Sales and Income Tax Data Collected Comparability: Quarterly vs. Annual Surveys

Dedrick Owens[1]
dedrick.l.owens@census.gov
U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233

Abstract

In this research, we consider whether sales and income tax data collected in the Public Sector Quarterly Tax Survey (QTAX) is comparable to that collected in Annual Survey of Local Government Finance (ALFIN). Ideally, the sales tax collected from QTAX in a year should be similar to that of ALFIN. The key element in assessing this question is the source(s) from which respective surveys obtain their data. As it stands, the collection sources are very different---QTAX data represents responses from state government representatives and ALFIN from state level annual reports or administrative records. In this research, we investigate the magnitude of the differences and discuss how to improve the data quality.

Key Words: Sales Tax, state and local government agencies, Quarterly Tax Survey, Survey of Local Government Finance.

1. Introduction

The Public Sector Quarterly Tax Survey (QTAX) covers the 50 state governments, all local government property tax collectors and local government non-property tax collectors. All state government tax collection is covered. A stratified cluster sample of local property tax collectors yields the estimates of property tax. A stratified sample of local tax imposers yields estimates of all non-property taxes.

The quarterly summary of selected non-property taxes (F-73) is a sample survey of local governments that impose at least one of the following three taxes:

- T09 – General sales and gross receipts tax
- T40 – Individual income tax
- T41 – Corporate net income tax.

F-73 is used to produce estimates of local government totals for T09, T40, and T41 at the national level. The previous F-73 survey, which was modified in 2013, collected information on 11 non-property taxes. The scope of F-73 was reduced from 11 taxes to three to reduce respondent burden and to increase data quality (Quarterly Summary**[4]**).

The Annual Survey of Local Government Finance (ALFIN) surveys the 10,056 local governments (counties, municipalities, townships, special districts, and school districts)

1. Disclaimer: Any views expressed are those of the author(s) and not necessarily those of the U.S. Census Bureau.

including the District of Columbia. The ALFIN provides periodic and comprehensive statistics about governments and governmental financial activities. It collects data on selected governments revenue, expenditures, and debt. Revenue data include taxes (e.g., property, sales, tobacco, motor vehicle, licensing and permit), charges, interest, and other earnings. Similar ALFIN components are labeled T09, T40, and T41. A new sample is selected every five years-- ending in '4' and '9'(Government Finance[**5**]).

In this paper we explore comparability of sales tax and gross receipts (T09) and individual income tax (T40) in the QTAX and ALFIN. Corporate net income (T41) is not considered due to the small number of units (with net corporate income) appearing in both samples.

2. Collection Methods

The ALFIN employs three data collection modes: mail canvass, internet collection, and central collection from state sources. Collection methods vary by state and type of government.

Reviews of government accounting records provide data for most state government agencies and the 48 largest (in terms of 1990 population) county and municipal governments. Data for local governments in about 27 states are consolidated and submitted by state agencies (central collections), usually as electronic transmissions or mutually developed questionnaires (Government Finance[**5**]).

Data for the balance of local governments are obtained via mail questionnaires sent directly to county, municipal, township, and special district governments.

In some cases, the data from central collections and mail canvass procedures were incomplete or questionable. If Census Bureau analysts were unable to obtain corrected data from original sources, they attempted to obtain data from comprehensive annual financial reports (CAFRs). In many instances, supplied/compiled data are supplemented with secondary sources, such as debt information from merchants.

For the QTAX, Government tax authorities report tax revenues by type of tax. Most local governments report only property tax collections and some report significant non-property revenues such as income and sales taxes. State governments report data for more than 25 types of taxes including personal income, sales, corporate income, motor fuel sales, motor vehicle license, and severance taxes. Local government non-property tax data (F-73) are solicited by mail and returned through mail or web collection.

3. Methodology

In this section we use sales tax data to demonstrate the methodology by which to test any statistical difference between data collected in QTAX and the ALFIN. We identify local government non-property tax units that are in both the QTAX and ALFIN samples. Sales tax data, though from the same local government, are reported with different frequency depending on survey. Local governments report sales tax annually in ALFIN and quarterly in the QTAX. Specifically, for the ALFIN we consider sales tax reported by individual agencies on a fiscal year basis and for QTAX the sum of quarterly response for the same period (e.g. quarters 4 prior year, 1, 2, and 3 current year).

To test the hypothesis that ALFIN sales tax collection (annually) is equivalent to the sum of quarterly reported sales tax reported in QTAX, we observe that we have two observations on each of n institutions found in both samples:

Table 1. Example Sales Tax (T09) data Layout for QTAX and ALFIN

Agency i	Y_i (QTAX)	X_i (ALFIN)
1	Y_1	X_1
2	Y_2	X_2
3	Y_3	X_3
4	Y_4	X_4
.	.	.
.	.	.
n	Y_n	X_n

3.1 Test Description

Paired data as such can be analyzed using the Wilcoxon Sign Rank Test (Conover[**1**], 352) when the underlying distribution of the data is not normally distributed. The assumptions are as follow:

a. $Z_i = Y_i - X_i$, and take as our model

$Z_i = \theta + e_i$, where i = 1, 2, 3, 4, ……, n,
Y_i is the weighted sum of quarterly QTAX sales tax for the i^{th} unit
X_i is weighted sales tax reported in ALFIN for the i^{th} unit,
n is the number of units appearing in both samples, and
θ is the unknown location shift

b. The distribution of each Z_i is symmetric about 0.
c. Z_i share a common mean.
d. The Z_i's are mutually independent.
e. Scale of Z_i is at least interval.

