Tax audit impact on voluntary compliance

Yongzhi Niu

New York State Department of Taxation and Finance

11. May 2010

Online at https://mpra.ub.uni-muenchen.de/22651/
MPRA Paper No. 22651, posted 11. May 2010 19:04 UTC
Tax Audit Impact on Voluntary Compliance

Yongzhi Niu, Ph. D.

New York State Department of Taxation and Finance, Building 8, Room 938,
W.A. Harriman Campus, Albany, N.Y. 12227
(e-mail: bdniu@budget.state.ny.us)

The author would like to thank Dr. Roger Cohen for his valuable comments and suggestions in preparing this paper. Thanks also go to Benjamin Holloway, Kathy Wagner, and Jim Sherman for data preparation assistance and their comments and edits of an earlier draft. The opinions expressed in this paper are those of the author and do not necessarily represent the views of the New York State Department of Taxation and Finance.

Abstract

This study examines the tax audit impact on voluntary compliance. It is different from those in the literature in several ways. First, models were built exclusively for investigating the voluntary compliance behavior shifts after a firm is audited. Second, apart from the theoretical approach and laboratory experiment approach used in the literature, this study applied the difference-in-differences non-experimental approach. Third, historical population data of a New York State economic sector were used in this study instead of experimental data or randomly selected sample data often used in the literature. The results of both Ordinary Least Squares (OLS) and Time Series Cross Section (TSCS) autoregressive modeling methods are presented. The results of both methods suggest that after an audit, a firm would report a higher sales growth rate. The TSCS approach shows that in the year of the audit, a typical firm would report a sales growth rate which is 2.63 percentage points higher than a firm that was not audited. The percentage would decline by a rate of 1/3 each year thereafter. The findings suggest that the audit productivity derived in many research papers, where only the direct audit collections are considered, may be underestimated. The results of this research may provide policy makers with extra incentives to strengthening the audit efforts to generate more revenues.
An interesting question facing tax authorities is whether a tax audit has a positive impact on future voluntary compliance, i.e., after an audit, will a firm change its filing behavior to report higher tax liabilities than it would without the audit?

Tax professionals have opposing answers to this question. Some say that an audit has a positive impact. They argue that after an audit, especially when an audit results in additional tax liabilities, a firm may think that the tax authorities are closely monitoring its activities and may feel that it will be caught again if it attempts to conceal revenue. For this reason, the firm may report higher sales revenue than it would have otherwise reported had it not been audited. Other tax professionals argue the opposite. They say that after an audit, a firm may think that it is less likely to be selected for re-audit in the future, which may provide the firm with incentives to cheat.

Both arguments have merit and both sides can provide evidence to support their stance based on the examination of individual firms. However, we would like to know the overall audit impact, not the impact on a particular firm. This study tries to answer the question in a more comprehensive way. In this study, one segment of the New York State economy, Food Services and Drinking Places (NAICS 722), is investigated to detect possible behavior shifts after firms are audited. More specifically, alternative approaches are explored to quantify the impact of an audit on reported sales, which we think is the best indicator for voluntary compliance.

In the literature, there are generally two distinctive approaches to studying tax compliance. The first approach is based on theoretical grounds; and the second is applying laboratory experiments to analyze the behavior shifts in compliance after simulating tax policy and regulation changes.

In the first approach, the bases of theoretical grounds are not uniform among the researches in the literature. One group of studies is based on the classical microeconomic theories where taxpayers are assumed to be rational and follow the rule of utility maximization in making their tax compliance decisions. Their ultimate goal is to maximize their expected utility under uncertainty (For example, see Allingham and Sandmo 1972). Another group of the studies introduces sociological and psychological factors, such as moral, shame, trust, political power, and game theory, into their theoretical considerations. They hope these factors can explain some
compliance phenomena which the simple utility models could not explain. (For example, see Bernasconi 1998, Kirchler, Hoelzl, and Wahl 2008, and Kirchler 2007). Under this approach, a wide range of interesting factors influencing tax compliance, besides audit, is studied. For example, some researchers propose that taxpayer uncertainty has a positive effect on compliance, (Alm, Jackson, and McKee 1992(2), and Beck and Jung 1989). Another example is the work by Erard and Feinstein (1994), who build a game-theoretic model of tax compliance which challenges the notion that honest taxpayers do not significantly influence most aspects of tax compliance systems.

