placed in one of the IRS business audit classes, 7,300 of whom were placed in one of the IRS farm audit classes, and 941,000 of whom were placed in one of the IRS nonbusiness, nonfarm classes.[17] Our sample includes 6,492 returns from the business audit classes, 1,945 returns from the farm audit classes, and 35,150 returns from the nonbusiness, nonfarm classes. In total, our sample contains 4,433 IRS audit cases, of which 1,073 were audits of returns falling in a business class (business audits), 148 were audits of returns falling in a farm class (farm audits), and 3,212 were audits of returns from a nonbusiness, nonfarm class (nonbusiness, nonfarm audits). Slightly less than one-quarter of all audits resulted in no additional assessment; however, approximately 40% of all farm audits resulted in no additional assessment. Approximately 6% of all audits in our sample resulted in a negative additional assessment; the mean negative assessment in the population was $1,330, while the median negative assessment was $283. The percentage frequencies of negative assessments in the various categories were quite similar to the overall frequency, although the mean and median levels of these assessments for the business category ($3,287 and $698, respectively) were much larger than the overall average. In each audit category the majority of federal audits resulted in a positive assessment. In particular, 71% of all audits in our sample, 73% of all business audits, 52% of all farm audits, and 72% of all nonbusiness, nonfarm audits resulted in a positive assessment. The mean positive assessment in the population was $3,073 for all audits combined, $5,502 for business audits, $4,310 for farm audits, and $2,472 for nonbusiness, nonfarm audits. The median positive assessment was $1,021 for all audits combined, $1,444 for business audits, $1,240 for farm audits, and $924 for nonbusiness, nonfarm audits.

The ODR does not classify taxpayers into different audit classes. Therefore in table 2, all subsequent tables, and our econometric analysis we have placed each Oregon return in the same class as the matching federal return. Our sample includes 1,667 independent Oregon audits, of which 802 are audits of filers whose federal returns were placed in an IRS business audit class. 77 are audits of filers whose federal returns were placed in an IRS farm class, and 788 are audits of filers in an IRS nonbusiness, nonfarm class.[18] Approximately one-third of all Oregon audits in our sample resulted in no additional assessment, as compared to one-quarter of all federal audits. The largest Oregon no-change rates were in the business (39%) and farm (48%) categories. Approximately 5% of all Oregon audits in our sample resulted in a negative additional assessment; the mean negative assessment in the population was $578, while the median negative assessment was $225. The frequency of negative state assessments and the average size of those assessments was somewhat smaller in the business and farm categories. The majority of all Oregon audits resulted in a positive additional assessment, just as for federal audits. Over 61% of all Oregon audits in our sample resulted in a positive assessment: the mean positive assessment in the population was $1,051, while the median assessment was $399. These figures are well below the corresponding figures for federal audits, which is not surprising because Oregon has essentially a flat tax rate of 9%, well below the federal rate for most income categories. Approximately 57% of all audits of business returns, 49% of all audits of farm returns, and 66% of all audits of nonbusiness, nonfarm returns in our sample resulted in a positive assessment: the mean positive assessments for these three groups in the population were $1,060 and $1,317, and $1,032 respectively, while the median assessments were $442, $582, and $328. Our sample also includes 1,158 Oregon piggyback audits, including 280 audits of business returns, 23 audits of farm returns, and 855 audits of nonbusiness, nonfarm returns.[19] The mean assessment for all piggyback assessments in the population was $762, which is somewhat higher than the mean assessment for all independent Oregon audits, while the median piggyback assessment was $324, again higher than the corresponding figure for independent audits. The mean piggyback assessment was $1,005 for business returns, $414 for farm returns, and $710 for nonbusiness, nonfarm returns, while the median assessments were $374, $172, and $319, respectively.

The focus of our analysis is on the relationship between ODR and IRS audits. To explore this relationship we have partitioned the returns in our sample into five groups: (1) returns selected for both a federal audit and an independent state audit; (2) returns selected for both a federal audit and a state piggyback audit; (3) returns selected for a federal audit but not selected for a state audit; (4) returns selected for an independent state audit but not selected for a federal audit; and (5) returns not selected for audit by either tax authority. For each of these groups, table 3 presents the population frequencies as well as the relevant mean and median audit assessments. The top two rows of the table classify returns according to whether they were or were not selected for an IRS audit, and the three lefthand columns classify returns according to whether they were not selected for any kind of Oregon audit, were selected for an independent Oregon audit, or were selected for a piggyback audit. Note that the upper right-hand corner entry of the interior of the table is empty, because it is logically impossible for a return to fall into the no-federal audit, state piggyback audit category. For each category the table lists the population frequency and percentage frequency, and where appropriate, the mean and median assessments associated with each kind of audit performed. For convenience the table also displays the marginal totals for each row and column category.

The figures in table 3 indicate that there were over one million resident filers in Oregon in 1987, of whom fewer than 1% were selected for any kind of audit. The IRS selected approximately 6,000 returns for audit; the mean federal assessment was $2,146, while the median assessment was $596.[20] The ODR selected slightly less than 2,000 returns for an independent state audit; the mean assessment for these audits was $626, while the median assessment was $79. In addition, the ODR performed slightly less than 1,500 piggyback audits; the mean assessment for piggyback audits was $762 and the median assessment was $324.

The most interesting figures in table 3 are the individual cell totals. Consider first the cell including returns selected for both an IRS and an independent ODR audit. A population total of 328 returns fall into this cell. This is a far larger number than would be predicted to fall into the cell if the IRS and the ODR had selected returns for audit at random; apparently the criteria used by the different agencies to select 1987 tax returns for audit were similar. The audit assessments associated with returns in this cell were quite high, higher than the assessments associated with returns in any other cell. In particular, the federal mean assessment for returns in this cell was $5,746, well above the overall federal mean of $2,146, while the federal median assessment for returns in this cell was $1,924, well above the overall median of $596. Similarly, the ODR mean (median) assessment for returns in this cell was $1,419 ($644), well above the mean (median) for all independent ODR audits of $626 ($79).

