Cite as Det. No. 02-0114, 22 WTD 174 (2003)

BEFORE THE APPEALS DIVISION
DEPARTMENT OF REVENUE
STATE OF WASHINGTON

In the Matter of the Petition For Correction of Assessment of	)	D E T E R M I N A T I O N
	)	No. 02--0114
. . .	)	Registration No. . . .
	)	FY . . . /Audit No. . . .
	)	Docket No. . . .

MISCELLANEOUS: AUDITS – TEST PERIODS – STATISTICAL SAMPLING – METHODOLOGY. Using a sampling of documents and projecting that sample over the entire period under review is an accepted and commonly used method to audit records. Over recent years the use of statistical sampling in sales and use tax audits has increased. In designing a statistical sampling program, consideration must be given to the cost of sampling, precision, and sample size. Contrary to the taxpayer's argument, we conclude the Department's election to use an 80% lower limit in assessing tax, with certain mandated sample sizes, was based on a conservative cost-benefit analysis and not on any flaw in methodology.

Headnotes are provided as a convenience for the reader and are not in any way a part of the decision or in any way to be used in construing or interpreting this Determination.

Mahan, A.L.J. – After agreeing to the use of random sampling in a use tax audit, the taxpayer protests the methodology used by the Department of Revenue (Department).[1]

ISSUES

1. When using statistical sampling as a means to estimate the amount of underpayment of use tax, is the Department required to use the lower end of a 95% confidence interval in determining tax liability?

[1] Identifying details regarding the taxpayer and the assessment have been redacted pursuant to RCW 82.32.410.

2. In the absence of a statute, rule, or written policy statement concerning sampling methods, must the taxpayer consent to the methodology used by the Department in its statistical sampling program?

3. Did the Department err in assessing tax on an invoice where retail sales tax had been paid and in assessing use tax on an invoice allegedly for computer training purposes?

FACTS

The taxpayer operates a hospital located in the State of Washington. The Department audited the taxpayer's records for the January 1, 1994 through September 30, 1998 period. Because of the large number of accounts payable records, the Department and the taxpayer initially agreed to a random, non-statistical sampling of the records. Under this plan, the Department randomly selected 1,000 records (out of a stratified population of 283,715) for review and projected that result over the entire audit period. The taxpayer disagreed with the results of the initial projection, and sought to apply a statistical evaluation to the random sample of 1,000 records. The Department countered that, to evaluate the results on a statistical basis, the audit plan would have to be changed and a larger sampling done. It was agreed that the sample size would be increased to 2,000 records. The Department used its recently created Computer Assisted Audit Program (CAAP) to identify files for review and to estimate the underpayment or overpayment of use or sales tax, based on a statistical sampling of the records randomly selected for review.

Following the results of its random sampling of the taxpayer's records, the Department projected a deficiency in the amount of $ . . . based on a statistical evaluation of the sample results. The total assessment, including the projected amount, totals $ . . . . The taxpayer timely appealed this assessment.

The taxpayer asserts the Department erred as a matter of law in making certain choices (discussed in further detail below) over how to statistically evaluate and project the sample results. It contends that the "confidence interval and a one-tailed test [as used by the Department], leaves a 20% chance that the true error rate is lower than the estimate. In other words, there is a 20% chance that the tax assessed is too high." Rather, the taxpayer asserts that a "two-tailed test" with a 95% confidence interval should be used. This would leave a 97.5% probability (under the lower limit of the confidence interval) that the tax assessed was too low and only a 2.5% probability that the tax assessed was too high.[2]

[2] We are aware of no state or federal taxing agency regularly using a confidence level of 97.5%, and the taxpayer has not presented any evidence that such a level is used by any taxing authority. Further, the taxpayer's description of the statistical procedures it wants used may be inaccurate. The Department takes the position that a lower limit is established as a result of a *one-tailed* test. The term "tail" we understand to be the upper and lower limit points of a confidence interval (as discussed further *infra*). A *two-tailed* test utilized in a statistical evaluation results in a point estimate, which is the mid-point between the upper and lower limits when using a confidence interval. *See* Will Yancey & Roger C. Pfaffenberger, *Use and Abuse of Sampling in Sales and Use Tax Audits*, 97 *COST* State Tax Report, Issue 6, pp. 2-9 (November 1997), reprinted in 13 State Tax Notes 1673 (December 29, 1997). Regardless,

In response, the Department stated it would be willing to go to a lower limit with a one-tailed 95% confidence level, provided that the sample size was substantially increaseed over the 2,000 records already selected. According to the Department, in order to maintain the same level of precision as attained in the revised sample, the sample size would have to be quadrupled. The taxpayer rejected this counter proposal. The taxpayer states it would not agree to an increased level of random sampling, and would prefer an actual review of all records to an increase in sample size.