The second approach, the laboratory experiment approach, has been gaining popularity. People adopt this approach because of the scarcity of real field data. In this approach, experiment participants, either students or real taxpayers, are provided specific information regarding audit, amnesty, or other tax policies. Then they are asked to file tax returns (fake or real tax returns). The tax return data are analyzed to reveal their compliance behavior. For example, Alm and McKee (2006) apply experimental methods to examine the individual compliance responses to a “certain” probability of audit, and conclude that the compliance rate rises if an individual knows he will be audited and the rate falls if he knows he will not be audited. Slemrod, Blumenthal, and Christian (2001) randomly select taxpayers and inform them that their filling will be “closely examined” and found evidence of taxpayers’ behavior changes in response to an increased probability of audit, although the responses are not uniform among different groups of taxpayers. Another experimental research by Alm, McKee, and Beck (1990) finds that the effectiveness of an amnesty program depends on the design of the program and the enforcement efforts post the amnesty. Alm, Jackson, and Mckee (1992(1)) use data from laboratory experiments to estimate the effects on compliance of the major fiscal instruments and conclude that, among others, there is a positive relationship between audit rate and compliance, but they caution about the generalization from the estimates based on the experiments. Another study by Mittone (2006) concludes that early experience of audits in taxpayers’ “tax life” is a more effective way to increase compliance than later audits. Yet another experimental research by Kastlunger, Kirchler, Mittone, and Pitters (2009) also suggests that, although the effectiveness of audits and fines cannot be completely confirmed, early audits in taxpayers’ “tax life” have a positive impact on compliance.
The studies mentioned above touch many aspects of the taxpayers’ compliance behavior; however, none of them deals exclusively with the changes in the voluntary compliance behavior after an audit. This study is different from those in the literature in several ways. First, we built models exclusively for investigating the voluntary compliance behavior shifts after a firm is audited. Second, apart from the theoretical approach and laboratory experiment approach mentioned above, we used the difference-in-differences non-experimental approach in our study, which is a new approach to the topic in the literature. Third, historical population data of a New York State economic sector were used in this study instead of experimental data or randomly selected sample data used in the literature.

The remainder of this paper consists of five sections. The first section discusses data used in this research and the second section discusses modeling methodologies we tried. The third section presents modeling results from two specific approaches: a pooled ordinary least squares (OLS) regression for all audit groups (an audit group is defined as aggregate of all firms which were audited in the same year) and a time series cross section (TSCS) regression with the first-order autoregressive method. The fourth section discusses the dollar amount of the audit impact based on the regression coefficients from the TSCS approach. Finally, the fifth section concludes the paper.

I. Data

A. Audit, Assessment, and Sales Tax Data

The audit, assessment, and sales tax data used in this study are from three data sets maintained by the New York State Department of Taxation and Finance. The first is the sales tax data set comprised of sales tax returns and account information, including taxpayer names, addresses, business descriptions, industrial code, sales revenue, and sales tax. From this file, all taxpayers in the Food Services and Drinking Places (NAICS 722) were selected.
The Food Services and Drinking Places sector was chosen for several reasons. First, the sales of individual firms in this sector are relatively stable and are not seriously influenced by economic cycles as those in many other sectors, such as auto or building material sales, so the model results may be less impacted by economic cycles. Second, the data of a large portion of firms in this sector are available for the whole period from 1998 to 2008, our studying period. Third, the number of firms audited each year from 2003 to 2008 in this sector is relatively stable and large enough to avoid sampling bias.

There are 6,886 firms from the Food Services and Drinking Places sector included in this study. They are chosen because they were in business and filed sales tax returns for the entire study period.