We believe these findings make sense and are quite consistent with our model of state and federal tax audit decisions and assessments. For a return to fall into this cell both the signal observed by IRS and the signal observed by the ODR must have exceeded their threshold values. Although the signals may be positively correlated, they are not identical, and each is likely to contain information about both the federal and the state assessment distributions, as is true in our model whenever w_s and w_f are positively correlated, ϵ_s and ϵ_f are positively correlated, or the cross correlations $\rho_{\eta_f \epsilon_s}$ and $\rho_{\eta_s \epsilon_f}$ are large and positive. As a result, the expected federal assessment is likely to have been larger when both signals exceeded their respective threshold values than when only the federal signal exceeded its threshold value, and the same is true for the state assessment. Although the mean assessments associated with returns in this cell were substantially larger than the overall means, the differential between the median assessments in this cell and the overall medians was even more dramatic, a finding that we interpret as follows. Whenever both authorities independently decide to conduct an audit, the probability that the assessment for that case will be negative or zero is substantially lower than if only one of the authorities prefers to audit. However, the probability of a very large "outlier" assessment is only slightly increased when both agencies decide to audit.[21]

Consider next the cell involving returns for which the ODR chose to conduct an independent audit, but the IRS chose not to conduct an audit. The mean and median assessments associated with ODR audits of returns in this cell were \$466 and \$32 respectively, well below the mean and median assessments associated with any other kind of ODR audits, including piggyback audits. These results are also consistent with our model.

Now consider the remaining two cells that involve returns selected for an audit, the cell including returns selected for an IRS audit but not selected for an ODR audit and the cell including returns selected for an IRS audit and an ODR piggyback audit. There are 4,175 cases where the IRS performed an audit and the ODR performed no audit of any kind, and there are 1,488 cases where the IRS performed an audit and the ODR performed a piggyback audit. Thus, the ODR chose not to conduct a piggyback audit in the majority of cases for which the IRS performed an audit. The mean federal assessment for IRS audit cases that were not followed up with an ODR piggyback audit was \$1,572, while the mean federal assessment for IRS audits that were followed up with a piggyback audit was \$2,965. The median figures for these two cases are \$294 and \$960, respectively. Although these figures indicate that the ODR did follow up on many of the most profitable federal audit cases, it nonetheless chose not to conduct a piggyback audit in many cases (several thousand) for which there was a sizeable federal assessment. We are not certain why the ODR chose not to conduct more piggybacks audits in 1987. However, we provide some additional statistics and discuss this issue further later in the paper, in subsection 3.3.

Tables 4A, 4B, and 4C duplicate the format of table 3 for each of the three main categories of returns in our sample: business returns, farm returns, and nonbusiness, nonfarm returns. Most of the qualitative features of these tables are similar to the features of table 3, so we do not discuss them in detail. However, we note that while the ODR performed many more independent audits than piggyback audits in both the business and the farm classes, it actually conducted more piggyback audits than independent audits in the nonbusiness, nonfarm classes. Apparently the ODR allocates most of its audit resources to the business and farm classes. We also note that the mean assessments on independent state audits were fairly similar across audit categories. In contrast, the mean assessment on federal audits was much larger for the business audit category than for either the farm or the nonbusiness, nonfarm category.

3.2 Model Estimation Results

In this subsection we present and discuss results from the estimation of our model of state and federal audit selection decisions and assessments. We present results for three different audit categories. The first category includes all returns in our sample that fall into the middle IRS business audit class; this IRS class includes all business returns with reported Schedule C total gross receipts between \$25,000 and \$100,000.[22] The second category contains all returns that fall into either of two IRS farm audit classes; we have pooled the returns for these two classes together in order to obtain an adequate number of degrees of freedom for estimation. The final category includes all observations for which the IRS placed the federal return in the middle IRS nonbusiness, nonfarm class; this IRS class includes all nonbusiness, nonfarm returns for which the calculated total positive income is between \$25,000 and \$50,000.[23] For each category we first present and discuss results from the estimation of the single-agency model, estimated separately for the IRS and for the ODR, then briefly discuss the estimation of the piggyback equation (10), and then present and discuss results from the estimation of the remaining parameters of the joint-agency model. All of our econometric results are based on a weighted analysis that accounts for the choice-based sampling scheme we employed in collecting our data.

Table 5 presents results from estimation of the single-agency model for each audit class, while table 6 presents the results for parameters included in the joint-agency model but not in the single-agency models, with one exception. The exception is that no results are presented in table 6 for $\rho_{u_f u_s}$, the parameter that measures the degree of linear association between audit assessments when both agencies make a non-positive audit assessment. Because our data include very few observations where both agencies make such an assessment, we were unable to reliably estimate this parameter. We have therefore restricted the value of this parameter to zero for all audit classes. Notice that table 5 does not list the explanatory variables included in x_{1f}, x_{2f}, x_{1s}, and x_{2s} for any of the classes and also does not present the estimates of the associated parameters β_{1f}, β_{2f}, β_{1s}, and β_{2s}. Our contract with the Oregon Department of Revenue requires us not to disclose any information that might be used by others to infer the audit selection criteria of either the IRS or the ODR. As a result, we cannot reveal either the explanatory variables or the parameter estimates associated with the explanatory variables included in the federal and state revenue assessment models. We will discuss certain qualitative features of these variables below. Note also that both IRS and ODR audit assessments are measured in thousands of dollars in our analysis. The standard error estimates in tables 5 and 6 have been adjusted to account for choice-based sampling. However, we have not adjusted the standard error estimates in table 6 to account for the use of a multi-stage estimation procedure. Consequently, the estimated standard errors in this table may tend to overstate the precision of our parameter estimates to some extent.

We first discuss the results for the business class. There are 2,640 observations in our dataset in this class, of which 441 include federal returns that were selected for an IRS audit, 221 include state returns that were selected for an ODR independent audit, and 127 include state returns that were selected for an ODR piggyback audit. To determine what variables to include in x_{1f} and x_{2f} we used a specification search to determine which tax return characteristics were significantly related to IRS audit decisions and assessments. We used a similar procedure to determine which variables to include in x_{1s} and x_{2s}. Ultimately, only a few x variables were included in the model, indicating that within this relatively narrowly defined audit class only a few return characteristics were relevant for explaining audit selection decisions and assessments. Further, the set of variables that were included in x_{1f} and x_{2f} only partially overlapped with the set of variables included in x_{1s} and x_{2s}.

indicating that the IRS and the ODR relied on somewhat different variables in making their audit selection decisions for 1987 returns.