The taxpayer also asserts the Department erred as a matter of law by using the methodology it did in the absence of the taxpayer's consent. Whether there was agreement on the projection method is disputed. The Department states it informed the taxpayer verbally as to the methodology to be used. Following the revised random sampling, the Department provided the taxpayer with a detailed letter describing the methodology to be used in projecting the sample, including specific information on confidence intervals and the use of an 80% lower limit. The Department clearly made great efforts to inform the taxpayer on the methodology to be used and to address any question or concern raised by the taxpayer.

However, the Department has no contemporaneous written records to show actual or constructive consent as to the methodology with respect to confidence intervals and limits. At the hearing on this matter, the taxpayer and the taxpayer's representatives stated they did not agree to the methodology to be used by the Department in projecting the sample. Given the relative recent development of the Department's program, the lack of contemporaneous records showing actual or constructive consent, and the contrary testimony by the taxpayer and its representatives, we find more likely than not, based on the evidence before us, that the Department did not have the taxpayer's actual or constructive consent with respect to the methodology to be used in projecting the sample. We further note that since this audit the Department has developed an engagement letter to inform taxpayers of the methodology to be used in its sampling program. Further, the Department is in the process of drafting a sampling manual, which may be used to further give notice to taxpayers of the Department's methodology.

The taxpayer also asserts the Department erred in not giving it credit for sales tax paid on one invoice and assessing tax on an invoice where the services provided were not subject to use or sales tax. At the hearing on this matter, the taxpayer provided copies of the disputed invoices. The invoice where tax was allegedly paid was not on the list of invoices on which the Department based the assessment of additional tax. The taxpayer believes the invoice at issue (which shows that the invoiced amount included tax) was paid under a later invoice number. However, the taxpayer at the hearing did not have any records to trace the payment of the invoice at issue to a later invoice. The taxpayer also presented an invoice for "professional services" rendered by a computer software and hardware company. The taxpayer believes that invoice was for computer training, but did not present at the hearing any records to show that

---

for the purposes of this decision we understand that the taxpayer wants a 95% confidence interval used in order to be 97.5% confident that the lower limit is no more than the true amount due.

such services were for training. In a post-hearing submittal, the taxpayer provided copies of documents concerning invoices and payments to the software and hardware company.

ANALYSIS

The goal of a sales and use tax audit is to identify the total amount of underpaid or overpaid tax for the period under review. For many businesses, in particular businesses with large numbers of transactions, it is a costly and time-consuming process for both the taxpayer and the Department to review all records for the entire period under review. The Department recognizes that a sampling of documents, rather than a review of all of the records for the entire period, and projecting the results over the entire period, is an accepted and commonly used auditing method to estimate the amount of tax underpaid or overpaid. The Department has often used test periods as a sampling method. *See, e.g.*, Det. No. 88-233, 6 WTD 59 (1988); Det. No. 87-354, 4 WTD 293 (1987); Det. No. 93-240, 13 WTD 369 (1994). Such methods are also known as block sampling.[3]

Another sampling method is statistical sampling. Statistical sampling requires random selection of the sample and the use of probability theory to evaluate the sample results. Over recent years the use of statistical sampling in sales and use tax audits has increased. Statistical sampling can add greater accuracy and precision to the efficiency that block sampling provides.[4] *See generally*, Faranak Naghavi, Jeri Mulrow, & Eric Falk, *Reversals and Use Tax Audits and Statistical Sampling: A Double Benefit*, 11 J. Multistate Tax'n 18 (2002). Such sampling must be done on a random and not on a block basis. If the number of transactions under audit is large,

---

[3] In block sampling, the auditor chooses a group or block of transactions that took place during a specified time period or in a certain location. The selection of the block is usually not the result of a random process. Generally, the block that is selected based on convenience and accessibility. *See* Jeri Mulrow, *Statistical Sampling as a Win-Win in Tax Audits*, 15 State Tax Notes 1491 (December 7, 1998).

[4] The potential differences in accuracy between block and statistical sampling was discussed in Will Yancey & Roger C. Pfaffenberger, *Use and Abuse of Sampling in Sales and Use Tax Audits*, 97 *COST State Tax Report* , Issue 6, pp. 2-9 (November 1997), reprinted in 13 *State Tax Notes* 1673 (December 29, 1997), as follows:

> The severe limitation of nonstatistical sampling is that it does not allow the auditor to make a quantitative estimate of sampling risk. An example of nonstatistical sampling is block sampling in which the auditors select a few days or weeks from the population which the auditor or taxpayer deems to be representative of the entire population. By not taking sample transactions over the entire audit period, block samples run the risk of producing sample information that is relevant only to the period for which the sample was taken. If the tax deficiency rate in the sample differs significantly from the population, the block sampling method will produce results that are not valid.
>
> *Statistical sampling* methods provide a quantitative estimate of the sampling risk. Statistical sampling requires that the person selecting the sample rely on a random sample selection process rather than his or her judgment about the extent to which the sample represents the population. The statistical sample might not be a good representation of the population in some instances, but this sampling risk can be quantified using statistical formulas derived from the theory of probability.

stratified random sampling can be used, which stratifies the population into subgroups according to specified attributes, for greater efficiency and accuracy.