The second data set is an audit data set. The audit data set contains a history of audits and audit results. The 6,886 firms selected from the sales tax data set were divided into two groups: experimental and control. If a firm experienced an audit during the study period, it was placed in the experimental group. Otherwise, a firm was placed in the control group. The experimental group includes 1,955 firms and the control group includes 4,931 firms. The audit data set also contains the first-contact date, which is the date on which a firm is initially informed that it is facing an audit. This date is used as the audit date to divide the 1,955 experiment firms into six sub-groups, from 2003 audit group to 2008 audit group.

The third data set is an assessment data set. In the case that the first-contact date for a firm is missing from the audit data set, the assessment date specified in the assessment data set will be used as the audit date. The summary of the control group and the six experimental groups is presented in Table One.
Table One: Summary of Audit Groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Audit Year</th>
<th>Number of firms</th>
<th>Sales in 2008 ($millions)</th>
<th>Sales Tax in 2008 ($millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Not audited</td>
<td>4,931</td>
<td>2,596.6</td>
<td>215.5</td>
</tr>
<tr>
<td>1</td>
<td>2003</td>
<td>286</td>
<td>1,070.9</td>
<td>90.1</td>
</tr>
<tr>
<td>2</td>
<td>2004</td>
<td>238</td>
<td>577.4</td>
<td>48.6</td>
</tr>
<tr>
<td>3</td>
<td>2005</td>
<td>512</td>
<td>1,010.6</td>
<td>84.6</td>
</tr>
<tr>
<td>4</td>
<td>2006</td>
<td>390</td>
<td>617.2</td>
<td>51.4</td>
</tr>
<tr>
<td>5</td>
<td>2007</td>
<td>263</td>
<td>289.1</td>
<td>24.1</td>
</tr>
<tr>
<td>6</td>
<td>2008</td>
<td>266</td>
<td>279.6</td>
<td>23.8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>6,886</strong></td>
<td><strong>6,441.5</strong></td>
<td><strong>538.1</strong></td>
</tr>
</tbody>
</table>

B. Economic Variables

The purpose of this study is to investigate the impact of audit on voluntary compliance. Reported sales revenue of firms was chosen as the dependent variable for our models since it is the best indicator of voluntary compliance. The sales revenue is greatly influenced by economic conditions. In a booming economy, for example, the increase in wages may boost retail sale transactions, which means a higher level of sales revenue.
To control the effect of general economic conditions on the sales revenue, several economic indicators were tried in our models as explanatory variables. One indicator is the total wages of New York State (Data source: US Bureau of Labor Statistics), as the sales revenue is closely tied to the total State wages. We also tried New York State wages in the private sector, New York State personal income, and New York State disposable income (Data source: Bureau of Economic Analysis) to determine which economic indicator has the closest relationship to the sales revenue.

II. Modeling Methodology

In the modeling process, numerous approaches were tried to determine the relationship between an audit and voluntary compliance. The general form of the models can be written as

\[ y = \alpha_0 + a_1 X + \sum_{j=1}^{m} \beta_j Z_j + \epsilon \]  

(E1)

where \( y \) may be the reported sales or may be some variations, such as difference or growth rate, of the reported sales. \( X \) is a dummy variable for audit (audit dummy), which is equal to 0 before the audit year and 1 for the audit year. To explore the best model fit, we tried a variety of assumptions about the audit dummy after the audit year. \( Z \)s represent factors other than the audit dummy affecting the sales revenue, such as New York State personal income or total wages. Like the dependent variable \( y \), \( Z \)s may take alternative forms, such as difference or growth rate. In some cases, \( Z \)s were dropped from our trial models. One reason for dropping them is that they are not statistically significant; another reason is that these economic variables may not be relevant in some cases because they are not audit group specific. In these cases, we assume that these variables have the same or similar impact on the audited firms and not-audited firms.