Now consider in more detail the results presented in table 5 for the business class. Three features of the results are especially noteworthy. First, the four correlations $\rho_{\eta_f u_f}$, $\rho_{\eta_f \epsilon_f}$, $\rho_{\eta_s u_s}$, and $\rho_{\eta_s \epsilon_s}$ are all positive, indicating that, for each of the IRS and the ODR, the value of the signal observed by the tax authority was positively correlated with both the likelihood of a positive assessment and, conditional on a positive assessment occurring, the magnitude of the assessment. The fact that the correlations are all relatively large suggests that both tax authorities possess extensive information about noncompliance behavior that we are unable to observe in our data. The largest of these correlations is $\rho_{\eta_s u_s}$, the correlation between the state signal and the likelihood of a positive state assessment, which equals .665; the other three correlations are all approximately equal to .2, though they are somewhat imprecisely estimated. Second, the correlation between the piggyback signal η_p and the stochastic disturbance ϵ_p in the piggyback equation is extremely large (.977). The fact that this correlation is so close to one indicates that the state knew almost exactly how much revenue it would earn from a piggyback audit of a 1987 return, a fact which is not particularly surprising, because the piggyback audit is based on the federal audit results and does not typically involve the examination of tax records. Third, the estimates of the audit cost variables $\lambda_f c_f$, $\lambda_s c_s$, and $\lambda_p c_p$ are all precisely estimated and sensible. The estimate of $\lambda_f c_f$ is 5.23 (recall that the revenue assessment variable is measured in thousands of dollars). From independent IRS data sources we have learned that the cost of a federal business audit of a 1987 return was approximately \$1,000. Hence we estimate that the shadow value associated with increasing the IRS audit budget in this audit class by one dollar is approximately five dollars, suggesting that if the IRS were allocated additional funds for business audits, net federal government revenue would increase. Our estimate of five for the shadow value of additional audit resources is far below previous estimates made by Dubin, Graetz, and Wilde (1990), which were based on the analysis of aggregate state-level data on federal audit assessments.[24] Our estimate is also somewhat below an estimate of eight made by Erard and Feinstein (1994b), based on estimation of a slightly different audit selection model and using data from the same business audit class. The estimate of $\lambda_s c_s$ is .684. Although we do not at the present time possess reliable information about the cost of independent state audits, we believe that these audits are less expensive than federal audits, and that c_s was probably in the neighborhood of \$500 for audits of 1987 returns. Therefore we estimate that the shadow value associated with increasing the ODR budget for independent business audits by one dollar is approximately equal to a dollar, suggesting that state revenue could not be increased through greater enforcement within this class. Finally, the estimate of $\lambda_p c_p$ is .271. As for independent state audits, we do not possess reliable information about the cost of state piggyback audits; however, we doubt that this cost exceeded \$100 for 1987 returns. Hence we estimate that the shadow value associated with increasing the ODR budget for piggyback business audits is between two and three dollars, which indicates that the ODR could have increased audit revenues within the business class by reallocating resources from independent audits to piggyback audits. This conclusion corroborates our earlier remark, made with respect to table 4A, that the ODR appears to conduct many independent audits but relatively few piggyback audits of business filers.

Now consider the results presented in table 6 for the business class. The estimates for all five of the correlations listed in table 6 are positive. The correlation $\rho_{\eta_f \eta_s}$ appears to be quite precisely estimated at .449. The estimates of the two cross correlations $\rho_{\eta_f \epsilon_s}$ and $\rho_{\eta_s \epsilon_f}$ are .189 and .316, respectively, which are similar in magnitude to the estimates of the single-agency correlations $\rho_{\eta_f \epsilon_f}$ and $\rho_{\eta_s \epsilon_s}$ reported in table 5. The large positive estimate of $\rho_{\eta_f \eta_s}$ indicates that much of the information that the IRS possesses about 1987 business filers that we are unable to observe is also possessed by the ODR. The positive estimate of $\rho_{\eta_f \epsilon_s}$ indicates that the information the IRS possesses is positively correlated with the size of the state assessment. An interesting question is whether this information adds anything to the knowledge already possessed by the ODR about the state assessment. To find out, we have used our results to compute an estimate of the partial correlation between the federal signal η_f and the state error term ϵ_s, conditional on the value of the state signal η_s.[25] The estimated value of this partial correlation is .11, which indicates that the federal signal contains some information about the state assessment that is not contained in the state signal. Thus, it would appear that the ODR could improve its business audit selection procedures if it were made privy to more federal information. Similarly, the large positive estimate of $\rho_{\eta_s \epsilon_f}$ indicates that the state may possess information that would be helpful to the IRS in predicting federal assessments. Indeed, our computations indicate that the implied partial correlation between the state signal η_s and the federal error term ϵ_f conditional on the value of the federal signal η_f is equal to .23, which confirms that the state signal contains information about federal assessments not contained in the federal signal. These findings are consistent with the figures presented in table 4B, which indicate that in the business classes revenue assessments are much higher, for both the IRS and the ODR, when both authorities conduct an audit than when only one of the two conducts an audit.

Two conclusions follow from these observations. First, the results in table 6 for the joint-agency model are consistent with the results presented in table 5 for the single-agency models. In particular, both sets of results indicate that both the IRS and the ODR possess extensive information about compliance characteristics of 1987 business filers beyond what we are able to infer from individual tax records. Second, although much of this information is possessed in common by the two authorities, each authority also possesses some unique information that is correlated with the other agency's revenue assessments. This second conclusion leads us to believe that for business classes additional information sharing between the two tax authorities might significantly improve audit selection procedures. The estimates of the two remaining correlations for the business class listed in table 6, $\rho_{u_f u_s}$, and $\rho_{\epsilon_f \epsilon_s}$, are both very large and positive (.788 and .592, respectively). These results are sensible and are consistent both with the hypothesis that taxpayers who cheat on their federal return also cheat on their state return and with the complementary hypothesis that taxpayers who make a mistake on their federal return carry the mistake over to their state return.

Having discussed the results for the business class in some detail, we now discuss the results for the farm and nonbusiness, nonfarm classes. Our dataset includes 1,945 observations pertaining to the farm audit class and 9,239 observations pertaining to the nonbusiness, nonfarm class. These data include information on 148 federal farm audits, 856 federal nonbusiness, nonfarm audits, 77 independent state farm audits, and 200 independent state nonbusiness, nonfarm audits. The data also include information on 23 farm and 279 nonbusiness, nonfarm state piggyback audit cases. As for the business class, we used specification searches to determine the variables to include in x_{1f}, x_{2f}, x_{1s}, and x_{2s} for our farm and nonbusiness, nonfarm classes. For the farm class our search resulted in the inclusion of a relatively small number of tax return characteristics as explanatory variables in both the federal and state assessment equations, just as for the business class. However, for the nonbusiness, nonfarm class, our search resulted in the inclusion of a fairly large number of explanatory variables in the federal assessment equations and a quite small number of explanatory variables for the state assessment equations. We do not think it is surprising that more tax return characteristics were needed to explain federal audit selection decisions and assessments in this class than in the business and farm classes, because this class includes a much wider variety of tax returns and patterns of noncompliance. We also do not find it surprising that few variables were important in explaining state audit selection decisions in this class. because the ODR appears to devote most of its independent audit resources to business and farm returns.