The determination of the size of the sample in a sales and use tax audit is an important consideration as the sample size affects the reliability of the results. Sample size determination is inevitably a trade-off between the cost of sampling and precision. Increasing sample size results in both a more costly audit and a more precise estimate.[5]

Sampling size is also related to the outcome with respect to the point or midpoint of the estimate and the "confidence interval" of the estimate. As explained in Will Yancey & Roger C. Pfaffenberger, *Use and Abuse of Sampling in Sales and Use Tax Audits*, 97 *COST* State Tax Report, Issue 6, pp. 2-9 (November 1997), reprinted in 13 State Tax Notes 1673 (December 29, 1997):

> A point estimate is a single number that is chosen to best estimate an unknown population parameter. In sales and use tax audits, the targeted population parameter is typically the total amount of underpaid taxes. . . . A point estimate of a population parameter does not provide information about the *reliability* of the estimator. To do that, it is necessary to provide a confidence interval estimate of the population parameter. We might say, for example, that our point estimate of the total amount of tax underpayment by a corporation in a three year period is $5.0 million, and that we are 95% confident that the total amount of tax underpayment is between $4.5 and $5.5 million. The width of the confidence interval provides a measure of the reliability of the point estimator. As the width of the confidence interval increases, the reliability we place in the point estimator value decreases.
>
> In the popular press, the confidence interval is generally reported in terms of the *margin of error* of the estimator (e.g., "It is predicted that a political candidate will receive 60% of the vote *with a margin of error of 4%*). The margin of error is approximately the amount to be added and subtracted from the point estimate to produce a 95% confidence interval estimate (e.g., for the example above, the 95% confidence interval is from 56% to 64% of the vote). The margin of error or the confidence interval half-width is determined based on the sample size, the degree of confidence required, and on the variation among the sampled observations.

---

[5] As stated in Herbert Arkin, *Sampling Methods for the Auditor: An Advanced Treatment* 153 (1982) with respect to the use of a lower limit in a tax audit:

> It will be remembered that the lower limit is calculated by subtracting the sampling error of the total from the point estimate of the total. Thus, the smaller the sample, the larger the sampling error and the less collected!
>
> This gives rise to a cost-benefit relationship of considerable importance. The larger the sample, the greater the amount collected, but also the greater the cost of the audit.

In using a statistical evaluation to project the sample results, adjustments can be based on different levels of precision. If the sample evaluation is sufficiently precise, the point estimate (mid-point) of the sample projection may be used. If precision is more than just a small percentage, the point estimate becomes more tenuous. In these cases, the lower limit of the confidence interval may be used.[6] *Id.*

States using statistical sampling have made different choices as to desired levels of precision, the use of point estimates or lower limits, and confidence levels. For example, New York uses the following procedures in its statistical sampling program:

> The department's guideline regarding audit assessments determined through the use of statistical sampling is to assess any tax liability or provide a refund based upon the mid-point of the estimate. The sample would be considered valid as long as the precision of the estimate calculated at a 90% confidence interval based on a two-sided test is no greater than 25% of the estimate.

New York State Department of Taxation & Finance, *Computer-Assisted Audits: Guidelines and Procedures for Sales and Use Tax Audits*, Publication 132 (revised October 2001).

In comparison, the Department uses different criteria in how it projects sample results. In this case the tax deficiency is based on the 80% lower limit (not the midpoint) with a precision goal of 30%. The benefit in selecting this confidence level would be a smaller sample size with a resulting cost reduction. By way of further trade-off, the Department used the lower limit of the confidence interval, which effectively has the Department assuming the risk that four out of five times taxpayers would probably underpay sales and use tax. In contrast, at the midpoint estimate an equal number of taxpayers would probably pay more than they actually owe and an equal number would pay less. Using the lower limit of a confidence interval, has been described as fairly adjusting the amount assessed due to the uncertainty in an estimate; "if the state concludes that the uncertainty adjustment is sacrificing too much revenue, it has the option of drawing larger samples at its own cost." Joseph C. Bright, Joseph B. Kadane, & Daniel S. Nagin, *Statistical Sampling in Tax Audits*, 13 Law & Social Inquiry 305, 324 (Spring 1988).[7]

Based on our review of the literature cited above and the facts and arguments presented, we find the taxpayer has not shown by a preponderance of the evidence that the assessment was

---

[6] This case concerns a deficiency assessment. The Department has a different policy with respect to overpayments. Whereas the Department largely assumes the risk of sampling error with respect to debits, the risk of sampling error is largely assumed by taxpayers with respect to overpayments.

