Predicting Aggregate Taxpayer Compliance Behavior

Alan Plumley, Internal Revenue Service; Brian Erard, Brian Erard & Associates; and Derek Snaidauf, IBM Global Business Services[1]

The IRS seeks to be able to estimate the impact that its service and enforcement interventions have on the voluntary compliance of taxpayers. A key method for doing that is to conduct field experiments among taxpayers as they respond to their real tax obligations. A typical experimental design includes tracking the behavior of both a test group and a control group, which are sufficiently identical in all relevant respects except that the intervention is applied only to the test group. We believe, however, given the nature of most IRS interventions and the diversity and geographically non-uniform distribution of the taxpayer population, that it will often be very difficult to construct a control group that is sufficiently identical to a test group in all relevant respects, except for the intervention (or in preventing the control group from being affected by the intervention). The goal of this research, therefore, was to develop the capability to control statistically for factors that influence taxpayer behavior, supplementing the role of control groups in future field experiments. To do this, we develop econometric models to predict aggregate reporting behavior among individuals. That is, we seek to estimate what taxpayers would have done in the absence of an intervention. The difference between what they would have done and what they actually did in any given time period is a measure of the change in voluntary compliance. If that measure of behavioral change improves as a result of an intervention, then the intervention is considered to have improved compliance.

Data

We extend in this paper the foundational studies by Plumley (1996) and Dubin (2007), which were focused on estimating the independent effects of many of the determinants of voluntary compliance. Our methodology is similar, but our focus is on developing accurate predictions of the dependent variable: taxpayer reporting behavior. To do this, we compiled a robust database containing detailed state-level longitudinal data on taxpayer characteristics and behavior, as well as IRS activities, from Tax Years 1982 through 2009. This includes data for 750 variables from over 20 different sources. Approximately 200 of these variables were updated from both Dubin's and Plumley's studies and about 550 additional variables, not included in these earlier databases, were incorporated to provide a richer set of potential determinants of taxpayer behavior. Beyond our current empirical work, the analysis database will serve as a valuable tool for researchers to employ when analyzing various issues relating to taxpayer filing and reporting behavior across states and over time. To ensure uniformity between all our variables, we combined the District of Columbia with Maryland, since data from the District of Columbia was not available for all variables and it was already combined with Maryland in previous studies.

In the case of the non-IRS source data, a major task was the construction of variables from the Annual Social and Economic Supplement of the Current Population Survey conducted by the Census Bureau. This data source includes nationally representative micro-level cross-sectional data on residents of housing units (homes, apartments, group living arrangements, etc.) based on survey responses. To construct the desired variables from this data source, it was necessary to construct filing (or potential filing) units out of the household residents for each year of the survey, assess whether a tax return was required to be filed, and construct measures of income, filing status, and other relevant factors at the level of the filing unit before aggregating to the state level.

Once the data were collected from the various sources, the data were processed, standardized, and validated to help ensure their accuracy. Ultimately, we created a comprehensive analysis database that included detailed state-level longitudinal data on taxpayer characteristics, income tax reporting behavior, IRS service and enforcement activities, and other factors. Where feasible, annual state-level observations were collected for the entire period from 1982 through 2009. However, in some cases variables could be measured only at a national level or only over a shorter time span. For example, variables available only at a national level include several of the political science variables, such as the ratio of Democrats to Republicans in the Senate and House and the average political ideology score of members in Congress, economic variables, including the GDP deflator and CPI, and finally certain IRS variables including web data and complexity measures. Variables that are available only over a shorter time span, either due to lack of data availability or a change in the data collection methodology, include burden, service activity measures, variables from the Bureau of Labor Statistics (BLS), income and offsets data, and criminal investigation measures.

Model Development

Using our robust database, we conducted a preliminary econometric analysis to explore its suitability as a source for predicting taxpayer reporting behavior. We followed a logical analytic progression beginning with specifications similar to previous studies and gradually introducing a number of methodological refinements to incorporate new variables, explore alternative functional forms, and test predictive performance. The focus in this paper is on the reporting of overall income and offsets on the tax return. Thus, our current study not only extends previous efforts for a longer time-frame, but it also integrates a richer set of variables and incorporates innovations to the econometric methodology. Ultimately, the results of our analysis provide a preliminary assessment of the feasibility of using state-level panel data to predict taxpayer reporting behavior.

Our econometric methodology builds on the prior work of Dubin et al. (1990), Dubin (2007), and especially Plumley (1996). As in those studies, we employ panel data regression techniques to explain the aggregate reporting behavior of taxpayers across different states and time periods as a function of various IRS activities and other relevant behavioral determinants.

In general terms our econometric specification is as follows:

$$Y_{it} = \alpha_i + \gamma_t + \beta_A^{'} A_{it} + \beta_O^{'} O_{it} + \varepsilon_{it} ,$$

where Y represents a measure of reporting behavior (such as total income reported, total offsets reported, net income reported, or income reported for a specified line item), A represents a set of IRS activities (including both enforcement and service activities), and O represents a set of other relevant measured determinants of reporting behavior. The subscripts "i" and "t" represent individual states and years, respectively, reflecting our objective of explaining the variation in reporting behavior across both states and time. In the above specification, the parameters β_A and β_O represent coefficients to be estimated. The term ε_{it} is an error term that is meant to capture the net impact of unobserved factors across states and over time on state-level reporting behavior. Finally, the terms α_i and γ_t represent possible sources of state-specific and year-specific heterogeneity. More specifically, α_i represents unobserved time-invariant differences across states and that drive inter-state differences in reporting behavior, while γ_t represents unobserved state-invariant differences across years that drive inter-temporal differences in reporting behavior.

Following Plumley (1996), we specified two alternative definitions of total income: (1) an "A" version that excluded income items that were subject to substantial changes in reporting requirements over the estimation period; and (2) a "B" version that included all taxable income sources, regardless of changes in reporting requirements. A comparable pair of measures for total offsets was also developed. We found that the levels of income and offsets were relatively steady over time under the more restrictive "A" definition, but tended to be somewhat more variable under the "B" definition. Our analysis for this paper focused on the "A" definition.

Fixed vs. Random Effects

The two most common approaches to modeling heterogeneity in panel data are fixed effects and random effects. In the context of the state-specific heterogeneity term α_i in our above specification, a fixed effects specification treats this term as a state-specific constant term in the analysis. In contrast, a random effects specification treats the value of α_i for each state as a random draw from a probability distribution. An advantage of the fixed effects specification is that it produces consistent estimates of the parameters of the model even when the α_i terms are correlated with one or more of the explanatory variables in the model. However, if these terms are not correlated with any of the explanatory variables, the random effects specification produces more efficient (precise) estimates; intuitively, the random effects specification exhausts fewer degrees of freedom, because it is not necessary to estimate the value of α_i ("nuisance parameter") for each state as one does with the fixed effects specification. The fixed effects specification also yields only conditional predictions, in the sense that it is limited to predicting observations that come from units for which a fixed effect has been estimated. However, as Plumley (1996) points out, since the units in our case are states and essentially all states are included in our analysis,[2] this is not a meaningful limitation for our application. Like Plumley (1996), we tend to favor the fixed effects approach for this study as it produces consistent estimates under a wider range of circumstances than the random effects approach. However, we perform some comparisons with the random effects approach to see how sensitive the findings are to the choice of method.