Consider now the results for the farm and nonbusiness. nonfarm classes reported in table 5. For the farm class the results indicate that the private signal of the IRS has a positive and significant correlation with the size of a positive audit assessment. In contrast. the private signal of the ODR for this class is positively and significantly related to the probability of a positive assessment. The estimated correlation between the state signal η_s and the error term u_s (.819) is very large and quite precise, indicating that the state possesses very good information about the likelihood of a positive farm audit assessment. We do not find it surprising that the state possesses better information than the federal government about the likelihood of farm noncompliance. because farmers are spread across the state. and while the ODR maintains more than a dozen separate tax offices. the IRS operates out of a single district office located in Portland. This is one filer group for which local information seems especially important. Interestingly. while the ODR seems to possess superior information about the likelihood of a positive audit assessment. the IRS seems to possess better information about the probable magnitude of such an assessment. Indeed. the estimated correlation between the ODR signal η_s and the error term ϵ_s converged to zero in estimation. indicating that ODR has no private information beyond what can be obtained from tax return data about the likely magnitude of a positive assessment. For the nonbusiness. nonfarm class the results indicate a positive correlation between each agency's private signal and the likelihood of a positive assessment. However. the values for $\rho_{\eta_f \epsilon_f}$ and $\rho_{\eta_s \epsilon_s}$ converged to zero in estimation for this audit class. indicating that neither agency possesses private information about the likely magnitude of a positive audit assessment for a nonbusiness. nonfarm return. The estimated value of $\rho_{\eta_p \epsilon_p}$ is very close to one for each of these classes, again indicating that the state has extremely accurate information about the potential revenue associated with a piggyback audit. The estimated values of the audit cost parameters $\lambda_f c_f$ and $\lambda_s c_s$ are approximately 2.5 and .95, respectively, both for the farm class and the nonbusiness, nonfarm class. Farm audits tend to be about as costly as business audits for the IRS, suggesting that the shadow value associated with increasing IRS farm audit resources by a dollar is approximately \$2.50, only about one-half as much as the corresponding figure for the IRS business class. Federal nonbusiness, nonfarm audits tend to be the least costly type of federal audit. Based on an estimate of \$600 as the average cost of such an audit, the implied shadow value of increasing IRS nonbusiness, nonfarm audit resources by one dollar is approximately four dollars, somewhat below the estimated shadow value for business audits but well above the estimated shadow value for farm audits. Apparently the IRS could have obtained greater revenue by performing more audits of business returns for tax year 1987 and less audits of farm returns. We assume that state farm audits cost about the same as state business audits (roughly \$500) and that state nonbusiness, nonfarm audits are somewhat less expensive (perhaps \$300). Based on these assumptions, we estimate that the shadow value of an additional dollar of state audit resources is about \$2 for a farm audit and \$3 for a nonbusiness, nonfarm audit. Our estimates imply that the state could have obtained greater revenue by allocating more resources to nonbusiness, nonfarm audits and less to business audits. Our estimate of $\lambda_p c_p$ is negative but very imprecise for the farm class. For the nonbusiness, nonfarm class the estimate is .28, quite similar to the value obtained for the business class. Based on an assumed piggyback audit cost of about \$100, we estimate that the shadow value of an additional dollar of piggyback audit resources for the nonbusiness, nonfarm class is between \$2 and \$3.

Now consider the results for the farm and nonbusiness, nonfarm classes presented in table 6. For both classes, the estimated correlation between the state and federal signals is high, indicating a substantial degree of overlap between the information sets of the two agencies. For the farm class, the estimated cross-correlations $\rho_{\eta_f \epsilon_s}$ and $\rho_{\eta_s \epsilon_f}$ are quite large indicating that each agency has information that pertains to the other agency's audit assessment distribution. To determine whether the information the IRS possesses would be useful to the ODR in formulating its audit strategy, we have again computed the implied partial correlation between the federal signal η_f and the state error term ϵ_s conditional on the value of the state signal η_s. The value of this partial correlation for the farm class is .54, indicating that the federal signal contains a great deal of information about state audit assessments not contained in the state signal. We have also computed the implied partial correlation between the state signal η_s and the federal error term ϵ_f conditional on the value of the federal signal η_f for the farm class. The value of this correlation is .60, which indicates that the state signal contains a great deal of information about federal assessments not contained in the federal signal. Thus, it appears that greater information sharing between the two agencies might result in substantial improvements to state and federal farm audit selection procedures. On the other hand, the estimated cross correlations for the nonbusiness, nonfarm class are small in absolute value and statistically insignificant. Apparently, there is less scope for improved audit selection through information sharing arrangements for this class. Just as for the business class, the estimates of the two remaining correlations listed in table 6, $\rho_{w_f w_s}$ and $\rho_{\epsilon_f \epsilon_s}$, are positive and generally quite large for the farm and nonbusiness, nonfarm audit classes, indicating a strong link between compliance behavior on federal and state tax returns.

3.3 Additional Findings

Over the past few months we have conducted a detailed analysis of the ODR individual income tax audit programs. Our analysis has generated two findings that we believe may be of some general interest and applicable to other states. First, as was previously indicated during our discussion of table 3, our data indicates that the ODR did not follow up on a significant number of cases for which the IRS conducted an audit and made a significant positive assessment. Table 7A provides some additional information about this finding. According to the statistics presented in the table, of the 198 federal audit cases in our data for which the IRS assessed more than \$10,000, Oregon did not follow up on 106, or more than one-half. Of the 249 federal audit cases in our data for which the IRS assessed between \$5,000 and \$10,000, Oregon did not follow up on 142, again more than one-half. Of the 1,428 federal audit cases for which the IRS assessed between \$1,000 and \$5,000, Oregon did not follow up on 845. Surprisingly, Oregon did follow up on many cases for which the IRS made far lower assessments, as is also indicated in the table. Of course it is possible that our data are wrong, and that Oregon did follow up on many of the audits in question. We are still discussing this possibility with Oregon officials, but, at the present time, have discovered no evidence in support of this hypothesis. We believe that the more likely explanation for the above finding is that the ODR either did not receive information about many of these audits or did not retain the information for later use. We wonder whether the results presented in table 7A are similar for other states.

Our second finding is that for tax year 1987 ODR audits of part-year and non-resident filers generated larger mean and median assessments than were generated by audits of any other filer groups. Table 7B presents additional information related to this point. The figures in the table indicate that the ODR conducted very few audits of part-year and non-resident filers for tax year 1987 but made large revenue assessments. If this second finding is correct and carries over to more recent years, it suggests that the ODR might be able to increase its revenue assessments by shifting some of its audit resources to these two groups. Again, we wonder whether the statistics for other states are similar.

