# Statistical Observations on Mass Appraisal

by Josh Myers
Josh Myers Valuation Solutions, LLC.

# About Josh

- Josh Myers is an independent CAMA consultant and owner of Josh Myers Valuation Solutions, LLC. Josh has experience working at the local jurisdiction level and at the CAMA software vendor level. He holds a Masters Degree in Statistics from the University of Virginia. Josh has given about a dozen major conference presentations. His article, co-authored with Wayne Moore, PH.D, "Using Geographic-Attribute Weighted Regression for CAMA Modeling," was published in the Journal of Property Tax Assessment and Administration in 2010 and won the 2011 IAAO Bernard L. Barnard Outstanding Technical Essay Award. Josh is also a member of the IPTI/IAAO Editorial Review Board. He resides in Chesapeake, Va., is a volunteer staff member at his Church, and helps coach a high school baseball team.

# My Goals Today

- The goal today is education rather than simply presenting new research. I want to give you several tools or ideas today that you can use back home. I will cover about a dozen topics, but that means I can't do so in a lot of detail.

- My goal is not to throw a lot of formulas at you today, although I am sure that would seem impressive to you. Instead, I want to discuss these ideas more conceptually so I don't have to spend a lot of time explaining formulas.

- There will be a five minute break after one hour. I will leave 20 minutes at the end of the presentation for all questions. During the question period at the end, feel free to ask me a question about any topic that you wish, whether we cover it today or not.

# Overview

- The Importance of Software
- General Ratio Statistics
- Population vs. Sample
- The Importance of Confidence Intervals
- Vertical Inequity
- Thoughts on the Coefficient of Dispersion
- Types of Residential Mass Appraisal Models
- Pros and Cons of the AVM
- The Importance of the Use of Location in Modeling
- General Modeling Thoughts