We tried numerous modeling specifications of E1 to improve the quality of the models, including Ordinary Least Squares (OLS) regressions for each audit group, OLS regressions for pooled data of all audit groups, and Time Series Cross Section (TSCS) models. In the TSCS modeling trials, alternative estimation methods were tried, including one and two-way effect methods, first-order autoregressive model (Parks method), and the moving average error correction model (Dal Silva method).
III. Modeling Results

In this section, two specific models we tried will be presented. The first is the pooled OLS regression for all audit groups, and the second is the TSCS model with the first-order autoregressive method. They are the best among the models we tried. They are chosen for their parsimonious nature and significant statistics. While the regression results of the two models are similar, we will show that the TSCS model is superior to the pooled OLS model. The dollar amount of the audit impact on the reported sales and sales tax will be presented in Section IV based on the estimated coefficients of the TSCS model.

In both models, the difference in the growth rate of reported sales between each of the audit groups and the non-audit group is used as the dependent variable. There are several advantages to using the growth difference other than the level or growth rate. First, it avoids the stationary controversy in using levels; second, many external factors, such as tax rate changes, can be dropped from the right hand side of the equation since it can be reasonably assumed that they have the same or similar impact on different audit groups; and third, it avoids the need of normalizing the level data to make the data comparable among different audit groups, which would be necessary if we use levels because of the differences in the number of firms and the level of sales among these audit groups.

A. Pooled OLS Model for All Audit Groups

Before the pooled OLS model was tried, we had tried separate OLS regressions for each of the six audit groups. The structural form of the model for Audit Group \( i \) is:

\[
y_i = \alpha_{0,i} + \alpha_{1,i} X_i + \varepsilon_i
\]

(E2)

where

- \( y_i \) = the difference in sales revenue growth rates between Group \( i \) (audited in year \( i \)) and Group 0 (control group);
- \( X_i \) = audit dummy for Group \( i \). It equals to 0 before the audit year and one in the audit year, and decays at a certain rate afterwards.
There are no economic variables on the right hand side of the equation because they are not statistically significant in all cases. The estimation results (not presented in this paper) show that the regression statistics for almost all audit groups is not satisfactory: the adjusted R-square is low, T statistics is not significant, and Durbin Watson is far from the ideal level. Overall, the relationship between the reported sales and audit could not be detected for most of the audit groups.

Another shortcoming of E2 is the very limited number of observations in each regression. There are only 11 observations (1998 to 2008) available in the data set and one observation is lost when the level is converted to the growth rate. For this, we think a pooled model may be more appropriate. In the pooled model, the observations for all audit groups are stacked together as one data set. In the case that we want to try the lagged dependent or lagged independent variables on the right hand side of the equation, they are specified as extra independent variables since the time series nature of the data has been removed in the pooled model.

Equation E3 is the specification of the pooled model, which is similar to Equation E2, except that it includes observations of all six audit groups instead of one audit group, and it has one regression instead of six regressions. Therefore, there is only one \( \alpha_0 \) and one \( \alpha_1 \) for all audit groups.

\[
y_i = \alpha_0 + \alpha_1 X_i + \epsilon_i \quad \text{(E3)}
\]

where

\( y_i \) = the difference in sales revenue growth rates between Group \( i \) (audited in year \( i \)) and Group 0 (control group);

\( X_i \) = audit dummy for Group \( i \). It equals to 0 before the audit year and 1 in the audit year, and decays at a certain rate afterwards.

Again, there are no economic variables on the right-hand side of Equation E3 because they are not statistically significant. The insignificance of the economic variables may indicate that they have the same or similar impact on the sales revenue growth for different audit groups.
Various assumptions of the audit dummy, $X_t$, were tested. The following is the regression results of Equation E3 with the assumption that the audit impact will decay at a rate of $1/2$ with each passing year after the audit year.

Statistics of Equation E3 (with $X$ decreasing at the rate of $1/2$ each year after the audit year):

\[
\begin{array}{ccc}
\alpha_0 & \alpha_1 \\
0.0079 & 0.0272 \\
\end{array}
\]

P value (0.0976) (0.0431)

Adj R-square: 0.0526

Durbin-Watson: 1.9073

Compared with the separate regressions for each of the audit groups represented by Equation E2, the pooled regression has two improvements. First, the degree of freedom increased substantially, from 8 to 58; second, both adjusted R-square and T are higher. It is clear that T statistics for the intercept, $\alpha_0$, is not significant at the 95 percent confidence level. Also, the Adjusted R-square, 5.26 percent, is still low. However, if considering the fact that $y$ is not the level of sales, nor the growth rate of sales, but the difference in the growth rate between the audit groups and the non-audit group, we would think that the low adjusted R-square may be acceptable.