4. Conclusion

Little is known about state audit programs and about the relationship between these state programs and federal tax enforcement activities. In this paper we use data provided by the Oregon Department of Revenue and the Internal Revenue Service to examine various aspects of these programs. In particular, we find that there is substantial - though not complete - overlap between the information employed by the ODR and the IRS for audit selection, and that the information used by one agency is not always made known to the other. This result suggests that the enforcement activities of each agency could be improved through greater information sharing. We also find that the ODR may have foregone substantial amounts of audit revenue by not following up on more returns for which IRS had made large audit assessments. Finally, we find that each agency could increase its collections by reallocating its enforcement budget among the various audit classes, although the efficient reallocation is not the same for the ODR and the IRS.

It should be remembered that these results are based on an analysis of the interactions between federal and state enforcement efforts in a single state (Oregon). The results might differ if other states were examined. Indeed, our survey of state audit programs clearly indicates that the individual income tax enforcement programs in Oregon are far more extensive and effective than in most other states. The shadow value of additional audit resources, the scope for resource reallocation, and the benefits from additional federal-state information sharing may actually be much higher for some of these other states than we have found for Oregon.

Appendix

In this appendix we analyze in somewhat greater detail the ***strategic*** version of our federal and state model, which was mentioned previously in subsection 2.2. Recall that in this version of the model, the state tax authority makes a decision whether to perform an audit in the first period, taking into account the possibility that it may have the opportunity to piggyback on a federal audit in the second period if it should choose not to audit in the first period. Figure 1 illustrates the decision tree associated with the auditing choice faced by the state tax authority in the first period of the model.

The top branch of the tree corresponds to the decision to conduct a period one audit: the expected value associated with this branch is $E(R_s|\eta_s) - \lambda_s c_s$, which is identical to the expected value associated with a period one audit in the nonstrategic model. The bottom branch of the tree corresponds to the decision to wait. In the nonstrategic model, the value associated with this branch is zero, and the state chooses to conduct a period one audit whenever $E(R_s|\eta_s) \geq \lambda_s c_s$. However, in the strategic model this branch has a more complex structure, and the value associated with waiting is typically positive.

To demonstrate that the value associated with this branch is generally positive, we describe the logical structure of this branch in detail. The wait branch leads first to a random event involving the federal authority's period one audit decision. If the federal authority chooses not to audit, the value associated with the wait branch of the tree is zero. However, if the federal authority chooses to audit, the federal assessment R_f is revealed, and the state authority faces a further decision in period two, whether or not to conduct a piggyback audit. The federal authority's audit selection decision is the same as in the nonstrategic model, and therefore there is a threshold value η_f^* such that the federal authority chooses to conduct a period one audit if and only if $\eta_f \geq \eta_f^*$. Since the signal η_s is observed by the state authority and is correlated with the signal η_f, the state authority assesses the probability of a federal audit as $Prob(\eta_f \geq \eta_f^*|\eta_s)$, which can be written as a function $L(\eta_s)$. Suppose now that the federal authority has conducted a period one audit and has made a revenue assessment R_f. At the beginning of period two the state observes the signal η_p and computes the net expected value of performing a piggyback audit to be $E(R_p|R_f, \eta_p) - \lambda_p c_p$. Since the net value associated with no piggyback audit is zero, the state will choose to perform a piggyback audit whenever $E(R_p|R_f, \eta_p) > \lambda_p c_p$, or whenever η_p exceeds the threshold value η_p^*. The threshold value η_p^* is itself a function of R_f, denoted $\eta_p^*(R_f)$. What expected value should the state authority then assign to the option of performing a piggyback audit in period two, as of period one? Since the option is relevant only when the federal authority conducts a period one audit, the state authority must evaluate the distribution of R_f, conditional on the signal η_s and on the event $\eta_f > \eta_f^*$, and then use this conditional distribution to compute the expected value of the piggyback option. Applying the logic of this argument, the expected value of the piggyback option may be denoted $Q(\eta_s)$ and be expressed as

$$(17) \qquad Q(\eta_s) = \int_{R_f=-\infty}^{+\infty} \int_{\eta_p=\eta_p^*(R_f)}^{+\infty} E(R_p | R_f, \eta_p) g(\eta_p) f(R_f | \eta_s, \eta_f > \eta_f^*) \, d\eta_p \, dR_f.$$

The expected value associated with the wait branch of the decision tree is then $L(\eta_s)Q(\eta_s)$, which in general is positive.

TABLE 1

Collection and Enforcement Activities by the States and the IRS for the Individual Income Tax, 1992

	State Individual Income Tax Audit Budget, Average Across the States	Total State Tax Agency Budget, Average Across the States	IRS State-Level Budget, Average Across the States
Agency Budget (dollars per capita)	$0.40	$12.61	$16.33
	For State Independent Audits, Average Across the States	For RARs and CP2000 Reports, Average Across the States	For the IRS
Audit Rate (percentage)	0.33%	0.75%	0.91%
	For the States from Independent Audits, Average Across the States	For the States from RARs and CP2000 Reports, Average Across the States	For the IRS
Additional Assessments for the Individual Income Tax (dollars per capita)	$3.78	$5.04	$23.66
	For States in the Survey, Average Across the States	For All States, Average Across the States	For the IRS
Individual Income Tax Collections (dollars per capita)	$539	$486	$3,738

Note: All averages are simple unweighted averages. State data on individual income tax collections, additional assessments, audit rates, and agency budgets are calculated from the survey of state tax administrators, as discussed in the text. IRS data are calculated from the Internal Revenue Service 1992 Annual Report.