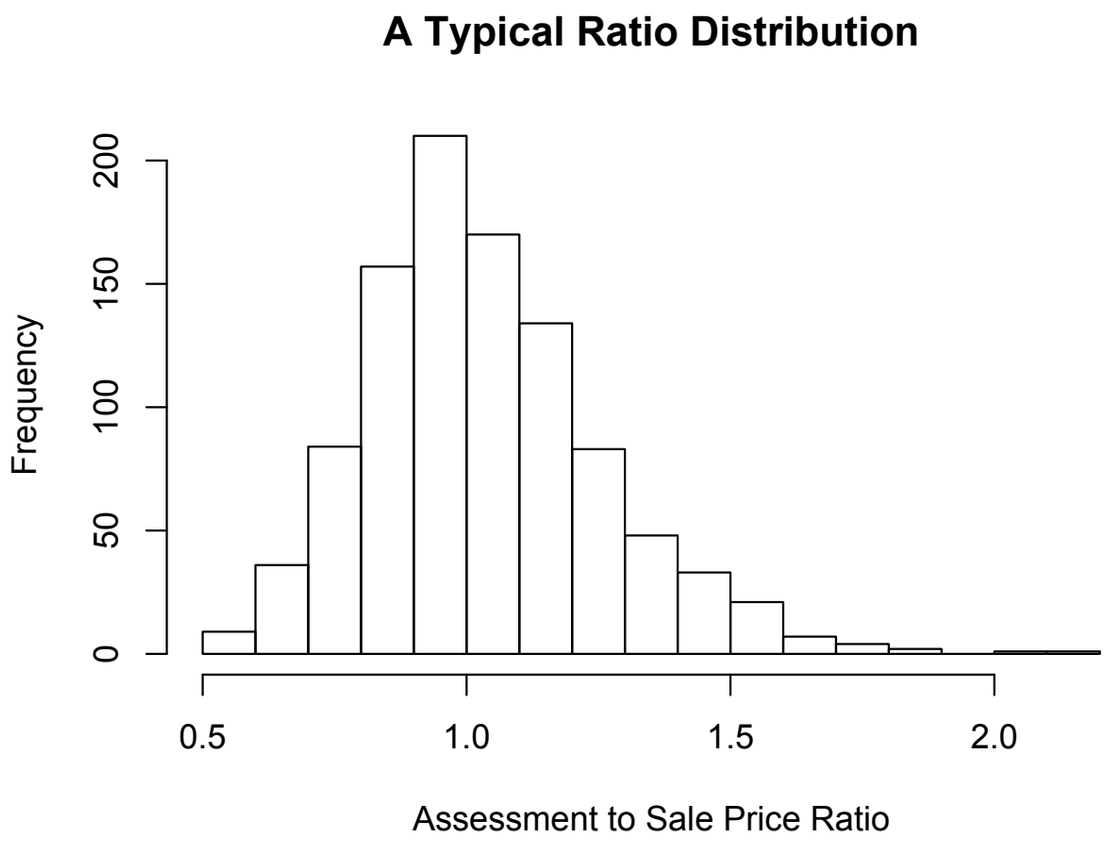
# The Importance of Software

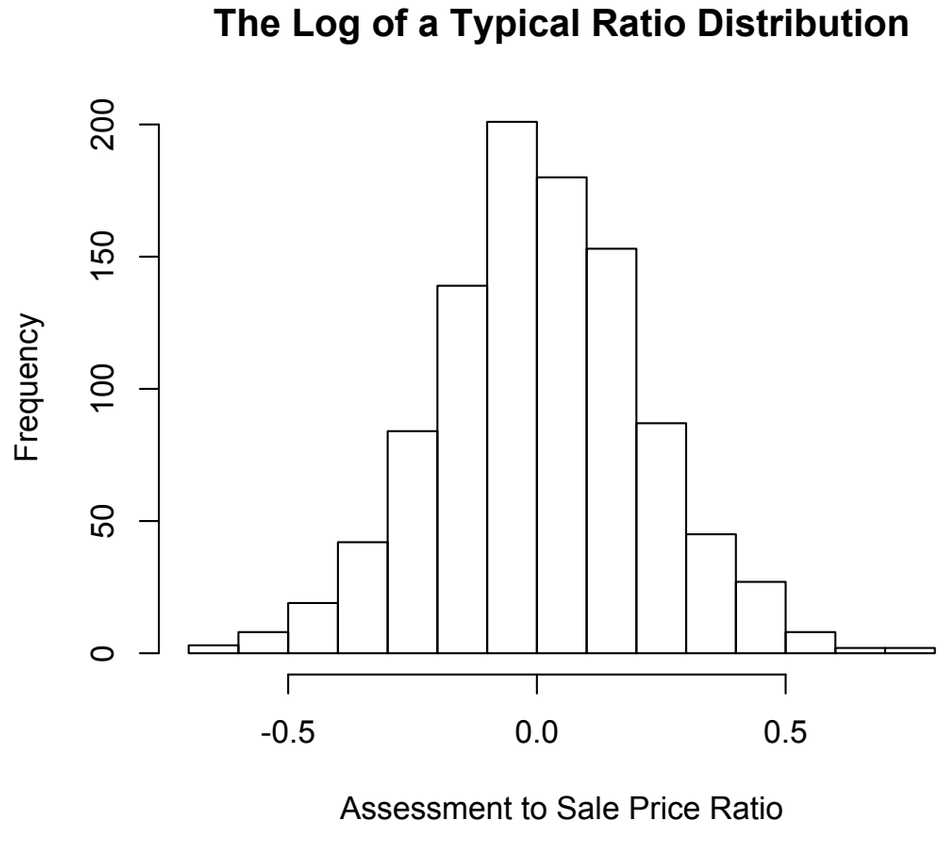
- Software is the key to changing the mass appraisal industry.
- This is because people will not implement good ideas until it is feasible to do so and only software grants them that power.
- Unless good ideas are included in software in an easy-to-use way, then they will never get adopted by the industry at large. We can present on them at conferences, publish papers about them, and place them in the standards, but without good software none of that matters.

# General Ratio Statistics: What is a Ratio?

- Assessment to Sale Price Ratio: Assessment / Sale Price
- We use ratios because they offer us a measure of how close the Assessment is to the Sale Price that is independent of the value of the Sale Price. The Ratio give us a measure of how close the assessment is to the Sale Price in terms of a percentage.
- The distribution of the Ratios usually appears to be approximately log-normal. This means that taking the logarithm of the Ratios produces an approximately normal distribution. This is useful in several different contexts.