One can also apply a fixed or random effects specification for the time-specific heterogeneity term, in which case one has what is known as a "two-way" fixed or random effects specification. An alternative approach we employ in much of our analysis is to model the term γ_t using one or more time trend terms; for instance:

$$\gamma_t = \gamma_1 t + \gamma_2 t^2 .$$

In this example, time-specific heterogeneity would be modeled using a quadratic trend.

Endogeneity

Both Plumley (1996) and Dubin (2007) recognize that the audit rate is likely to be an endogenous explanatory variable. To account for this, they employ an instrumental variables approach. We follow Plumley in using measures involving state level measures of direct examination time as instruments; specifically, our instruments are the direct examination time percentage (the share of examiners' time directly devoted to examination activities) and the lagged value of the average direct examination time. For our fixed effects specifications, we employ the standard two-stage least squares approach to estimation. In our random effects specifications, we employ the instrumental variables approach proposed by Balestra and Varadharajan-Krishnakumar (1987).

Another explanatory variable that is likely to be endogenous in our model is the combined state and federal marginal tax rate. Owing to the graduated federal (and in some cases, state) tax structure, the state level marginal tax rate will tend to be lower when state level income reporting is low. In our analysis, we experiment with using the combined state and federal marginal tax rate based on a fixed national measure of the income distribution as an instrument. We find that we get extremely similar results when we directly substitute this instrument for the endogenous measure in our analysis. Since the latter approach simplifies prediction, we use it in our prediction exercises.

Other Statistical Issues

Our work goes beyond the previous studies to address a host of statistical issues, including the use of: specifications with ratio dependent and explanatory variables versus alternative functional forms; short versus long panels; and year dummies versus trend terms. For the most part, many of our results are reasonably robust against these alternatives. For instance, we generally obtained qualitatively similar parameter estimates (in terms of coefficient signs and statistical significance) when we substituted alternative functional forms for the base-case ratio specifications. Details of these analyses are presented later in this paper.

Predictive Accuracy

Since an important objective of this study is to evaluate the potential of our alternative specifications to forecast future reporting behavior, we have developed two alternative methodologies for measuring forecasting performance. The first is based on a "leave-one-out" prediction methodology in which one year of data at a time is left out of the estimation sample and the resulting parameter estimates are then used to predict reporting behavior within each state for the excluded year. Successively leaving out each year from the estimation sample produces a set of out-of-sample predictions of reporting behavior for each state and year, which can then be compared against actual reporting behavior. Under our second methodology, we exclude the last four years of the data sample from estimation and then use the resulting parameter estimates to forecast reporting behavior in each of these four years. A comparison against the actual reporting behavior provides an assessment of forecasting performance one, two, three, and four years into the future.

For both our leave-one-out and step-ahead forecasting approaches, we focus on two alternative measures of out-of-sample predictive performance. The first is the mean absolute deviation of the out-of-sample prediction of reported income in each state and time period from the true value of reported income. The second is the root mean-squared error (i.e., the square root of the average squared deviation of the out-of-sample prediction from the actual value). Both of these measures are normalized by dividing them by the average value of reported income over all states and time periods. We refer to the first measure as the "absolute deviation as a percentage of income". The second measure is known in the statistics literature as the "coefficient of variation of root mean-squared error".

A limitation of modeling the time-specific heterogeneity term using fixed effects is that the value of the fixed effect would not be known for years outside of the sample period, which makes forecasting difficult. We therefore employ trend terms rather than yearly fixed effects in much of our analysis. However, a comparison of our results based on our longer panel analyses indicates that certain parameter estimates (notably, the audit rate coefficient) are sensitive to whether yearly fixed effects or trend terms are employed. To investigate the impact of this choice on predictive performance, we have developed an econometric approach to forecasting with yearly fixed effects. Under this approach, we predict the value of the fixed effects for years outside of the sample period based on the estimated sample period fixed effects. We use a Box-Jenkins time series approach (autoregressive integrated moving average, or ARIMA, analysis) to model the fixed effects. Results of our analysis indicate that an autoregressive process of order 2 provides a reasonable fit to the data in the specification we have investigated.

Another complication of our analysis for prediction purposes is the presence of endogenous explanatory variables. Consider a fixed effects specification of the following form:

$$Y_{it} = \alpha_i + \gamma_t + \beta_A A_{it} + \beta_O^{'} O_{it} + \varepsilon_{it} \ ,$$

where the variable A represents the audit rate – an endogenous explanatory variable. We can consistently estimate the parameters of this model using an instrumental variables approach. Suppose that we then substitute the predicted values of the coefficients in for the actual values and attempt to predict Y as:

$$\hat{Y}_{it} = \hat{\alpha}_i + \hat{\gamma}_t + \hat{\beta}_A A_{it} + \hat{\beta}_O^{'} O_{it} \ .$$

In general, this will not be a consistent predictor of Y, because the error term ε will be correlated with A. Consequently, the conditional expectation of $(Y - \hat{Y})$ given O and A will (asymptotically) converge to the value $E(\varepsilon \mid A)$ the value of the conditional expectation of the error term given the audit rate A. Since ε and A are correlated, this expectation will not be equal to zero. To address this issue, we employ a two-stage approach to prediction motivated by the Durbin-Wu-Hausman specification test for endogeneity. In the first stage, we regress the audit rate against all of the explanatory variables of the model as well as the instruments (just as in the first stage of two-stage least squares estimation). We obtain the residual (u) from this regression. In the second stage, we estimate the following regression specification:

$$Y_{it} = \alpha_i + \gamma_t + \beta_A A_{it} + \beta_O^{'} O_{it} + \lambda u_{it} + \varepsilon_{it} \ .$$

Under the Durbin-Wu-Hausman test, one performs a t-test of whether the coefficient λ is equal to zero. The intuition for this test is that this extra term involving the residual u accounts for the correlation between ε and A, so that if λ =0, there is no correlation and, hence, A is exogenous. Although we have not seen this specification used in the econometric literature for purposes of prediction, it can also serve this function. In particular, this extra term involving the residual u directly accounts for the conditional expectation of ε given A that was left out of the above prediction formula and was the source of inconsistent estimation.

Estimation Results

We have employed a systematic approach to estimation to explore the sensitivity of our methodology to the choice of time period, the selection of explanatory variables, the specification of functional forms, and the use of fixed vs. random effects. A comparison of the results provides evidence of the degree to which the methodology is robust to different modeling assumptions and yields some insights about productive areas for further data collection and modeling refinements.