TABLE 2

Audit Results by IRS Audit Category

	Total	Business	Farm	Nonbusiness, Nonfarm
Unweighted Number Returns	43,587	6,492	1,945	35,150
Weighted Number Returns	1,012,023	63,611	7,307	941,105
Federal Audit Cases				
Overall Frequency	4,433	1,073	148	3,212
No-Change Frequency	1,007	210	62	735
Neg. Change Cases:				
Frequency	259	78	9	172
Median Change	283	698	319	209
Mean Change	1,330	3,287	614	713
Pos. Change Cases:				
Frequency	3,167	785	77	2,305
Median Change	1,021	1,444	1,240	924
Mean Change	3,073	5,502	4,310	2,472
State Independent Audits				
Overall Frequency	1,667	802	77	788
No-Change Frequency	563	312	37	214
Neg. Change Cases:				
Frequency	83	30	2	51
Median Change	225	214	184	229
Mean Change	578	428	184	651
Pos. Change Cases:				
Frequency	1,021	460	38	523
Median Change	399	442	582	328
Mean Change	1,051	1,060	1,317	1,032
State Piggyback Audits				
Frequency	1,158	280	23	855
Median Change	324	374	172	319
Mean Change	762	1,005	414	710

Note: Frequencies are unweighted; dollar amounts are weighted to reflect population totals

TABLE 3

IRS-Oregon Audit Interactions (All Returns)

		OREGON			
		No Audit	Independent Audit	Piggyback Audit	Row Totals
	No Audit	1,004,408 (99.25%)	1,624 (.16%) Mean: 466 Median: 32	Empty	1,006,032 (99.41%)
IRS	Audit	4,175 (.41%) Mean: 1,572 Median: 294	328 (.03%) Fed. Mean: 5,746 Fed. Median: 1,924 State Mean: 1,419 State Median: 644	1,488 (.15%) Fed. Mean: 2,965 Fed. Median: 960 State Mean: 762 State Median: 324	5,991 (.59%) Mean: 2,146 Median: 596
	Column Totals	1,008,503 (99.66%)	1,952 (.19%) State Mean: 626 State Median: 79	(As Above)	1,012,023 (100%)

Note: Figures weighted to reflect population totals

TABLE 4A

IRS-Oregon Audit Interactions

(Returns in IRS Business Audit Classes)

		OREGON			
		No Audit	Independent Audit	Piggyback Audit	Row Totals
	No Audit	61,829 (97.2%)	679 (1.07%) Mean: 430 Median: 1	Empty	62,508 (98.27%)
IRS	Audit	693 (1.09%) Mean: 3,131 Median: 337	126 (.20%) Fed. Mean: 5,830 Fed. Median: 2,431 State Mean: 1,478 State Median: 857	284 (.45%) Fed. Mean: 4,551 Fed. Median: 1,527 State Mean: 1,005 State Median: 374	1,103 (1.73%) Mean: 3,805 Median: 777
	Column Totals	62,522 (98.29%)	805 (1.27%) State Mean: 594 State Median: 52	(As Above)	63,611 (100%)

Note: Figures weighted to reflect population totals

TABLE 4B

IRS-Oregon Audit Interactions

(Returns in IRS Farm Audit Classes)

		OREGON			
		No Audit	Independent Audit	Piggyback Audit	Row Totals
IRS	No Audit	7,094 (97.09%)	63 (.86%) Mean: 252 Median: 0	Empty	7,157 (97.95%)
	Audit	113 (1.55%) Mean: 874 Median: 0	14 (.19%) Fed. Mean: 13,987 Fed. Median: 2,007 State Mean: 4,797 State Median: 777	23 (.31%) Fed. Mean: 1,757 Fed. Median: 728 State Mean: 414 State Median: 172	150 (2.05%) Mean: 2,233 Median: 1
	Column Totals	7,207 (98.63%)	77 (1.05%) State Mean: 645 State Median: 0	(As Above)	7,307 (100%)

Note: Figures weighted to reflect population totals

TABLE 4C

IRS-Oregon Audit Interactions

(Returns in IRS Nonbusiness, Nonfarm Classes)

OREGON

		No Audit	Independent Audit	Piggyback Audit	Row Totals
IRS	No Audit	935,485 (99.40%)	882 (.09%) Mean: 508 Median: 76	Empty	936,367 (99.50%)
	Audit	3,369 (.36%) Mean: 1,274 Median: 310	188 (.02%) Fed. Mean: 5,077 Fed. Median: 1,698 State Mean: 1,305 State Median: 448	1181 (.13%) Fed. Mean: 2,607 Fed. Median: 849 State Mean: 710 State Median: 319	4,738 (.50%) Mean: 1,757 Median: 570
	Column Totals	938,854 (99.76%)	1,070 (.11%) State Mean: 648 State Median: 99	(As Above)	941,105 (100%)

Note: Figures weighted to reflect population totals

TABLE 5

Estimation Results for the Single-Agency Model

IRS Results

	Business Class	Farm Class	Nonbusiness, Nonfarm Class
Correlation $\rho_{\varepsilon_f u_f}$	.1992 (0.181)	.1571 (0.216)	.1824 (0.107)
Correlation $\rho_{\varepsilon_f e_f}$	.2695 (0.158)	.3910* (0.097)	0.00-
Standard Deviation σ_{u_f}	.9841* (0.252)	1.334* (0.403)	1.465* (0.445)
Standard Deviation σ_{e_f}	2.205* (0.166)	2.210* (0.298)	1.776* (0.089)
Shadow Cost $\lambda_f c_f$	5.227* (1.560)	2.518* (1.270)	2.347* (0.341)
Number of Observations	2,640	1,945	9,239
Number of Audit Cases	441	148	856
Log-Likelihood Value	-318.4	-268.6	-375.3

Oregon Results

	Business Class	Farm Class	Nonbusiness, Nonfarm Class
Correlation $\rho_{\varepsilon_s u_s}$	.6654* (0.229)	.8191* (0.087)	.3296 (0.262)
Correlation $\rho_{\varepsilon_s e_s}$	.2144 (0.125)	0.00-	0.00-
Correlation $\rho_{\varepsilon_p e_p}$	.9772* (0.007)	.9690* (0.012)	.9876* (0.006)
Standard Deviation σ_{u_s}	.4951* (0.184)	.3868* (0.068)	.7307* (0.219)
Standard Deviation σ_{e_s}	1.810* (0.179)	2.015* (0.333)	1.938* (0.178)
Standard Deviation σ_{e_p}	.4409* (0.065)	.1391 (0.078)	.3785* (0.057)
Shadow Cost $\lambda_s c_s$	.6842* (0.176)	.9382* (0.458)	.9612* (0.260)
Shadow Cost $\lambda_p c_p$	.2705* (0.039)	-.4124 (0.687)	.2801* (0.052)
Number of Observations	2,640	1,945	9,239
Number of Independent Audits	221	77	200
Number of Piggyback Audits	127	23	279
Log-Likelihood Value	-174.5	-136.8	-121.1

(adjusted standard errors in parentheses)