# General Ratio Statistics: What is a Ratio?



# General Ratio Statistics: Measures of Center

- Median Ratio: Finding the middle of the distribution of ratios. The ratios are sorted, and then the middle ratio is found (the mean of the middle two ratios is taken if the number of ratios is even).

- Mean Ratio: Sum(Ratios) / N, where N is the number of ratios. The ratios are summed, and then that sum is divided by the number of ratios.

- Given the nature of the typical ratio distribution, it makes more sense to use the median rather than the mean as a measure of center. This is because the median ratio is robust (outlier-resistant). The median ratio is almost always less than the mean ratio because the typical ratio distribution is right skewed, as stated previously.

# General Ratio Statistics: Measures of Spread

- Coefficient of Dispersion (more on this later): average absolute difference between the ratios and the median ratio all divided by the median ratio. COD = 100 * Mean(Abs(Ratio - Median(Ratio)))/Median(Ratio)
- Coefficient of Variation: standard deviation of the ratios divided by the mean ratio. COV = SD(Ratio)/Mean(Ratio)
- The COD is usually a better measure of spread than the COV because it is a little more robust to outliers.
- We also have the Interquartile Range (the difference between the third and first quartiles) and the Range (the difference between the minimum and the maximum).
- Graphical Measures of Spread: box-plots and histograms.

# Ratio Distribution and Outliers

- How do you detect outlying ratios?
- One method is to take the log the ratios and use the traditional formula: (Q1 - 5*IQR, Q3 + 5*IQR) where ratios outside of this interval are considered outliers, Q1 is the first quartile, Q3 is the third quartile, and IQR = Q3 - Q1.
- Another method is to not log the ratios and instead use an outlier formula that expects the interval to be uneven.

# Population vs. Sample

- A population is the full set of data. A population has to be strictly-defined under certain conditions by the analyst based on the situation at hand.
- A sample is a subset of the population that was collected according to a certain protocol.
- In mass appraisal, one example of a population is the set of all single-family homes in a given county and one example of a sample from that population is those single-family homes in the county that sold in the last year.
- One of the primary goals of statistics is to use the sample to make conclusions about the population. The sample is used to estimate the population. The assumption is that the sample is representative of the population.

# Confidence Intervals

- There are two important ways to describe a population statistic: a point estimate and an interval estimate. Both a point estimate and an interval estimate are calculated from the sample.

- A point estimate is the best guess of the true population statistic derived by calculating that statistic based on the sample. An example of a point estimate is the Mean or Standard Deviation.

- It is highly unlikely that the true population statistic will be the same as its point estimate. Therefore we need some way to tell what range the true population estimate is likely in. This is the interval estimate.

- The confidence interval is one interval estimator we use.

# Confidence Intervals

- We calculate confidence intervals for a given point estimator by using a formula or by bootstrapping (a computer-intensive resampling process).
- Without a confidence interval, a point estimate doesn't mean all that much.
- An example from politics.
- What this means for us: we need to calculate a confidence interval for every point estimator we use. For example, we shouldn't calculate median ratios without doing a confidence interval for the median ratio.

# Vertical Inequity

- We are usually talking about market value vertical inequity when we say vertical inequity.
- The question is whether Ratios are equal across the range of market values (or between strata of market values).
- There are many methods to detect vertical inequity like Price-Related Differential (equal to the mean ratio divided by the sale-price weighted mean ratio), regression-based methods that produce statistical tests (the Gloudemans Price-Related Bias Method and the Clapp Instrument Variable Method are the best), and graphical measures (the McMillen Quantile Regression Plots are the best), among others.

# Vertical Inequity

- All of the currently available methods for detecting vertical inequity display some sort of bias – either toward regressivity or progressivity – depending on either the quality of the assessed value, the quality of the sale price, or both, the choice of the proxy for market value, and the formulation of the method itself.

- There are two solutions to potentially fix this bias problem. These two solutions are not methods of detecting vertical inequity in themselves but rather ways to replace the proxy for market value that is used in the existing methods.

# Vertical Inequity

- One solution is to produce a second estimate of market value, using a different modeling methodology, and use that as the proxy.
- A second solution is to estimate the qualities of the assessed value and the sale price and then use the appropriate linear combination of the two of them as the proxy.
- This was all explained in depth at my presentation at the 2013 GIS/CAMA Technologies Conference. You can access this presentation on my website.

