

**Preliminary Results of a
Comparison of the Sample Median and Weighted Mean as Estimators of the Population
Weighted Mean Assessment/Sales Ratio**

by

Robert C. Denne and Robert J. Gloude mans

*This is a slightly modified version of a paper prepared for the IAAO Assessment Standards
Committee on March 31, 2005*

We conducted an extensive empirical analysis to compare the relative efficacy of the sample median and the sample weighted mean as estimators of the population weighted mean assessment-to-sale price ratio and, by implication, the total market value of a property class, as is required for equalization purposes.

We analyzed data sets from three sources: four from the State of Idaho, two from a major metropolitan county in the U.S., and the remainder (comprising 61 different datasets) from an anonymous Midwestern state with township-based assessment administration.

Idaho Data Sets

This analysis used the data provided by Alan Dornfest, Alex Chizewsky, and Pete Davis (DCD) from their study of assessment ratio data in Idaho as reported in the “Alternative Methods of Addressing Ratio Outliers” reported in *Journal of Property Tax Assessment and Administration*, vol 1, no 4 (2004).

The data set included 21,092 sales from all Idaho counties with a median ratio of .960, weighted mean of .944, COD of 9.0, and PRD of 1.009. DCD treated the sample as a “pseudo population” and drew 50 samples of 10, 17, and 50 sales to test various outlier trimming strategies. We followed the same approach comparing the accuracy of the sample median and weighted mean as estimators of the known population weighted mean (.944), except that we drew more samples and added sample sizes of 30 and 100.

We drew 50,000 samples without replacement of each size (250,000 samples in all) and computed and compared the sample medians and weighted means to the known population weighted mean. The average absolute errors over all 50,000 samples were as follows (the estimator with the lower average error is highlighted):

	10	17	30	50	100
Median	0.0299	0.0253	0.0210	0.0184	0.0166
Wtd Mn	0.0333	0.0266	0.0210	0.0166	0.0121

The population is characterized by good dispersion (COD of 9.0) and rather good vertical equity (PRD of 1.009). The median is the superior estimator for small samples, the measures produce

identical results at a sample size of 30, and the weighted mean is the superior estimator for larger samples (50 and 100). Neither estimator consistently outperforms the other.

In addition, DCD constructed a “subpopulation” of 1,000 sales that they used to test various outlier trimming methods for purposes of both direct and indirect equalization. The subpopulation used for indirect equalization had a median ratio of .966, weighted mean of .950, COD of 9.3, and PRD of 1.010. From this subpopulation DCD drew 50,000 samples of 10, 17, 30, and 50 to test various trimming scenarios. We also drew 50,000 samples of each of their sizes, added a sample size of 100, and again compared the median and weighted mean as estimators of the population weighted mean (.950). The average errors were as follows:

	10	17	30	50	100
Median	0.0299	0.0259	0.0215	0.0192	0.0172
Wtd Mn	0.0331	0.0267	0.0205	0.0161	0.0113

The relative performance of the statistics on the subpopulation is similar to that obtained on the population of all sales. The median provides a more accurate estimator of the population weighted mean for small samples and the sample weighted mean is superior for larger samples (30+).

In addition, we selected two additional subpopulations in the DCD study for additional analysis. One is school district with 1,758 sales, a median ratio of .919, weighted mean of .899, COD of 11.0, and PRD of 1.022. The other is a property category with 1,145 sales, median ratio of .935, weighted mean of .913, COD of 13.5, and PRD of 1.029. These two subpopulations introduce more dispersion and price-related bias. In the second case, the PRD is only in borderline compliance with IAAO guidelines. Results for the two subpopulations are as follows:

School District: N = 1,758, Median = .919, Wtd Mean = .899, COD = 11.0, PRD = 1.022

	10	17	30	50	100
Median	0.0372	0.0317	0.0264	0.0234	0.0213
Wtd Mn	0.0413	0.0329	0.0257	0.0202	0.0144

Property Category: N = 1,145, Median = .935, Wtd Mean = .913, COD = 13.5, PRD = 1.029

	10	17	30	50	100
Median	0.0441	0.0368	0.0298	0.0263	0.0238
Wtd Mn	0.0458	0.0362	0.0280	0.0220	0.0155

As might be expected, the weighted mean begins to emerge as superior in the presence of increased vertical bias, although the median retains its advantage for the smallest samples.