\section*{B. Time Series Cross Section (TSCS) Regression}

In addition to the fact that the Adjusted R-square is low in the pooled OLS model presented above, another shortcoming of the model is that it disregards the time series nature of the data, as we mentioned previously. In the remaining part of this section, we will use the time series cross section models to overcome the shortcomings. We tried various TSCS modeling techniques and found that the best is the TSCS first-order autoregressive model (Parks Method) specified below:

\[
y_{it} = \alpha_0 + \alpha_1 X_{it} + \alpha_2 Z_t + u_{it} \quad \text{(E4-a)}
\]
\[ u_{it} = \rho_i u_{i,t-1} + \varepsilon_{it} \]  
(E4-b)

where

\( t = 1999, 2000, \ldots, 2008 \), represents the years the data cover;

\( y_i = \) difference in sales revenue growth rate between Group \( i \) (audited in year \( i \)) and Group 0 (control group);

\( X_i = \) audit dummy for Group \( i \). It equals to zero before the audit year and equals to \( (\frac{2}{3})^n \)

since the audit year, where \( n = 0, 1, \ldots, N \), the number of years since the audit year;

\( Z = \) growth rate of New York State total wages.

Statistics of Equations E4-a and E4-b:

<table>
<thead>
<tr>
<th>( \alpha_0 )</th>
<th>( \alpha_1 )</th>
<th>( \alpha_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1216</td>
<td>0.0263</td>
<td>-0.1052</td>
</tr>
</tbody>
</table>

P value

\( (0.0012) \quad (0.0001) \quad (0.0030) \)

\( \rho_{2003} \quad \rho_{2004} \quad \rho_{2005} \quad \rho_{2006} \quad \rho_{2007} \quad \rho_{2008} \)

\( 0.1918 \quad 0.0873 \quad 0.5200 \quad -0.4013 \quad -0.4066 \quad 0.3906 \)

R-square: 0.4989

The statistics shows that the TSCS autoregressive model has significant improvements over the pooled OLS model. First, the R-square increases to 0.4989; Second, the T statistics for the intercept and each of the independent variable is more significant, especially the T statistics for \( \alpha_1 \), the parameter we are most interested in. The low p-value for each coefficient indicates that the independent variables are significant contributors to the variation of the dependent variable and should not be dropped from the modeling.
One advantage of the TSCS autoregressive model is that it takes into consideration the influence of the past behavior on the current period. A close examination of the autocorrelation coefficients, $\rho$s, tells us that a significant portion of past error components for most audit groups do contribute to the error term of the current period.

IV. The Audit Impact

It is interesting that, although the modeling methods are different, the pooled OLS model represented by Equation E3 and the TSCS autoregressive model represented by E4-a and E4-b produce similar coefficients for the audit dummy, 0.0272 and 0.0263, respectively. Since the TSCS autoregressive model is considered superior to the pooled OLS model, in this section we will use the results of the TSCS model to compute the dollar amount of the audit impact on the reported sales and sales tax.

The audit dummy plays an important role in calculating the audit impact on reported sales and sales tax. We choose the audit dummy with a decaying rate of 1/3 for the TSCS model because it gives the best fit. The regression results show that the audit impact is 2.63 percentage points in the audit year for the experimental groups, i.e., the relative growth rate (relative to the growth rate of the control group) of an experimental group is 2.63 percentage points higher than in the case if the experimental group were not audited, and the impact would decay at the rate of 1/3 each year afterwards. The scale of the impact is presented in Table Two and Figure One.