*Significant at the 5% level

-The parameter is constrained to zero

TABLE 6

Joint Agency Model Results

	Business Class	Farm Class	Nonbusiness, Nonfarm Class
$\rho_{\psi f \psi s}$	.4486* (0.043)	.5485* (0.074)	.4611* (0.033)
$\rho_{\psi f e s}$	.1890 (0.155)	.4509 (0.270)	-.0485 (0.095)
$\rho_{\psi s e f}$	.3156* (0.087)	.6763* (0.154)	-.0486 (0.087)
$\rho_{w f w s}$	.7876* (0.051)	.6371 (1.111)	.7087* (0.078)
$\rho_{e f e s}$	.5921 (0.315)	.2722 (0.322)	.8795* (0.056)
Log-Likelihood Value	-461.0	-348.3	-470.0

(partially adjusted standard errors in parentheses)

*Significant at the 5% level

TABLE 7A

Taxpayers in Sample with a Positive Federal Audit Assessment by Whether They Were Subjected to a State Audit of Any Kind

Federal Audit Assessment (in dollars)	Number of Cases with No State Audit	Number of Cases with a State Audit
1 - 100	183	61
101 - 500	405	320
501 - 1,000	596	275
1,001 - 5,000	845	583
5,001 - 10,000	142	107
Over 10,000	106	92
All Cases	2,277	1,438

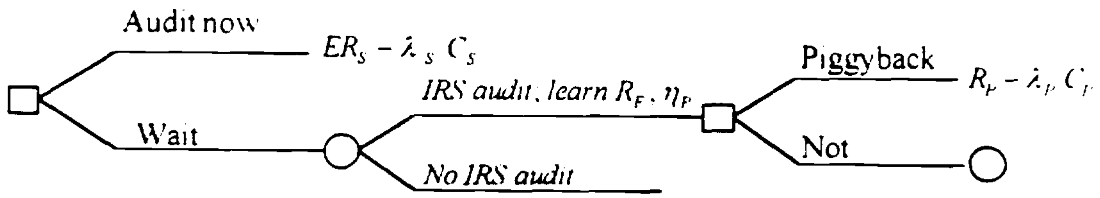
TABLE 7B

State Audit Results by Type of Form

	Form 40F	Form 40S	Form 40P	Form 40N
Number of Audits	1,913	100	36	75
Mean Audit Assessment	733	340	1,175	1,871
Median Audit Assessment	88	77	215	744

Note: Figures weighted to reflect population totals

Figure 1


Audit now
$ER_S - \lambda_S C_S$
Wait
IRS audit; learn R_F, η_P
No IRS audit
Piggyback
$R_P - \lambda_P C_P$
Not

Endnotes

* We thank officials at the Oregon Department of Revenue and the Internal Revenue Service for the enormous amount of help they provided in assembling the dataset used in this paper. We also thank the many state tax administrators who filled out our tax audit survey and returned it to us. Finally, we thank participants at the NBER preconference meeting in August 1994 and at the NBER Tax Policy Analysis Conference in January 1995 for helpful comments and suggestions, especially James Wetzler and Jim Poterba.

1. Note that compliance and enforcement issues surrounding state sales taxes have been examined in some detail. See Due and Mikesell (1993).

2. These states are Alabama, Arizona, Arkansas, California, Connecticut, Florida, Hawaii, Illinois, Indiana, Iowa, Kansas, Louisiana, Minnesota, Missouri, Nebraska, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Oregon, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Utah, Vermont, Washington, Wisconsin, and Wyoming. Several states (Maryland, Michigan, and Massachusetts) have declined to respond to the survey. Most of the remaining states have indicated that they will eventually provide some information. It should also be noted that the survey is not limited to the individual income tax: other major state taxes and their associated audit programs are also surveyed.

3. It should be noted that the IRS state budget information includes all expenditures for IRS service offices located in the state, including agency costs for taxes other than the individual income tax, and so may not be directly comparable to state individual income tax audit costs. If instead the comparison is between total state tax agency budgets and IRS state budgets, where the state agency budgets include all costs associated with all state taxes, then state budgets are on average closer (though still somewhat smaller) than IRS state budgets.

4. The only two previous studies of which we are aware that distinguish negative from zero assessments are Alexander and Feinstein (1987) and Erard (1995).

5. Alexander and Feinstein (1987) and Erard (1995) both present a more elaborate model of taxpayer reporting errors and underreports. In particular, both studies present models in which errors are symmetrically distributed around zero, negative

assessments are always due to error, and positive assessments are due either to error or to intentional evasion.

6. It should also be noted that the amount the taxpayer is assessed during an audit may differ from the amount that is eventually received by the tax authority, either because the assessment is later reduced following an appeal by the taxpayer, or because the taxpayer fails to pay the assessed amount. We leave it to future research to incorporate these issues into a model of compliance and enforcement.

7. We note that either one of the correlations may be fully identified, but then only the absolute value of the other correlation is identified, because of the form of the likelihood function.

8. This assumption is always satisfied when $\rho_{\eta w}$ and $\rho_{\eta \epsilon}$ are both positive. The condition is also satisfied if one correlation is positive and the other negative, provided that the negative correlation is not too large in absolute value relative to the positive correlation. Our econometric estimation has never produced estimates for which the condition fails to hold.

9. In actual practice, the state may need to contact the taxpayer if the information on the taxpayer's state return is insufficient to determine the tax consequences of the federal audit results.

10. The two λ values also might differ if the shadow value associated with state budgetary resources is different in time periods one and two.

11. A consideration of the covariance matrix associated with this multivariate normal distribution may clarify the restrictions we have imposed. Order the disturbances as η_f, η_s, w_f, w_s, ϵ_f, and ϵ_s. The covariance matrix is six by six. Consider the upper triangular portion of this matrix. The first four diagonal entries are one. The next two diagonal entries are $\sigma^2_{\epsilon_f}$ and $\sigma^2_{\epsilon_s}$. In the first row the remaining elements are $\rho_{\eta_f \eta_s}$, $\rho_{\eta_f w_f}$, $\rho_{\eta_s w_s} \rho_{\eta_f \eta_s}$, $\sigma_{\eta_f \epsilon_f}$, and $\sigma_{\eta_f \epsilon_s}$. In the second row the remaining elements are $\rho_{\eta_s \eta_f} \rho_{\eta_f w_f}$, $\rho_{\eta_s w_s}$, $\sigma_{\eta_s \epsilon_f}$, and $\sigma_{\eta_s \epsilon_s}$. In the third row the remaining elements are $\rho_{w_f w_s}$, $\rho_{\eta_f w_f} \sigma_{\eta_f \epsilon_f}$, and $\rho_{\eta_f w_f} \sigma_{\eta_f \epsilon_s}$. In the fourth row the remaining elements are $\rho_{\eta_s w_s} \sigma_{\eta_s \epsilon_f}$ and $\rho_{\eta_s w_s} \sigma_{\eta_s \epsilon_s}$. In the fifth row the remaining element is $\sigma_{\epsilon_f \epsilon_s}$.