Our preliminary econometric models explore the reporting of a broad measure of the overall total amount of income reported on tax returns before any statutory adjustments or deductions. We first present the estimation results for our base specification and extensions for our model of total income reporting. We then discuss the predictive performance of selected specifications.

Base Specification and Extensions for Total Income Reporting

We begin by specifying the model of income reporting presented by Plumley (based on his "A" definition of total income) using the same time period (1982-1991). Consistent with his approach, we have employed a limited definition of income that controls for some of the changes in federal income reporting requirements over time. In addition, we have included certain control variables in our analysis to account for various changes in federal tax laws, such as the forms of income that are excluded from taxation, the amount allowed for dependent exemptions, various features of the Tax Reform Act of 1986 (captured through a dummy variable), and certain other tax changes (captured through either trend terms or yearly fixed effects).[3]

In his model, Plumley employed a two-way fixed effects model (state and year effects) to explain state level income reporting on required returns (returns that were legally required to be filed). We examine how the results are impacted by substituting trend terms for the year effects. We then explore the sensitivity of the results to using updated measures for some variables, excluding certain variables that we were unable to update for future years, and including some new or substitute explanatory variables. Next, we extend the analysis to different time periods and examine the role of some additional explanatory variables.

We have observed that both Dubin (2007) and Plumley (1996) have relied extensively on ratio variables in their analyses. As summarized by Wiseman (2009), the use of ratio measures in regression analysis is controversial, and there is a growing literature demonstrating that such measures can sometimes lead to spurious and inconsistent findings. We have therefore estimated some alternative specifications that do not rely as heavily on ratios. For instance, we have investigated specifications in which the natural log of reported income is regressed against the natural log of personal income and other explanatory variables rather than using the ratio of reported income to personal income as the dependent variable as is done in Plumley's study. We have also estimated specifications in which many of the ratio explanatory variables have been replaced by variables that separately account for their numerators and denominators.

Table 1 below provides a preliminary sensitivity analysis of the model of income reporting presented by Plumley.[4] The first column includes the results for his original specification. The dependent variable in this specification is the ratio of income reported on returns that were required to be filed to total personal income. In the second column, various modifications have been made, including dropping his information returns matching (lnirp), criminal investigations (lncid), taxpayer service calls (tps_callspc), and IRS return preparation variables (tps_retpreppc) for which we do not have updated measures for subsequent years. Also in the second column, trend terms have replaced the yearly fixed effects; the marginal tax rate variables (mtr15k and mtr57k) have been replaced by a combined state-federal measure of the marginal tax rate (which has been instrumented); a broader measures of the value of dependent exemptions has replaced the measure of the value of child exemptions (childexemptspct); and updated versions of certain other variables (such as soleprops—the percentage of sole proprietors) have been introduced. Overall, the results are not very sensitive to these changes, although the estimates of the coefficients of soleprops and lnaud (the natural log of one plus the audit rate) have become less precise.

In the original Plumley specification, the audit start rate was employed as an explanatory variable. In column 3, the audit close rate was substituted.[5] This has only a very modest impact on the results. In column 4, the specification in column 3 is estimated using random effects rather than state-level fixed effects. The results are quite comparable.

The first column in Table 2 repeats the information in column 3 of the previous table for the case in which the audit close rate is employed. The second column extends the original time period (1982-1991) to a longer time span (1982-2007). While many of the coefficients have the same signs and similar levels of precision in the longer panel, there are some noteworthy exceptions. In particular, the coefficients of the audit rate and the marginal tax rate change signs and become significant. In the case of the marginal tax rate, the new negative coefficient is intuitive, suggesting that high marginal tax rates lead to less compliance. However, the new negative coefficient on the audit rate is counter-intuitive. One would expect, all else equal, that a higher audit rate would yield relatively greater (not less) compliance. In the third column, we have included some additional trend terms in our specification. This does not substantively alter the results. In column 4, we apply random effects estimation to the specification from column 3. This also has only a modest impact on the results. Finally, in column 5, we employ a two-way fixed effects specification that includes year dummies rather than trend terms. This specification change does have an important impact on the results. In particular, the coefficient of the audit rate now becomes positive and significant. Apparently, the year dummies are able to capture certain state-invariant changes in taxpayer circumstances that the trend terms cannot. We have performed a Wald test of the joint significance of the year dummies in our specification, and the evidence strongly supports the alternative hypothesis that the year dummies are jointly significant explanatory variables. We later examine whether the inclusion of year dummies translates into an improved forecasting performance over the use of trend terms.

In Table 3, we experiment with some additional variables not included in the original Plumley specification. In column 1, the specification includes explanatory variables describing the percentages of potential returns for which the primary taxpayer has some college education, is male, and is a homeowner. In addition, population density and the Gini coefficient based on CPS family income (a measure of income inequality within the state) are included as explanatory variables. Only the population density is found to be significant over the 1982-1991 estimation period. In the second column, the estimation period is extended to 2004. With these additional data points, all of the new explanatory variables are found to be statistically significant. However, as with Table 2, the coefficient of the audit rate becomes negative and significant when the time period is extended. In column 3 of the table, the rate of criminal sentences for tax evasion and money laundering is included as an explanatory variable for the 1988-2004 period. This variable is not found to have a significant impact on reporting behavior. In the fourth column, year dummies are employed rather than trend terms. The criminal sentence rate remains insignificant in this specification. However, consistent with previous findings, the audit rate coefficient becomes positive when year dummies are employed (however, the estimate is statistically insignificant). We have again employed a Wald test of the joint significance of the year dummies and the evidence again strongly supports the alternative hypothesis that these variables are jointly significant.

TABLE 1. Model of Income Reporting Presented by Plumley (1996) and Some Variations

	(1)	(2)	(3)	(4)
	Original Specification	Various Changes	Audit Close Rate	Random Effects
Lnaud	11.259 (1.52)	9.358 (0.82)	9.298 (1.29)	7.075 (0.96)
filingrate	0.302 (5.35)**	0.300 (4.75)**	0.308 (5.65)**	0.314 (5.62)**
fthresholdpct	0.935 (2.39)*	1.286 (5.21)**	1.111 (4.28)**	0.986 (3.89)**
mtr15k	1.292 (1.18)			
mtr57k	-1.421 (0.68)			
childexemptspct	1.582 (1.90)			
Lnburden	8.489 (1.18)	7.924 (3.27)**	7.039 (2.72)**	6.710 (2.62)**
Soleprops	2.635 (2.78)**	0.702 (1.41)	0.666 (1.50)	0.595 (1.40)
soleproptfs	-0.056 (2.96)**	-0.020 (1.97)*	-0.018 (2.10)*	-0.018 (2.13)*
Paidprep	-0.124 (2.89)**	-0.106 (2.25)*	-0.116 (3.34)**	-0.133 (3.99)**
Lnirp	-9.160 (1.63)			
Lncid	1.122 (3.17)**			
tps_callspc	-0.006 (1.78)			
tps_retpreppc	0.055 (0.69)			
Singles	0.114 (0.58)	0.240 (2.55)*	0.251 (3.09)**	0.273 (3.35)**
under30	-0.099 (1.06)	-0.020 (0.22)	0.038 (0.42)	0.002 (0.03)
over64	-0.060 (0.56)	0.021 (0.19)	0.084 (0.69)	0.033 (0.28)
Pcbirths	0.725 (2.99)**	0.659 (3.60)**	0.809 (4.32)**	0.609 (3.65)**
exclincomepct	-0.502 (1.21)	-0.815 (1.53)	-0.583 (1.50)	-0.673 (1.75)
unemplrate	-0.473 (3.03)**	-0.485 (2.44)*	-0.448 (3.29)**	-0.400 (2.96)**
Trend		0.739 (1.15)	0.552 (1.51)	0.268 (0.73)
tra86dum		-6.142 (3.45)**	-9.384 (4.50)**	-8.381 (4.33)**
Tratrend		1.211 (4.80)**	1.288 (5.20)**	1.382 (5.56)**
depamountpct		0.418 (0.59)	1.204 (1.60)	0.524 (0.76)
c_marg		71.440 (4.44)**	15.070 (0.62)	11.798 (0.52)
Constant	10.459 (0.09)	-67.978 (3.32)**	-51.611 (2.22)*	-39.961 (1.72)
Observations	490	490	490	490