Table Two: Audit Impact on Reported Sales Growth Rate (pp = percentage points):

<table>
<thead>
<tr>
<th>Audit year</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
<th>Year 5</th>
<th>Year 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact</td>
<td>2.63 pp</td>
<td>1.75 pp</td>
<td>1.17 pp</td>
<td>0.78 pp</td>
<td>0.52 pp</td>
<td>0.35 pp</td>
</tr>
</tbody>
</table>

Figure One:
The audit impact diminishes over time. After five years, it becomes insignificant (below 0.5 percentage point).

Given the fact that \( y \) is the difference in the sales growth rate between an audit group and the non-audit group,

\[
y_t = \frac{s_{a,t}}{s_{a,t-1}} - \frac{s_{n,t}}{s_{n,t-1}} \quad (E5)
\]

where \( s_a \) is the sales revenue of an audit group and \( s_n \) is the sales revenue of the non-audit group, the change in \( y_t \) can only be caused by the change of \( s_{a,t} \) since the other three items in the right-hand side of Equation E5 could not be affected by the current year audit to the audit group. If \( y_t \) is going to increase by a certain percentage points, for example, 2.63 percentage points as in the case of our TSCS model suggests, the only way to make it happen is that \( s_{a,t} \) increases by the amount of \( 0.0263 \times s_{a,t-1} \).

The dollar amount of the impact for each audit group is calculated in the following three steps. In the first step, a new series of the assumed sales growth rates without audit impact was calculated based on the information in Table Two. For example, if the reported sales growth rate
of the 2003 audit group is 7.00 percent in 2003 and 6.50 percent in 2004, then the assumed growth rate without the audit impact would be 4.37 percent \((= 7.00 - 2.63)\) in 2003 and 4.75 percent \((= 6.50 - 1.75)\) in 2004. In the second step, a series of sales without audit impact was obtained by multiplying the reported sales in the base year (the year before the audit year) by the series of the assumed sales growth rates (without the audit impact) obtained from the first step. And finally, the audit impact is obtained by deducting the series of sales without audit impact from the reported sales. The dollar amount of the impact on reported sales and sales tax for all audit groups is presented in Table Three.

**Table Three:**

<table>
<thead>
<tr>
<th>Year</th>
<th>Reported Sales</th>
<th>Sales without Audit Impact</th>
<th>Audit Impact on Sales</th>
<th>Sales Tax Collected</th>
<th>Sales Tax without Audit Impact</th>
<th>Audit Impact on Sales Tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>2,741.5</td>
<td>2,741.5</td>
<td>-</td>
<td>221.4</td>
<td>221.4</td>
<td>-</td>
</tr>
<tr>
<td>2000</td>
<td>2,923.3</td>
<td>2,923.3</td>
<td>-</td>
<td>236.1</td>
<td>236.1</td>
<td>-</td>
</tr>
<tr>
<td>2001</td>
<td>3,007.7</td>
<td>3,007.7</td>
<td>-</td>
<td>243.5</td>
<td>243.5</td>
<td>-</td>
</tr>
<tr>
<td>2002</td>
<td>3,006.1</td>
<td>3,006.1</td>
<td>-</td>
<td>242.8</td>
<td>242.8</td>
<td>-</td>
</tr>
<tr>
<td>2003</td>
<td>3,091.3</td>
<td>3,070.6</td>
<td>20.8</td>
<td>250.7</td>
<td>249.1</td>
<td>1.7</td>
</tr>
<tr>
<td>2004</td>
<td>3,220.3</td>
<td>3,170.9</td>
<td>49.4</td>
<td>269.0</td>
<td>264.8</td>
<td>4.1</td>
</tr>
<tr>
<td>2005</td>
<td>3,471.9</td>
<td>3,377.5</td>
<td>94.4</td>
<td>295.1</td>
<td>287.2</td>
<td>7.9</td>
</tr>
<tr>
<td>2006</td>
<td>3,593.2</td>
<td>3,452.9</td>
<td>140.3</td>
<td>300.7</td>
<td>289.0</td>
<td>11.7</td>
</tr>
<tr>
<td>2007</td>
<td>3,709.6</td>
<td>3,529.0</td>
<td>180.6</td>
<td>310.0</td>
<td>295.1</td>
<td>15.0</td>
</tr>
<tr>
<td>2008</td>
<td>3,844.8</td>
<td>3,625.8</td>
<td>219.0</td>
<td>322.6</td>
<td>304.4</td>
<td>18.2</td>
</tr>
<tr>
<td>Total</td>
<td>20,931.1</td>
<td>20,226.6</td>
<td>704.6</td>
<td>1,748.1</td>
<td>1,689.5</td>
<td>58.6</td>
</tr>
</tbody>
</table>