12. As described previously, we have imposed restrictions on the other two cross correlations, $\rho_{\eta_f w_s}$ and $\rho_{\eta_s w_f}$. In future analysis, we may allow these parameters to be free as well. However, when all four cross correlations are free, the single-agency

model estimates are not consistent in the joint-agency model, which significantly complicates estimation; see our discussion of estimation strategy in the main text.

13. The results of Alexander and Feinstein (1987), Feinstein (1991), and Erard (1995) all indicate that detection is quite imperfect. See the discussion of empirical results, in particular footnote 21, for some information about the relationship between federal and state assessments among those individuals and households selected for audit by both authorities.

14. In fact, all that is required for the estimation procedure to yield consistent estimates is that either the two equalities in the text hold and/or the following two equalities hold: $\rho_{\eta_f \epsilon_s} = \rho_{\eta_f \eta_s} \rho_{\eta_s \epsilon_s}$ and $\rho_{\eta_s \epsilon_f} = \rho_{\eta_s \eta_f} \rho_{\eta_f \epsilon_f}$.

15. Approximately 8.5% of our sample of Oregon returns failed to match with any IRTF record. When weighted, these numbers indicate that approximately 9.3% of all Oregon returns in the population would fail to match with any IRTF record. About 70% of all returns that fail to match are part-year or non-resident returns, which we exclude from our analysis in any case.

16. Excluding these kinds of returns reduced our sample by 5,511 observations, reduced the number of federal audits by 813, the number of independent Oregon audits by 94, and the number of Oregon piggyback audits by 48. We discuss part-year and non-resident filers in subsection 3.3.

17. It is important to note that many filers who possess some business or farm income nonetheless are not placed in a business or farm audit class. The IRS has specific, somewhat complex rules for assigning returns to audit classes. In general, a return is assigned to a business or farm class only if total business or farm gross receipts are sufficiently large relative to the taxpayer's total positive nonbusiness income, as calculated based on the information reported on the return.

18. The ODR has informed us that our sample may include a small number of coding errors in which piggyback audits were misclassified as independent audits.

19. Our dataset also contains information about Oregon audits based on the IRS information returns matching program, but we neither list these audits in table 2 nor use them in our analysis.

20. Note that the mean and median figures in table 3 are based on information about all audits, including those that resulted in either a negative assessment or no additional

assessment.

21. The returns in this cell provide interesting information about the ability of tax examiners to detect noncompliance, because each return is subjected to two independent audits, resulting in two separate assessments. We explored the relationship between the federal and state assessments for the 283 returns (unweighted frequency) in this cell. Since Oregon's tax rate is (approximately) a flat 9%, we multiplied the Oregon assessment by 2.5, so that the average Oregon assessment was similar to the average federal assessment. We considered first the 184 cases for which the federal assessment exceeded $1,000. For these cases, 14 Oregon assessments were non-positive, 25 (of the adjusted assessments, multiplied by 2.5) were less than one-quarter as large as the federal assessment, 17 were between one-quarter and one-half as large, 107 were between three-quarters and one and one-half times as large, and 21 were more than one and one-half times as large. Next we considered the 35 cases for which the federal assessment was non-positive. For these cases, 17 Oregon assessments were also non-positive, 7 were between $0 and $400, and 6 were greater than $1,000. Since the federal tax rate varies with income, we examined how our results varied with variations in the reported federal tax balance and adjusted gross income, but found that the results were not sensitive to the levels of these variables.

22. There are two other IRS business audit classes. One consists of all business returns for which the reported total gross receipts are below $25,000, and the other consists of all business returns for which the reported total gross receipts are above $100,000.

23. There are four other nonbusiness, nonfarm IRS classes. The majority of returns in our sample from this class reported some schedule C income.

24. It should be noted that the estimate by Dubin, Graetz, and Wilde (1990) is intended to account for both the direct gain in audit revenue and any revenue resulting from increased deterrence. In contrast, our measure accounts solely for the direct gain in audit revenue.

25. The partial correlation coefficient is defined as $\frac{\rho_{\eta_f \epsilon_s} - \rho_{\eta_f \eta_s} \rho_{\epsilon_s \eta_s}}{\sqrt{1-\rho^2_{\eta_f \eta_s}}\sqrt{1-\rho^2_{\epsilon_s \eta_s}}}$.

References

Alexander, Craig and Jonathan S. Feinstein. (1987) "A Microeconometric Analysis of Income Tax Evasion." Mimeo, M.I.T..

Alm, James, Roy Bahl, and Matthew N. Murray. (1993) "Audit Selection and Income Tax Underreporting in the Tax Compliance Game." *Journal of Development Economics*, Vol. 42, No. 1, pages 1-33.

Beron, Kurt, Helen Tauchen, and Anne D. Witte. (1991) "The Effects of Audits and Socioeconomic Variables on Compliance." In *Why People Pay Taxes*, edited by Joel Slemrod. Ann Arbor: University of Michigan Press, pages 67-89.

Clotfelter, Charles. (1983) "Tax Evasion and Tax Rates: An Analysis of Individual Returns." *Review of Economics and Statistics*, Vol. 65, No. 3, pages 363-373.

Dubin, Jeffrey A., Michael A. Graetz, and Louis L. Wilde. (1990) "The Effect of Audit Rates on the Federal Individual Income Tax." *National Tax Journal*, Vol. 43, No. 4, pages 395-409.

Due, John F. and Mikesell, John L. (1993) *Sales Taxation: State and Local Structure and Administration*. Baltimore: The Johns Hopkins University Press.

Erard, Brian. (1995) "Self Selection with Measurement Errors: A Microeconometric Analysis of the Decision to Seek Tax Assistance and Its Implications for Tax Compliance." Mimeo, Carleton University.

Erard, Brian and Jonathan S. Feinstein. (1994a) "Honesty and evasion in the tax compliance game." *RAND Journal of Economics*, Vol. 25, No. 1 (Spring), pages 1-19.

Erard, Brian and Jonathan S. Feinstein. (1994b) "Econometric Models of Compliance and Enforcement: Reporting Behavior and Audit Selection Decisions." Mimeo.

Erard, Brian and Jonathan S. Feinstein. (1995) "The Role of Moral Sentiments and Audit Perceptions in Tax Compliance." Forthcoming in *Public Finance*.

Feinstein, Jonathan S. (1991) "An econometric analysis of income tax evasion and its detection." *RAND Journal of Economics*, Vol. 22, No. 1 (Spring), pages 14-35.