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

TABLE 2. Results of Estimation of Model Using a Longer Panel

	(1)	(2)	(3)	(4)	(5)
	Audit Close Rate	Longer Sample	Additional Trend Terms	Random Effects	Year Dummies
lnaudnw	9.298	-8.134	-7.995	-6.982	8.789
	(1.29)	(4.40)**	(4.51)**	(3.73)**	(2.36)*
filingrate	0.308	0.351	0.358	0.323	0.310
	(5.65)**	(10.62)**	(11.05)**	(9.73)**	(10.21)**
fthresholdpct	1.111	0.249	0.595	0.283	0.718
	(4.28)**	(1.38)	(3.36)**	(1.74)	(3.94)**
c_marg	15.070	-67.941	-3.899	-21.186	-12.561
	(0.62)	(6.52)**	(0.31)	(2.10)*	(0.65)
depamountpct	1.204	1.957	1.285	0.677	1.488
	(1.60)	(4.22)**	(2.95)**	(1.60)	(3.81)**
lnburden	7.039	4.887	5.298	4.860	5.153
	(2.72)**	(5.51)**	(6.40)**	(5.80)**	(6.03)**
soleprops	0.666	-0.315	-0.629	-0.500	0.525
	(1.50)	(0.97)	(2.03)*	(1.78)	(1.78)
soleproptfs	-0.018	0.008	0.012	0.007	-0.013
	(2.10)*	(1.33)	(2.21)*	(1.48)	(2.33)*
paidprep	-0.116	-0.070	-0.081	-0.093	-0.002
	(3.34)**	(2.69)**	(3.37)**	(4.10)**	(0.11)
singles	0.251	0.224	0.234	0.224	0.163
	(3.09)**	(4.04)**	(4.48)**	(4.24)**	(3.47)**
under30	0.038	-0.012	-0.105	-0.183	0.029
	(0.42)	(0.22)	(1.93)	(3.41)**	(0.56)
over64	0.084	0.007	-0.071	-0.142	0.058
	(0.69)	(0.10)	(1.11)	(2.22)*	(0.96)
pcbirths	0.809	0.847	0.613	0.441	0.337
	(4.32)**	(7.09)**	(5.40)**	(4.29)**	(3.08)**
exclincomepct	-0.583	-0.828	-0.539	-0.767	-0.726
	(1.50)	(3.00)**	(2.09)*	(3.00)**	(3.08)**
unemplrate	-0.448	-0.338	-0.420	-0.357	-0.241
	(3.29)**	(3.82)**	(4.91)**	(4.08)**	(2.70)**
trend	0.552	-0.953	-1.036	-1.271	
	(1.51)	(4.68)**	(5.47)**	(6.93)**	
tra86dum	-9.384	-6.291	3.112	3.555	
	(4.50)**	(5.58)**	(2.05)*	(2.57)*	
tratrend	1.288	0.731	-0.333	-0.125	
	(5.20)**	(4.27)**	(1.62)	(0.63)	
dum91			1.795	1.827	
			(3.87)**	(3.75)**	
Constant	-51.611	10.515	-7.172	17.365	-22.213
	(2.22)*	(0.88)	(0.63)	(1.65)	(1.84)
dum92			1.369	1.323	
			(3.06)**	(2.84)**	
trendsq			0.035	0.033	
			(8.21)**	(8.51)**	
Observations	490	1274	1274	1274	1274

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

TABLE 3. Variables Added to the Model of Income Reporting Presented by Plumley (1996)

	(1)	(2)	(3)	(4)
	New Variables 1982-1991	New Variables 1982-2004	Cidsentrate 1988-2004	Year Dummies
lnaudnw	9.716 (1.29)	-13.751 (5.89)**	-9.996 (4.32)**	3.871 (1.14)
filingrate	0.292 (4.71)**	0.389 (11.26)**	0.414 (10.46)**	0.409 (11.64)**
fthresholdpct	1.052 (3.64)**	0.967 (4.65)**	1.026 (3.79)**	1.647 (6.91)**
c_marg	-25.591 (0.72)	-20.222 (1.47)	47.021 (2.30)*	37.828 (1.28)
depamountpct	1.297 (1.68)	0.695 (1.40)	1.356 (2.68)**	1.762 (4.15)**
lnburden	8.724 (2.99)**	2.896 (3.01)**	2.715 (3.02)**	2.730 (3.47)**
soleprops	0.481 (1.04)	-0.311 (0.94)	0.097 (0.21)	0.617 (1.59)
soleproptfs	-0.013 (1.53)	0.005 (0.84)	-0.002 (0.23)	-0.013 (1.81)
paidprep	-0.150 (4.09)**	-0.074 (2.74)**	0.029 (0.82)	0.036 (1.20)
singles	0.196 (1.91)	0.155 (2.17)*	0.202 (2.78)**	0.130 (2.11)*
under30	0.024 (0.24)	-0.219 (3.46)**	-0.048 (0.74)	0.070 (1.21)
over64	0.082 (0.65)	-0.083 (1.12)	0.011 (0.14)	0.061 (0.93)
pcbirths	0.431 (1.97)*	0.413 (3.22)**	0.457 (2.24)*	0.050 (0.29)
exclincomepct	-0.419 (1.05)	-1.000 (3.51)**	-0.950 (3.17)**	-0.236 (0.91)
unemplrate	-0.791 (3.83)**	-0.332 (3.44)**	-0.559 (4.58)**	-0.516 (4.23)**
trend	-1.733 (1.64)	-1.386 (6.39)**	-1.961 (6.99)**	
tra86dum	3.001 (0.40)	-2.061 (1.03)		
tratrend	-1.276 (0.88)	0.388 (1.50)		
trendsq	0.295 (1.85)	0.007 (0.97)	0.043 (4.37)**	
collegepct	-0.031 (0.41)	0.131 (2.68)**	0.105 (2.16)*	0.057 (1.32)
malepct	0.055 (0.51)	0.206 (2.69)**	0.124 (1.56)	0.105 (1.58)
homeownerpct	-0.081 (0.90)	-0.158 (2.73)**	-0.092 (1.58)	-0.052 (1.04)
popdensity	0.105 (3.42)**	0.044 (5.22)**	0.041 (4.23)**	0.020 (2.25)*
gini_faminc	-1.773 (0.10)	29.430 (3.17)**	29.943 (3.13)**	1.342 (0.15)
dum91		0.913 (1.77)	1.560 (3.27)**	
dum92		0.621 (1.26)	1.516 (3.36)**	
cidsentrate			-0.041 (0.13)	-0.322 (1.26)
Constant	-33.145 (1.31)	7.290 (0.54)	-28.571 (1.85)	-46.565 (3.11)**
Observations	490	1127	833	833