[7] *See also* Herbert Arkin *Sampling Methods for the Auditor: An Advanced Treatment* 153 (1982), where the author in discussing a tax audit stated:

> Thus, if a claim is made based on a sample, for the collection of a sum of money, the claimant must specify the *lower* limit of his confidence interval estimate in order to have a defensible figure. It can be said with a high degree of assurance that *at least* this amount is due. The actual amount due is probably greater than that figure but cannot be determined.

fundamentally flawed. The Department's election to use an 80% lower limit appears based on a conservative cost-benefit analysis and not on any flaw in methodology. More specifically, without increasing sample size, the taxpayer in the present case seeks to employ a 95% confidence interval with tax assessed at the lower limit of 97.5%. This would result in a reduced tax assessment while greatly increasing the risk of an underpayment of tax. The taxpayer cites to no authority whatsoever in support of such a practice or to any legal authority for the proposition that such a policy choice could be imposed on the Department. Accordingly, its appeal on this ground is denied.

We then must address whether, in the absence of the taxpayer's consent regarding the methodology to be used in the statistical sampling, we can sustain the assessment in this case. Unless records are unavailable, the Department in cases involving block sampling requires the consent of the taxpayer. *See, e.g.*, Det. No. 88-233, 6 WTD 59 (1988); Det. No. 87-354, 4 WTD 293 (1987); Det. No. 93-240, 13 WTD 369 (1994). We have not previously addressed whether this same principle applies in cases involving statistical sampling.

Because of increased accuracy and ability to quantify risk, an assessment based on statistical sampling presents a better case for being sustained in the absence of consent. *See Underwood v. Fairbanks North Star Borough*, 674 P.2d 785, 788 (Alaska 1983) (if a sample size is large enough to be statistically reliable, an estimate based on statistical sampling is admissible into evidence to support the assessment). This may be particularly true where a lower limit is used and the taxing authority assumes a risk of underpayment of tax. *See* Joseph C. Bright, Joseph B. Kadane, & Daniel S. Nagin, *Statistical Sampling in Tax Audits*, 13 Law & Social Inquiry 305, 313 (Spring 1988); *see also* Ill. Private Letter Ruling No. 95-0377, Ill. PLR Lexis 520 (1995) (on an administrative basis, statistical sampling can be used in the absence of consent and explicit statutory authority, even when adequate records are kept); *Shetakis Distributing Co., Inc., v. Nevada*, 108 Nev. 901, 839 P.2d 1315 (1992).

However, to resolve the case before us, we need not address whether a different rule should apply to assessments based on statistical sampling. In the present case, the various policy choices as to precision and confidence intervals were not clearly established by the Department prior to the audit. In the absence of such established policy directives with respect to statistical sampling, we are not inclined to deviate from the general principle regarding sampling in the absence of consent. As these policy choices become established by the Department, it may be necessary to revisit the issue whether a taxpayer's consent to the methodology used by the Department is required under Washington law.[8]

Accordingly, the case is remanded to the Audit Division for a review of all records. In accordance with RCW 82.32.110, the taxpayer must make written and electronic records available for examination.

[8] With the use of engagement letters and other material that identify the Department's policy choices and methodology, we do not anticipate consent on methodology to be at issue in the future. Rather, we anticipate that issues in the future will center on claimed errors regarding classification of invoices (such as made by the taxpayer in the present case), missing invoices, or other extraordinary items that may need reclassification or correction.

On remand, the Department's risk of an underpayment of tax will probably decrease, but at greater cost. The taxpayer will probably risk a greater tax liability at also a greater cost. To the extent the amount of the assessment does increase based on a review of more records, the taxpayer will not have the election to accept the lower estimate of an underpayment of tax.

As to the claimed errors with respect to either tax paid on invoices where no credit was given or for services that should not have been subject to tax, the taxpayer has the burden to maintain and produce records necessary to determine the amount of tax that is due. RCW 82.32.070. The records provided in a post-hearing submittal will be forwarded on remand to the Audit Division for a determination whether they adequate identify the amount of tax that is due.

DECISION AND DISPOSITION:

The taxpayer's petition is granted in part and denied in part.

Dated this 18th day of July 2002.