We use this test to determine if ALFIN and QTAX Sales Tax taken from the same units have the same mean. That is, we test the hypothesis:

$H_o: E(Z) = 0 \left(i.e., E(Y_i) = E(X_i)\right)$ – exp. QTAX equal exp. ALFIN Sales Tax in unit i
$H_1: E(Z) \neq 0$ exp. QTAX and ALFIN Tax not equal

Define R_i, the Signed Rank, for each pair (X_i, Y_i) as follows:

R_i = the rank assigned to (X_i, Y_i) if Z_i is positive (i.e, $Y_i > X_i$).
R_i = the negative of the rank assigned to (X_i, Y_i) if Z_i is negative.

Form the sum of the positive signed ranks and denote as:

$$T^+ = \sum(R_i \ where \ Z_i \ is \ positive)$$

Given the large sample study, we use the normal approximation and find the test statistic to be:

$$T = \frac{\sum_{i=1}^{n} R_i}{\sqrt{n(n+1)(2n+1)/6}}$$

accept if $-z_{\frac{\alpha}{2}} < T < z_{\frac{\alpha}{2}}$
reject Ho if $T > z_{\alpha/2}$ or $T < -z_{\alpha/2}$

When we consider survey data, symmetry under the null hypothesis means the probability of each unit's difference being positive (QTAX Sales Tax reported > ALFIN Sales Tax reported) is the same ($\frac{1}{2}$) as it probability of being negative. Ties---instances where ALFIN data collected is equal to QTAX data---are removed to conform to the symmetry assumption. Under the null one should observe approximately the same number of positive ranks as negative ranks. In the absence of symmetry (under the alternative hypothesis) we might observe an extremely large (or small number) of positive signed ranks. A preponderance of higher (or lower) positive ranks reflect a high number instances of large discrepancies between QTAX and ALFIN data obtained from the units.

3.2 Estimation of Location Shift

The parameter θ is the estimated difference or "shift" in the distribution of the Y_i's relative to the X_i's. The estimate of shift associated with Signed Rank Statistic, given by the Hodges-Lehmann Estimator (Hollander and Wolfe,**[9]**,33), can be expressed as:

$$\hat{\theta} = median\left\{\frac{Z_i+Z_j}{2}, \ i \le j\right\}, \text{ j =1,…n.}$$

This median of average Zi's estimates the population median (of differences between Y_i and X_i). We use it to estimate any differences in QTAX data reported relative to ALFIN on a given survey item.

3.3 Computation of Confidence Interval

A $(1-\alpha)$ confidence interval for θ is formulated using the n(n+1)/2 averages, $(\frac{Z_i+Z_j}{2})$, $i \le j, i = 1,2,\dots n.$ For n greater than 50, we define $w_p = [n(n+1)/4] + z_p\sqrt{n(n+1)(2n+1)/24}$ as the p^{th} quantile of the sign rank statistic T, where z_p is the p^{th} quantile of the standard normal. Here p is some number between zero and one.

Order averages from smallest to largest. The w_{pth} smallest average and $w_{(1-p)th}$ largest average constitute the $(1-\alpha)$ confidence bounds. For 2014 T09 (with sample size 752), the average associated with $w_{.05}$ (the 131,761th) forms the lower bound and the average associated with $w_{.95}$ (the 151,366th) the upper bound. The 90 percent confidence interval for 2014 T09 is (-0.61M, 0.63M).

A SAS procedure for estimation of the Location Shift and the $(1-\alpha)$ confidence interval (found in Han**[8])** was used to produce the results found in Table 3. This was needed due to the complexity of computations with large samples.

4. Test Results

In this section we examine distributional characteristics of data from the 2014 ALFIN sample and the 2014 QTAX sample. Table 2 provides a summary of key statistics for T09 and T40. The mean, median, and STD are shown in millions (M).

Table 2. Example Summary Statistics for Units appearing in 2014 ALFIN and QTAX Samples

	Sales Tax (T09)		Individual Income Tax (T40)	
	QTAX	ALFIN	QTAX	ALFIN
Mean	66.8M	65.6M	141.2M	141.1M
Median	22.9M	21.9M	15.7M	11.8M
STD	280.8M	272.4M	840.0M	871.1M
Kurtosis	414	429	132.4	130.2
Skewness	18.5	18.9	11.3	11.2
n	752	752	144	144

Note the similarities in descriptive statistics of T09 for QTAX and ALFIN. While T40 means for QTAX and ALFIN are similar as well, there is a distinguishable positive shift in the QTAX median relative that of ALFIN. Kurtosis measures whether the data are heavy or light tailed relative to the normal distribution, and skewness the presence of symmetry. Items T09 and T40 have kurtosis and skewness values suggesting that the data are non-normal. Consider as a single sample the differences Z_i (of sales tax QTAX vs ALFIN) of survey data collected from agency i. The Shapiro-Wilk test of normality (as shown in Pappas**[2]**) yields a p-value less than .10, confirming that the Z_i are not normally distributed.