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

In Table 4, we experiment with including some more new explanatory variables. These include an annual national measure of hours by IRS personnel in taxpayer-facing service occupations and annual national measure of tax return complexity. The state level measure of the rate of criminal sentences for tax evasion and money laundering that was introduced in Table 4 is also included. The sample period extends from 1996-2005 as this is the period for which we have measures of these three new variables. The first column of Table 4 presents a base level specification for that sample period that excludes the new variables, while column 2 includes them. The results for taxpayer-facing service hours and complexity are somewhat unexpected, suggesting that more service hours leads to lower taxpayer reporting and that greater tax system complexity leads to higher reporting. The estimated impact of criminal sentences on reporting behavior is positive for this time period, although it is not very precisely estimated. In column 3, the criminal sentence rate variable is dropped, allowing us to include an additional two years in the sample period. The signs of the estimated coefficients on service hours and complexity are unchanged, although their magnitude has been reduced. In column 4, the trend term has been replaced by year dummies in the base specification without any of the new variables. As observed in previous specifications, the use of year dummies results in a change in the signs of the estimated audit rate and marginal tax rate coefficients. In column 5, the new criminal sentence has been included in the two-way fixed effects specification. It was not possible to include the service hours and complexity variables in this specification, because these national level estimates are perfectly collinear with the year dummies. With the two-way fixed effects specification, the sign of the criminal sentence variable has reversed, although the estimate is statistically insignificant.

In Table 5, we experiment with a state level measure of attempted calls to the IRS help line. This is similar to the measure of calls handled by taxpayer service (tps_callspc) that was used by Plumley for the period from 1982-1991. Our measure is available for the 2002-2007 time period. The first column of the Table provides a base specification for this period that excludes the attempted calls variable, while the second column includes this variable. The results indicate that telephone assistance is positively associated with income reporting, although the coefficient estimate is not very precise. In columns 3 and 4, we repeat this exercise, this time using a two-way random effects specification rather than including a time trend. In this specification, the coefficient of the calls attempted variable becomes negative, but insignificant. Also observe that the audit rate and marginal tax rate coefficients have increased substantially compared to the earlier specification involving the time trend. It appears that the results for this time period (2002-2007) are rather fragile.

We have also experimented with functional forms. For instance, we have estimated variants of our specifications in Tables 4 and 5 in which the dependent variable is the log of income reported rather than the ratio of income reported to total personal income. In one variant, the natural log of total personal income is included as an additional explanatory variable and the other variables are the same as in the previous specification. In the other variant, many of the ratio variables are eliminated. In their place are separate measures of the numerators and denominators of these ratios. We have experimented with year dummies and trend terms as alternatives in these specifications. As with the results in this section, a Wald test supports the joint significance of the year dummies. The estimation results are qualitatively similar to the results presented in Tables 4 and 5.

Predictive Performance of Models of Total Income Reporting

An important objective of this project is to develop a preliminary assessment of the predictive capability of the panel data modeling approach. We begin by evaluating how well alternative econometric specifications of total income reporting forecast out of sample when they are based on the same 1982-1991 period employed in the Plumley study. We then explore how the forecasting performance changes when the models are estimated over a longer time span.

The first two columns of Table 6 below respectively present results from a two-way fixed effects version and a one way fixed effects with trends version of the parsimonious specification of income reporting behavior provided earlier in the first column of Table 2. The results for the two specifications are quite similar. In these specifications, both the natural log of the audit close rate and the combined state-federal marginal tax rate are treated using instrumental variables. Since having two instrumented variables complicates the prediction process to some extent, we experiment in column 3 with directly using the instrument for the state-federal marginal tax rate in the specification rather than as an instrumental variable. This variable represents the computed combined state-federal marginal tax rate based on a fixed national distribution of income in 1995. The results indicate that this approach yields very similar estimates to those shown for the instrumental variables specification in column 2. In column 4, we extend the specification in column 3 to include some additional explanatory variables that were not included in Plumley's original specification. Of these additional variables, only population density proves to be statistically significant.

TABLE 4. Experimentation with Some Additional New Explanatory Variables

	(1)	(2)	(3)	(4)	(5)
	Base Model 1996-2005	New Variables 1996-2005	Longer Period (no cidsentrate) 1996-2007	Base Year Dummies 1996-2005	Cidsentrate Year Dummies 1996-2005
lnaudnw	-9.938 (3.08)**	-12.306 (3.11)**	-2.209 (0.70)	18.132 (2.43)*	18.223 (2.44)*
filingrate	0.407 (7.84)**	0.408 (7.59)**	0.310 (6.39)**	0.364 (7.21)**	0.369 (7.29)**
fthresholdpct	2.240 (7.07)**	2.163 (6.74)**	0.945 (3.82)**	2.084 (6.86)**	2.116 (6.94)**
c_marg_95_fixed	89.795 (5.05)**	76.055 (3.88)**	60.857 (3.42)**	-81.169 (2.88)**	-84.842 (2.99)**
depamountpct	0.970 (1.51)	1.167 (1.76)	0.841 (1.42)	-0.268 (0.42)	-0.338 (0.53)
lnburden	1.356 (1.54)	0.603 (0.65)	1.671 (1.90)	2.321 (2.67)**	2.314 (2.67)**
soleprops	-0.427 (0.47)	-0.639 (0.68)	0.163 (0.21)	1.009 (1.07)	0.932 (0.99)
soleproptfs	0.009 (0.57)	0.012 (0.75)	-0.004 (0.33)	-0.020 (1.21)	-0.019 (1.14)
paidprep	0.114 (2.51)*	0.144 (3.08)**	0.038 (0.92)	0.135 (3.04)**	0.134 (3.03)**
singles	0.234 (2.77)**	0.249 (2.92)**	0.181 (2.38)*	0.187 (2.40)*	0.187 (2.40)*
under30	0.135 (1.82)	0.147 (1.97)*	0.119 (1.77)	0.167 (2.44)*	0.166 (2.43)*
over64	0.000 (0.00)	0.018 (0.21)	-0.001 (0.02)	-0.020 (0.27)	-0.025 (0.33)
pcbirths	0.281 (0.86)	0.548 (1.60)	0.153 (0.57)	0.433 (1.41)	0.398 (1.29)
exclincomepct	0.002 (0.00)	-0.085 (0.21)	-0.681 (2.10)*	-0.556 (1.63)	-0.588 (1.71)
unemplrate	-1.183 (8.13)**	-1.109 (7.39)**	-1.009 (7.79)**	-0.782 (5.09)**	-0.765 (4.96)**
collegepct	0.058 (1.05)	0.068 (1.23)	0.116 (2.37)*	0.018 (0.36)	0.014 (0.29)
malepct	0.014 (0.15)	0.013 (0.14)	0.049 (0.55)	0.036 (0.42)	0.036 (0.43)
homeownerpct	-0.026 (0.36)	-0.000 (0.01)	0.006 (0.09)	0.065 (0.99)	0.068 (1.03)
popdensity	0.004 (0.36)	0.009 (0.78)	-0.007 (0.72)	-0.008 (0.68)	-0.008 (0.72)
trend	0.139 (0.86)	-0.559 (1.70)	-0.013 (0.04)		
hoursrvrate		-19.377 (3.02)**	-8.777 (1.75)		
complexity		13.540 (3.15)**	7.056 (2.28)*		
cidsentrate		0.723 (1.77)			-0.338 (0.93)
Constant	-70.995 (3.80)**	-69.674 (3.68)**	-31.680 (1.88)	-31.592 (1.83)	-30.355 (1.75)
Observations	490	490	588	490	490