# Vertical Inequity: Pros and Cons of the PRD

- Most people are content to just use the PRD, but it has been under attack recently, sometimes unfairly.
- Pros: Easy to Calculate and Communicate, Widely Accepted (not a great reason)
- Cons: Overly simplistic in that only the overall lean to regressive or progressive is measured; It is hard to attach meaning to relative values of the PRD; It cannot prove whether a certain amount of Vertical Inequity is statistically significant (unless it is converted to a statistical test via a bootstrap confidence interval, as it is possible to do); It is affected by outliers (but basically all methods are).
- Sounds bad, but competing methods also have problems...

# Other Kinds of Vertical Inequity

- There are other kinds of vertical inequity beside market value vertical inequity. This would be inequity on a different scale beside market value. For example, grade - as you go up the grade levels, are properties assessed equivalently?
- Even though this kind of vertical inequity is bad in itself, it may in fact be the cause of market value vertical inequity. Therefore, it is important to investigate all types of vertical inequity to see where the root problem lies.
- There are different ways to detect this, with most of them being quite similar to that of market value.

# More Thoughts on the COD

- The COD is important because it is our primary measure of equity. Everybody uses it. That is why I thought it important to discuss it in more detail.

- We talked about the importance of confidence intervals earlier for point estimators (the COD is a point estimator). There are two basic ways to calculate confidence Intervals for COD: use the Bonnet-Seier Formula (Bonett and Seier, "Confidence Interval for a Coefficient of Dispersion in Nonnormal Distributions", Biometrical Journal 48, 2006) or run a Bootstrap Confidence Interval using computer methods.

- The Bonnet-Seier confidence interval is superior because its ease of calculation and its accuracy.

# More Thoughts on the COD

- The COD is not exactly the typical ratio error from the median ratio. That is the median absolute percentage error. The COD is equal to the median absolute percentage error under the conditions than the median ratio and the mean ratio are both equal to 1.
- The COD is not resistant to outliers. I am not sure why the COD is this strange combination of mean and median.
- Situations with bad quality sale prices or bad overall data quality cannot produce good CODs, so high CODs do not necessarily mean that your appraisal model is poor (a performance analysis, breaking down your results and comparing your results to a proven model, can help say for sure whether it is your model that is the problem).

# More Thoughts on the COD

- The COD is actually a measure of both horizontal and vertical inequity, not just a measure of horizontal equity as some often claim. But, why is this true?

- First, we see that the COD is proportional to the Variance under a normal distribution.

- Second, we see that a regression model of the form f(Ratio) = f(Proxy) can be modeled in order to determine the nature of the market value vertical inequity.

- Third, we know that the total sum of squares for any regression relationship can be partitioned into two parts, the error sum of squares and the regression sum of squares, in this way: SST = SSE + SSR.

# More Thoughts on the COD

- Fourth, by definition the SSR/SST is equal to the proportion of the variation, as measured by the model, that is due to vertical inequity. Likewise, the SSE/SST gives us the proportion of the variation, as measured by the model, that is due to horizontal inequity.

- Fifth, the SST = SSE + SSR also partitions the variance, and so therefore it also partitions the COD under a normal distribution since the COD is proportional to the variance under a normal distribution. Therefore, SSE/SST can also be used to find the percentage of the COD that is due to horizontal inequity and the SSR/SST can also be used to find the percentage of the COD that is due to vertical inequity.

# More Thoughts on the COD

- Sixth, this means that the COD has a component that is due to vertical inequity and a component that is due to horizontal inequity.
- Finally, therefore, the COD is not a pure measure of horizontal inequity and it should not be treated as such.

# BREAK

- Ahhhhhhhhhhh.....

# Types of Residential Models

- Cost; Locationally-Adjusted Cost (TCM or ProVal Method)
- Ordinary Least Squares (MRA); Segmented Ordinary Least Squares (Segmented MRA)
- Feedback
- Non-Linear Models
- GWR/GAWR
- Comp Sales
- Combo of MRA and Comp Sales; Spatial Regression
- Response Surface; Kriging

# To AVM or not to AVM?

- What are the benefits and drawbacks of using a Statistical Mass Appraisal Model (AVM) instead of a more "manual" approach? (Remember, the Locationally-Adjusted Cost Approach is not really "manual" to begin with).