* Only firms audited from 2003 to 2008 and with data available for the whole study period (1998 to 2008) are included.

Table Three shows that among the reported sales of $20,931 million by the 1,955 firms audited during the period from 2003 to 2008, about $705 million is from the audit impact on the voluntary compliance, accounting for about 3.4 percent of the total sales of these firms. Among
the $1,748 million sales tax collected from these firms from 2003 to 2008, $59 million, or 3.4 percent, is from the audit impact.

There are total 6,886 firms are included in this study: 1,955 audited firms and 4,931 not-audited firms. If we consider all of the 6,886 firms under study, the total reported sales from 2003 to 2008 are $35,680 million: $20,931 million for the audited firms and $14,749 million for the not-audited firms. The audit impact of $705 million accounts for about 2.0 percent of the total reported sales of the 6,886 firms. Similarly, the audit impact on the sales tax, $59 million, also accounts for 2.0 percent of the total sales tax.

The audit impact on the audited firms since 2003 is presented in Figure Two and Figure Three. Figure Two depicts the audit impact on reported sales while Figure Three the audit impact on sales tax. From these figures, we can see that the audit impact becomes stronger as time passes. This is because for each year after 2003, there are two factors contributing to the total audit impact, one being the audit impact on the newly-audited firms during the current year and the other being the audit impact on the firms which were previously audited.

**Figure Two:**

![Audit Impact on Reported Sales](image)

*Only firms audited from 2003 to 2008 and with data available for the whole study period (1998 to 2008) are included.*
V. Summary and Conclusions

Many modeling approaches were explored in this research to investigate the audit impact on reported sales. Several variations of the sales measure were tried as the dependent variables to
determine which one could be better explained by our models. In addition, both the constant and sliding audit dummies with various decay rates were tried to determine how big the scale of the impact is and how long it would last after a firm is audited.

For the Food Services and Drinking Places (NAICS 722) Sector in New York State, we found that both the pooled OLS model and the Time Series Cross Section autoregressive model produce similar estimation coefficients for the audit dummy and that the TSCS model with a sliding audit dummy provided the best adjusted R-square and T statistics. The audit impact on reported sales and on the sales tax was calculated according to the regression results of the TSCS model.

With an audit, a typical firm would report a higher sales growth rate in the year of the audit, 2.63 percentage points higher than in the case that the firm were not audited. The percentage would decay at a rate of 1/3 each year afterwards, i.e., extra $1.75 (=2.63 \times \left(\frac{2}{3}\right)^1$) percentage points in the second year, extra $1.17 (=2.63 \times \left(\frac{2}{3}\right)^2$) percentage points in the third year, and so forth.

In terms of the dollar amount, the audit impact contributed to $705 million of increased reported sales from 2003 to 2008 for the 1,955 audited firms. This is approximately 3.4 percent of the total reported sales of these firms during this period. New York State collected about $59 million more in sales tax during the same period due to the higher reported sales, about 3.4 percent of the total sales tax collected from these 1,955 firms.

This study does find a positive relationship between the audit and the voluntary compliance. The findings suggest that the audit productivity may be underestimated in many studies in the literature. It reminds us that when considering the productivity of the audit work, besides the direct audit collections, we should also take the audit impact on the voluntary compliance into consideration. For this reason, the finding may provide tax professionals and tax authorities with incentives to strengthen the audit power and to better structure the audit organizations to generate more revenue for the State.
REFERENCES