Vacant and Commercial Data Sets - Large Metro County

We applied the same testing scheme to vacant and commercial data used in a ratio study for a major metropolitan county. The data were edited to remove invalid sales and outliers. As is typical for vacant and commercial properties, both data sets are characterized by much wider dispersion than the residential data sets described above (CODs of 22.4 and 22.9). The PRD for vacant land is out of compliance with IAAO standards. Results were as follows:

Commercial Data: N = 600, Median = .973, Wtd Mean = .964, COD = 22.9, PRD = 1.018

	10	17	30	50	100
Median	0.0749	0.0600	0.0443	0.0343	0.0244
Wtd Mn	0.0997	0.0840	0.0686	0.0552	0.0397

Vacant Land: N = 515, Median = 1.005, Wtd Mean = .979, COD = 22.4, PRD = 1.041

	10	17	30	50	100
Median	0.0729	0.0554	0.0424	0.0343	0.0301
Wtd Mn	0.0942	0.0760	0.0597	0.0471	0.0331

With wider dispersion, the median outperforms the weighted mean as an estimator of the population weighted mean at all sample sizes, even for vacant land, where there is a sizeable differences between the sample median and weighted mean and the PRD is 1.041. The differences in relative accuracy are much wider than for the more homogeneous residential populations.

Anonymous Midwestern Townships

Almy, Gloudemans, Jacobs & Denne has completed a ratio study for 61 townships with at least 500 residential sales in an anonymous Midwestern state dominated by township assessing. Again, the data were edited to remove invalid sales and outliers. The appendix presents key sales ratio statistics for the 61 townships. We conducted similar simulation studies for these 61 residential “populations”, drawing 50,000 samples of 10, 17, 30, 50, and 100 each (61 x 5 x 50,000 = 15.25 million samples in all). The number of townships in which each of the two estimators emerged as the more accurate estimator of the population weighted mean is shown below (a tie is defined as an absolute difference in average errors of less than .001):

	10	17	30	50	100
Median	32	27	26	23	15
Wtd Mean	24	28	27	28	32
Tie	5	6	8	10	14

As in the Idaho residential data, the median is more accurate for the smallest samples (10) and the weighted mean is more accurate for the largest samples (50+). The measures are equally good predictors for the mid-sized samples (17 and 30).

A closer examination revealed, as also suggested in the Idaho data, that the weighted mean is relatively more accurate when vertical equity is relatively poor. Of the 61 townships, 23 have PRDs that exceed the IAAO standards. For these 23 townships, the weighted mean produced the lowest relative error in the majority of cases, including sample sizes of 10. However, where PRDs comply with IAAO standards, the median produced more accurate estimates of the population weighted mean at sample sizes of 10, 17, 30, and 50 (only for samples sizes of 100 was the sample weighted mean better able to hone in on the population weighted mean than the sample median).

Conclusions

It is empirically clear that neither estimator is universally superior to the other. From a conceptual point of view, the population weighted mean is unquestionably the correct parameter to measure, and it may seem intuitively obvious that the sample weighted mean should be its best estimator. However, that statistic suffers from larger sampling error than the median, which usually can only be offset at larger sample sizes. Where dispersion is high, as in the vacant and commercial samples, this inherent deficiency may not be offset even at large sample sizes.

As a general matter, when vertical inequity is relatively severe, pulling the median and weighed mean several points apart, the weighted mean usually provides the better estimate. When vertical equity is relatively good, the median generally provides the better estimate -- except at large sample sizes.

Which measure is most likely more correct in a particular situation depends on the interplay of sample size, influence sales, dispersion, and price-related bias. We posit that the weighted mean should be used in preference to the median for indirect equalization when price-related bias can be demonstrated, which implies reasonably large sample sizes. To dogmatically argue that only one measure is acceptable in all situations is clearly unsupported.