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

TABLE 5. Inclusion of State-level Measure of Attempted Calls to the IRS Help Line

	(1)	(2)	(3)	(4)
	Base 2002-2007	callattemptpct	Base 2002-2007, Year Dummies	Year Dummies, callattemptpct
lnaudnw	18.554	16.233	67.785	71.291
	(1.95)	(1.71)	(1.77)	(1.77)
filingrate	0.218	0.233	0.191	0.189
	(2.77)**	(2.97)**	(1.57)	(1.54)
fthresholdpct	0.138	0.245	0.090	0.083
	(0.40)	(0.71)	(0.11)	(0.10)
c_marg_95_fixed	19.546	-12.034	-252.328	-257.324
	(0.92)	(0.42)	(2.70)**	(2.66)**
depamountpct	1.238	1.301	0.541	0.455
	(1.30)	(1.38)	(0.43)	(0.35)
lnburden	2.751	2.908	4.192	4.220
	(2.91)**	(3.11)**	(2.94)**	(2.91)**
soleprops	2.180	1.992	-0.788	-0.832
	(1.67)	(1.54)	(0.36)	(0.37)
soleproptfs	-0.040	-0.037	0.012	0.013
	(1.82)	(1.71)	(0.33)	(0.35)
paidprep	-0.202	-0.166	-0.087	-0.095
	(3.17)**	(2.49)*	(0.96)	(1.02)
singles	-0.008	-0.014	-0.033	-0.029
	(0.08)	(0.14)	(0.24)	(0.21)
under30	-0.071	-0.079	0.027	0.029
	(0.73)	(0.81)	(0.21)	(0.22)
over64	-0.041	-0.038	-0.107	-0.115
	(0.41)	(0.38)	(0.79)	(0.81)
pcbirths	-0.159	-0.128	1.078	1.125
	(0.39)	(0.32)	(1.50)	(1.51)
exclincomepct	0.097	0.083	-0.660	-0.677
	(0.23)	(0.21)	(0.89)	(0.90)
unemplrate	-0.431	-0.420	-0.271	-0.266
	(1.90)	(1.88)	(0.74)	(0.72)
collegepct	0.078	0.057	0.056	0.067
	(1.17)	(0.85)	(0.58)	(0.64)
malepct	0.503	0.522	0.464	0.454
	(3.48)**	(3.65)**	(2.41)*	(2.29)*
homeownerpct	0.008	-0.007	0.071	0.075
	(0.10)	(0.08)	(0.65)	(0.66)
popdensity	-0.144	-0.141	-0.149	-0.151
	(4.33)**	(4.29)**	(3.37)**	(3.30)**
trend	0.893	1.011		
	(4.29)**	(4.61)**		
callattemptpct		0.131		-0.088
		(1.58)		(0.46)
Constant	11.707	10.871	75.980	78.444
	(0.51)	(0.48)	(2.02)*	(1.99)*
Observations	294	294	294	294

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

TABLE 6. Two-way Fixed Effects and a One-way Fixed Effects with Trends

	(1)	(2)	(3)	(4)
	Base Year Dummies	Base Trends	Base Trends c_marg_95_fixed	Base Trends c_marg_95_fixed Additional Variables
lnaudnw	9.728 (1.36)	9.298 (1.29)	8.494 (1.23)	6.938 (1.04)
filingrate	0.281 (4.58)**	0.308 (5.65)**	0.315 (5.99)**	0.332 (6.45)**
fthresholdpct	0.888 (2.14)*	1.111 (4.28)**	1.131 (4.27)**	1.247 (4.85)**
c_marg	-22.506 (0.67)	15.070 (0.62)		
depamountpct	1.168 (1.48)	1.204 (1.60)	1.220 (1.62)	1.399 (1.88)
lnburden	10.473 (3.40)**	7.039 (2.72)**	6.827 (2.71)**	6.252 (2.55)*
soleprops	0.403 (0.84)	0.666 (1.50)	0.643 (1.44)	0.705 (1.62)
soleproptfs	-0.013 (1.42)	-0.018 (2.10)*	-0.017 (2.03)*	-0.018 (2.13)*
paidprep	-0.147 (3.99)**	-0.116 (3.34)**	-0.117 (3.36)**	-0.126 (3.75)**
singles	0.270 (3.25)**	0.251 (3.09)**	0.251 (3.09)**	0.206 (2.11)*
under30	0.056 (0.59)	0.038 (0.42)	0.040 (0.44)	0.003 (0.03)
over64	0.060 (0.49)	0.084 (0.69)	0.073 (0.61)	0.094 (0.77)
pcbirths	0.782 (3.63)**	0.809 (4.32)**	0.846 (4.58)**	0.484 (2.38)*
exclincomepct	-0.546 (1.27)	-0.583 (1.50)	-0.548 (1.41)	-0.483 (1.27)
unemplrate	-0.603 (3.47)**	-0.448 (3.29)**	-0.450 (3.26)**	-0.603 (3.99)**
trend		0.552 (1.51)	0.511 (1.45)	0.248 (0.73)
tra86dum		-9.384 (4.50)**	-9.695 (5.08)**	-11.228 (5.75)**
tratrend		1.288 (5.20)**	1.305 (5.21)**	1.463 (5.59)**
c_marg_95_fixed			10.674 (0.51)	16.170 (0.79)
collegepct				-0.019 (0.26)
malepct				0.064 (0.61)
homeownerpct				-0.082 (0.96)
popdensity				0.102 (3.48)**
gini_faminc				-3.727 (0.22)
Constant	-49.943 (2.09)*	-51.611 (2.22)*	-50.076 (2.17)*	-43.105 (1.77)
Observations	490	490	490	490