We use the signed rank test since we have non-normal paired data. Table 3 reflects results of statistical tests designed to detect meaningful differences (at the α=.10 significance level) in QTAX and ALFIN data reported. The sample size for T09 and T40 are different because not all local governments tax personal income—the T40 sample includes only units that levy this tax.

Table 3. Test Results for T09 and T40 in the QTAX vs ALFIN

Sample Size (n)	Signed Rank Statistic (T)	p-value	HL-Shift	90% CI for Shift	Significant (S) Not-Significant (NS)

(T09) 2013 717	50178	0.0001	-4.3M	(-5.3M, -3.4M)	(S)
(T40) 2013 131	-897	0.0389	-1.6M	(-3.5M, -0.3M)	(S)
(T09) 2014 752	739	0.9014	0.04M	(-0.61M, 0.63M)	(NS)
(T40) 2014 144	1687	0.0006	3.7M	(1.6M, 6.0M)	(S)
(T09) 2015 784	16220	0.0105	1.0M	(0.35M, 1.8M)	(S)
(T40) 2015 140	1916	0.0001	5.1M	(3.1M, 8.6M)	(S)

The estimated difference (or shift) in 2014 T09 data collection in QTAX relative to ALFIN is approximately .04M (or $40,000). However, given the p-value (.9014) of the associated test we find that there is insufficient evidence to conclude that the survey shift is significantly different from zero. In fact, we are 90 percent confident that the true difference lies within the interval (-.61M, 0.63M)---which includes zero. A test of any value outside the interval (e.g. test if difference in QTAX data reported vs ALFIN is 1.2M) would yield a result that is statistically significant. For both 2013 and 2015 data collection T09 in QTAX is significantly different than in ALFIN---since a test for a zero (difference) is outside both the 2013 T09 interval (-5.3M, -3.4M) and 2015 T09 interval (0.35M, 1.8M).

Similarly, we find a test (of zero difference) for T40 data collected in QTAX vs ALFIN to be significant in all years studied. A test value of zero difference is outside the T40 confidence intervals for 2013, 2014, and 2015.

5. Conclusion

We initiated this study to assess the degree of comparability in Sales Tax (T09) and Individual Income (T40) Tax data collected in the ALFIN and QTAX. The question of whether there is consistency between the surveys' data collection is conclusive. We found statistically significant differences in T09 and T40 data collected in QTAX vs ALFIN in nearly all instances profiled.

We have just begun to study issues related to the comparability of data collected in ALFIN and QTAX. There will be further inquiry as to the reason for the differences in collection. A full enumeration of both QTAX and ALFIN in Census of Government years may provide important insight. The QTAX response rate must be improved, as the presence of fewer missing values may further close the gap. Finally, we will explore developing a composite estimate from Annual and Quarterly estimates.

References

1. Conover, W. J. (1999), Practical Nonparametric Statistics, 3rd Ed., John Wiley & Sons, Inc. pg 111-257.
2. Pappas, P.A. and DePuy, V., An Overview of Non-parametric Tests in SAS: When, Why, and How. http://analytics.ncsu.edu/sesug/2004/TU04-Pappas.pdf.
3. Non-Parametric Univariate Tests: Paired data. www.unm.edu/~marcusj/PairedNP.pdf
4. Quarterly Summary of State & Local Taxes. https://www2.census.gov/govs/qtax/qtaxmethodology_updatedimp.pdf
5. State & Local Government Finance website. https://www.census.gov/govs/local/ http://www.itl.nist.gov/div898/handbook/index.htm. Sect 7.2.1.3.
6. Chapter 7 ELSTAT – SAS. https://www.sas.com/storefront/aux/en/spelstat/62097_excerpt.pdf
7. Geyer, Charles J., Stat 5102 Notes: Nonparametric Test and Confidence Interval. http://www.stat.umn.edu/geyer/old03/5102/notes/rank.pdf.
8. Han, L.H., Calculating the point estimate and confidence interval of Hodges-Lehmann's median using SAS software. (SUSAG 2008: Proceedings of the SouthEast SAS Users Group, St. Pete Beach, FL). http://analytics.ncsu.edu/sesug/2008/ST-154.pdf
9. Hollander, M. and Wolfe, D.A. (1973), Nonparametric Statistical Methods, New York: John Wiley & Sons, Inc. pg 26-57.
10. NIST/SEMATECH e-Handbook of Statistical Methods. http://www.itl.nist.gov/div898/handbook/index.htm. Sect 7.2.1.3.
11. Serfling, R.J., Inference Procedures for One Sample and Paired-Data Location Problems. http://www.utdallas.edu/~serfling/3332/OneSampleLocation.pdf