Finally, we also note that neither the median nor the weighted mean is an ideal estimator for indirect equalization: one pays no heed to dollar amounts and one is overly sensitive to outliers. The profession should leave the door open to more robust alternatives.

Appendix – Anonymous Township Sales Ratio Statistics

TOWNSHIP	N	MEDIAN	WM	DIFF	COD	PRD
1	516	.940	.959	-.019	17.6	100.9
2	519	.999	1.005	-.006	18.3	102.2
3	565	.978	1.012	-.034	25.4	105.7
4	572	.974	.977	-.003	9.0	100.5
5	572	.997	1.018	-.021	16.0	101.2
6	575	.920	.958	-.038	22.0	103.6
7	603	.984	.991	-.007	22.2	106.8
8	631	.986	.996	-.010	18.6	104.0
9	654	.988	.989	-.001	11.9	101.2
10	719	1.004	1.006	-.002	9.0	100.8
11	720	1.000	1.004	-.004	12.3	101.3
12	727	.986	.980	.006	8.7	101.0
13	780	1.008	1.014	-.006	15.1	102.6
14	811	.949	.956	-.007	16.6	101.9
15	886	1.002	.983	.019	54.8	134.8
16	886	.963	.977	-.014	18.2	104.0
17	902	1.000	1.003	-.003	28.9	109.3
18	1009	1.147	1.135	.012	33.3	116.4
19	1011	.986	1.012	-.026	18.8	101.0
20	1022	.956	.969	-.013	16.4	99.2
21	1093	1.002	1.014	-.012	22.4	107.1
22	1117	.998	1.011	-.013	9.4	100.5
23	1139	1.084	1.089	-.005	42.9	119.4
24	1145	.989	.998	-.009	9.9	100.5
25	1194	.950	.974	-.024	21.8	104.4
26	1242	.987	.991	-.004	13.3	101.3
27	1257	.942	.959	-.017	18.0	103.4
28	1330	1.012	1.014	-.002	17.3	104.1
29	1348	.976	.986	-.010	10.2	100.3
30	1388	1.032	1.046	-.014	24.9	107.9
31	1415	.986	.987	-.001	9.8	101.0
32	1416	.992	.999	-.007	10.2	99.7
33	1430	.988	.991	-.003	7.0	101.1
34	1484	.963	.969	-.006	13.8	101.2
35	1505	1.047	1.051	-.004	16.0	102.0
36	1641	.964	.984	-.020	23.9	104.9
37	1723	.998	.994	.004	12.2	101.6
38	1747	.948	.972	-.024	16.3	102.1
39	1803	.986	.999	-.013	15.7	101.0
40	1858	.977	.983	-.006	15.5	101.9
41	2132	1.000	.999	.001	8.5	101.2
42	2210	.966	.973	-.007	11.4	101.4
43	2448	.981	.979	.002	25.5	105.3
44	2477	.989	.994	-.005	8.9	100.6
45	2552	.939	.953	-.014	14.2	102.3
46	2605	1.000	1.008	-.008	9.4	100.4
47	2660	.999	1.029	-.030	14.6	100.3
48	2732	1.011	1.011	.000	41.7	120.6
49	3138	.962	.964	-.002	12.3	101.5
50	3196	.979	1.006	-.027	21.7	103.2

TOWNSHIP	N	MEDIAN	WM	DIFF	COD	PRD
51	3481	.984	.986	-.002	8.6	101.1
52	3731	.962	.964	-.002	10.7	101.0
53	3830	.978	.983	-.005	6.3	100.2
54	3992	.963	.970	-.007	9.0	101.4
55	4803	.988	1.003	-.015	23.2	106.0
56	4907	.951	.948	.003	12.5	103.3
57	5194	.975	.985	-.010	8.5	100.2
58	5756	.921	.900	.021	16.9	105.3
59	6152	.971	.989	-.018	20.3	103.7
60	6860	.986	.984	.002	8.2	100.8
61	7216	.958	.969	-.011	16.5	103.2