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

Table 7 presents measures of the predictive performance of the more and less parsimonious specifications presented in columns 3 and 4 of Table 6. These measures are based on the leave-one-out prediction approach described earlier. The results indicate a rather similar forecasting performance across the two specifications. In both cases, the average absolute deviation of the forecast from the true level of income represents a little more than 2 percent of income, and the coefficient of variation of the root mean-squared error (CV of RMSE) is approximately 3.6 percent. As discussed above, the use of ratio specifications such as those presented in Table 6 is rather controversial. We have therefore estimated alternative versions of these models in which the dependent variable has been specified in natural log rather than ratio form and, in one variant, where many of the ratio explanatory variables have been replaced with separate variables representing the numerators and denominators of the ratios. The parameter estimates from these specifications have been fairly comparable in terms of signs and statistical significance. Further, the predictive performance of these specifications have turned out to be quite similar to that of the original specifications based on ratios.

TABLE 7. Leave-One-Out Predictive Performance of Models in Table 6

Specification from Table 6 Column #	Absolute Deviation as a % of Income Reported	CV of RMSE
3	2.12	3.57
4	2.18	3.62

We now explore the forecasting performance of specifications estimated from a longer panel. Table 8 summarizes the estimation results of some selected specifications that have been estimated from data spanning the period from 1982 to 2007 (or 2004 for the specifications that include the Gini coefficient as an explanatory variable). As with Table 6, the first two columns of Table 8 respectively provide a two-way fixed effects specification and a one way fixed effects with trend terms specification of a parsimonious model of income reporting behavior. While many of the parameter estimates are comparable in sign and significance across these two specifications, they do produce conflicting estimates of the coefficients of the natural log of the audit close rate and the combined state-federal marginal tax rate. This discrepancy between the results of the two alternative specifications is consistent with similar findings presented above. In column 3 we verify that directly substituting the instrument for the combined marginal tax rate as an explanatory variable yields comparable results to those presented in column 2. That specification is extended to include some additional explanatory variables in column 4. Finally, column 5 presents the results of estimating the specification in column 4 using a two-way fixed effects specification rather than using trend terms. Once again, the use of two-way fixed effects yields more intuitive coefficient estimates for the audit and marginal tax rate variables.

To investigate whether this translates into improved predictive performance, we have employed the leave-one-out prediction methodology described in Section 4.3 for each of the specifications in columns 3, 4, and 5. In the case of the year dummy specification in column 5, we have used an autoregressive process of order 2 to forecast the value of each year dummy when the corresponding year is left out of the estimation sample. The leave-one-out forecasting results are summarized in Table 9. All of the specifications predict reasonably well out of sample, with an average absolute forecast deviation of less than 3% of income reported and a CV of RMSE of 4.3 to 5.1 percent. Overall, the predictive performance is slightly weaker for the longer panel specifications summarized in Table 9 than for the comparable shorter panel specifications summarized earlier in Table 7. Interestingly, the specification based on year dummies performs slightly less well than those based on trend terms.

We have also used the results of the last two specifications presented in Table 8 to develop one, two, three, and four step-ahead forecasts. These specifications were estimated using a sample period of tax year 1982 through tax year 2000, and the results were then employed to develop state level forecasts of income reported for tax years 2001 through 2004. These forecasts were compared against actual levels of income reported to produce measures of the average absolute forecast deviation as a percentage of income reported and the CV of RMSE. The results are summarized in Table 10 below.

TABLE 8. Results of Estimation Using a Longer Panel

	(1)	(2)	(3)	(4)	(5)
	Base Year Dummies	Base Trends	Base Trends c_marg_95_fixed	Base Trends c_ marg_95_fixed Extra Variables	Year Dummies
lnaudnw	8.789 (2.36)*	-3.791 (2.15)*	-3.890 (2.20)*	-9.124 (3.89)**	4.881 (1.42)
filingrate	0.310 (10.21)**	0.350 (12.13)**	0.350 (12.04)**	0.393 (12.92)**	0.366 (12.22)**
fthresholdpct	0.718 (3.94)**	0.576 (3.57)**	0.552 (3.59)**	1.063 (5.99)**	0.994 (5.51)**
c_marg	-12.561 (0.65)	8.787 (0.75)			
depamountpct	1.488 (3.81)**	1.630 (3.99)**	1.643 (4.00)**	0.736 (1.65)	1.427 (3.46)**
lnburden	5.153 (6.03)**	5.039 (6.50)**	5.030 (6.45)**	3.236 (3.72)**	3.594 (4.15)**
soleprops	0.525 (1.78)	-0.148 (0.52)	-0.188 (0.67)	-0.160 (0.54)	0.589 (2.07)*
soleproptfs	-0.013 (2.33)*	0.003 (0.53)	0.003 (0.71)	0.001 (0.25)	-0.014 (2.71)**
paidprep	-0.002 (0.11)	-0.064 (2.80)**	-0.066 (2.92)**	-0.069 (2.87)**	-0.018 (0.77)
singles	0.163 (3.47)**	0.210 (4.32)**	0.209 (4.28)**	0.135 (2.06)*	0.114 (1.94)
under30	0.029 (0.56)	-0.042 (0.85)	-0.043 (0.86)	-0.214 (3.80)**	-0.077 (1.36)
over64	0.058 (0.96)	-0.007 (0.11)	-0.009 (0.16)	-0.056 (0.81)	0.073 (1.15)
pcbirths	0.337 (3.08)**	0.575 (5.34)**	0.582 (5.28)**	0.349 (2.91)**	0.269 (2.34)*
exclincomepct	-0.726 (3.08)**	-0.513 (2.11)*	-0.512 (2.09)*	-0.855 (3.25)**	-1.012 (4.19)**
unemplrate	-0.241 (2.70)**	-0.392 (5.05)**	-0.396 (5.11)**	-0.361 (4.33)**	-0.186 (2.14)*
trend		-0.453 (2.40)*	-0.488 (2.80)**	-0.969 (4.70)**	
tra86dum		-3.506 (3.35)**	-3.625 (3.51)**	-2.569 (2.46)*	
tratrend		0.626 (4.17)**	0.653 (4.59)**	0.603 (4.05)**	
dum92on		-4.875 (10.34)**	-4.811 (10.47)**	-3.757 (6.47)**	
c_marg_95_fixed			5.826 (0.58)	13.295 (1.26)	-2.461 (0.16)
collegepct				0.140 (3.17)**	0.055 (1.24)
malepct				0.189 (2.69)**	0.182 (2.88)**
homeownerpct				-0.195 (3.70)**	-0.127 (2.62)**
popdensity				0.045 (5.79)**	0.037 (5.20)**
gini_faminc				21.313 (2.46)*	4.573 (0.54)
Constant	-22.213 (1.84)	-17.122 (1.58)	-15.592 (1.50)	-6.228 (0.51)	-17.692 (1.51)
Observations	1274	1274	1274	1127	1127

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

TABLE 9. Leave-One-Out Predictive Performance of Selected Models in Table 8

Specification from Table 8 Column #	Absolute Deviation as a % of Income Reported	CV of RMSE
3	2.47	4.64
4	2.42	4.31
5	2.78	5.10

TABLE 10. Step-Ahead Predictive Performance of Selected Models in Table 8

Specification from Table 8 Column #	2001		2002		2003		2004	
	Absolute Dev'n %	CV of RMSE	Absolute Dev'n %	CV of RMSE	Absolute Dev'n %	CV of RMSE	Absolute Dev'n %	CV of RMSE
4	2.11	3.09	2.78	3.84	4.22	6.06	2.96	4.49
5	2.45	3.90	2.95	4.23	3.28	4.96	2.22	3.92

Generally, the forecasting performance is reasonably strong. As expected, the performance tends to decline to some extent as one predicts further out, although the four-year ahead forecast performance for 2004 is comparable to the one-year ahead performance for 2005 in the specification involving year dummies.

We have also estimated variants of the specifications summarized in Table 8 that rely less on ratio variables. The results for these variants were qualitatively similar to those based on the ratio variables.

Conclusions

We have found that the forecasting performance of our preliminary models of overall income reporting is reasonably strong. This performance is slightly stronger for our shorter panel (1982-1991) than our longer panel (1982-2007), although the performance is reasonably good in both cases.

Overall, the results of our analysis indicate that it is possible to develop reasonably good forecasts of what overall state level income reporting behavior would be in the absence of a major innovation, such as a significant change in service level or quality. However, the lack of a reasonably lengthy time series of high quality state-level measures of IRS service activities limits the potential for our current models to predict how such activities influence reporting behaviors. Fortunately, compiling such data is a high priority for the IRS Service-Compliance Initiative going forward.

References

Balestra, P., and J. Varadharajan-Krishnakumar, "Full information estimations of a system of simultaneous equations with error component structure," *Econometric Theory*, 1987, 3: 223–246.

Dubin, Jeffrey A., "Criminal investigation enforcement activities and taxpayer noncompliance," *Public Finance Review*, 2007. 35: p. 500-529.

________, M. Graetz, and L. Wilde (1990). "The Effect of Audit Rates on the Federal Individual Income Tax, 1977-1986," *National Tax Journal* 43:395-409.

Plumley, Alan H., "The Determinants of Individual Income Tax Compliance: Estimating the Impact of Tax Policy, Enforcement and IRS Responsiveness," *Internal Revenue Service, Publication 1916 (Rev. 11-96),* Washington, DC, 1996.

Wiseman, R.M. "On the Use and Misuse of Ratios in Strategic Management Research" in *Research Methodology in Strategy and Management,* Volume 5, ed. by D.D. Berg and D.J. Ketchen, Jr., pp. 75-110, Bingley, U.K.: Emerald Group Publishing, Ltd., 2009.

Endnotes

[1] RAS Office of Research (IRS), Brian Erard & Associates, and IBM Global Business Services, respectively. The views expressed in this paper are those of the authors, and do not necessarily reflect the views of the Internal Revenue Service.

[2] One exception is Alaska, which is excluded because of compatibility issues resulting from the need for all recipients of Alaska Permanent Fund Dividends—including children—to file federal tax returns. Also, Maryland and DC have been combined.

[3] See the listing of variable definitions provided at the end of the paper.

[4] Our results for this specification differ from Plumley (1996), as we use a more standard approach to instrumental variables estimation. The variable definitions are provided at the end of the paper.

[5] The audit start rate is defined as the number of audits started in a given year, expressed as a percentage of the total number of returns filed in the calendar year before the beginning of the audit. The audit close rate, which is the more standard measure, is defined as the number of audits completed in a given year, expressed as a percentage of the total number of returns filed in the calendar year before the closure. Plumley (1996) theorized that the information about audits that gets "rippled" into the general population at the start of an audit affects people's perception of the *probability* of an audit, while the information communicated when the audit is closed has more to do with the *consequence* of an audit. As a practical matter, since the two audit rates are so highly correlated, they appear to be fairly interchangeable in an analysis such as this, so it makes sense to use the audit closure rate, which is more readily available.

Appendix

Definitions of Included Variables

Variable	Definition
lnaud	Natural log of one plus the audit start rate
filingrate	Returns Filed/ Returns Required
fthresholdpct	Income below filing threshold among all potential returns as a % of PI
mtr15k	Marg. tax rate @ $15K taxable income (weighted by Singles & Marrieds)
mtr57k	Marg. tax rate @ $57K taxable income (weighted by Singles & Marrieds)
childexemptspct	Value of allowed dependent child exemptions/Personal Income
lnburden	Natural log of average burden (in dollars) based on the IMF Population
soleprops	% of Potential Returns having non-farm proprietorship income
soleproptfs	SoleProps x percentage of non-farm employment in Trade, Finance & Service sectors
paidprep	% of Returns Filed prepared by paid practitioner
lnirp	Natural log of information returns matching
lncid	Natural log of criminal investigations
tps_callspc	Taxpayer service calls handled per thousand of population
tps_retpreppc	Returns prepared by taxpayer service calls per thousand of population
singles	% of Potential Returns likely to qualify for Single filing status
under30	% among Potential Returns under age 30
over64	% among Potential Returns over age 64
pcbirths	Number born per thousand of population
exclincomepct	Excluded Income/Personal Income
unemplrate	Unemployment Rate
trend	Trend
tra86dum	TRA86 dummy variable equal to one for years subsequent to 1986
tratrend	Interaction of trend and tra86dum (trend times tra86dum)
depamountpct	Total Value of the Dependent exemption as a percent of personal income
c_marg	Combined Marginal Tax Rate Based on the Actual distribution of Reported Income / IMF Population
lnaudnw	Natural log of the audit close rate
dum91	Dummy variable for 1991
dum92	Dummy variable for 1992
trendsq	Trend squared
collegepct	% among Potential Returns having at least some college
malepct	% of Potential Single & HeadHhld Returns associated with males
homeownerpct	% of Potential Returns associated with homeowners
popdensity	Population density
gini_faminc	Smoothed state-level gini coefficient based on family income
cidsentrate	Total sentenced violations as a percentage of population
c_marg_95_fixed	Combined marginal tax rate based on 1995 national distribution of reported income / IMF population
hoursrvrate	National measure of number of hours worked by IRS employees in taxpayer-facing service occupations
complexity	National measure of the complexity of individual returns based on word counts of IRS individual income tax code
callattemptpct	Total call attempts as a percent age of the overall state population