- Here are a few things to consider. First, models make mistakes just as humans do, so they are not infallible. Second, people love to have the perception of control, even if the results are not optimal.

# To AVM or not to AVM?

- Pros of the AVM: has the potential to be highly accurate; usually results in huge time savings; usually removes many of the problems resulting from human biases.
- Cons of the AVM: usually hard to explain to appraisers and/or taxpayers; can produce some wildly-incorrect estimates; results are only as good as the modeler
- The Cons of the AVM are greatly exasperated when they are implemented improperly or without the sufficient involvement of the appraisal staff. A knowledgable modeler is needed to create and manage the AVM. Staff should be trained in how to understand the AVM and their key role in the process should be clearly explained. After the AVM has been run, results should be reviewed by appraisal staff to check for wildly-incorrect estimates.

# Use of Location in Modeling

- It is highly important for accuracy to incorporate location in Mass Appraisal Modeling. Remember, location, location, location! This is important for all types of models.
- Example: Charlottesville vs. Chesapeake

# Use of Location in Modeling

- But, do GIS parcel coordinates add any locational explanatory power over simply having locational binary variables (like the neighborhood, sub-market, or market area for the parcel)?
- Wayne Moore and Josh Myers did a research project on this in 2010. In short, the project compared the results of several top modelers, with only neighborhood binary variables used to account for location, to GWR and GAWR models, with parcel coordinates used to account for location.
- Here are some results from that research.

# Use of Location in Modeling

*Summary of Equity Findings for AVM Types with the Research Data Set*

| | Overall findings | | | | *QMR* findings by quintile | | | | | COD findings by quintile | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AVM | Med | COD | PRD* | VEI | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| AEP | 0.93 | 9.95 | 1.01 | 6.3 | 0.91 | 0.94 | 0.97 | 0.96 | 0.91 | 10.05 | 9.34 | 9.33 | 9.19 | 11.11 |
| COST | 0.93 | 14.91 | 0.99 | 10.8 | 0.88 | 0.92 | 0.98 | 0.98 | 0.96 | 12.10 | 14.62 | 15.00 | 14.40 | 15.56 |
| GAWR | 0.98 | 7.67 | 1.01 | 4.9 | 1.01 | 0.97 | 0.97 | 0.98 | 0.97 | 8.48 | 6.51 | 6.55 | 6.13 | 10.14 |
| GWR | 0.99 | 8.39 | 1.01 | 6.6 | 1.03 | 0.98 | 0.98 | 0.98 | 0.96 | 8.71 | 7.93 | 7.44 | 7.17 | 9.72 |
| MRA | 0.99 | 9.91 | 1.03 | 13.7 | 1.00 | 1.01 | 1.04 | 0.99 | 0.90 | 8.93 | 8.94 | 8.92 | 9.11 | 11.71 |
| NoRCN | 0.98 | 7.93 | 1.01 | 5.5 | 1.02 | 0.98 | 0.97 | 0.98 | 0.96 | 8.39 | 6.98 | 6.70 | 6.45 | 10.39 |
| TCM | 0.94 | 10.06 | 1.01 | 4.5 | 0.95 | 0.95 | 0.97 | 0.93 | 0.92 | 10.44 | 8.95 | 8.78 | 9.50 | 12.29 |

# Use of Location in Modeling

*Comparison of Horizontal Equity Performance between AVMs using the Research Data Set*

| AVM Test group | Test group sample size | $M_{COD}$ | Tukey-Kramer multiple-comparison test results COD performance is different from: |
|---|---|---|---|
| GAWR | 31 | 7.559 | MRA, AEP, TCM, COST |
| NoRCN | 31 | 7.780 | MRA, AEP, TCM, COST |
| GWR | 31 | 8.239 | AEP, TCM, COST |
| MRA | 31 | 9.549 | GAWR, NoRCN, COST |
| AEP | 31 | 9.641 | GAWR, NoRCN, GWR, COST |
| TCM | 31 | 9.731 | GAWR, NoRCN, GWR, COST |
| COST | 31 | 14.447 | GAWR, NoRCN, GWR, MRA, AEP, TCM |

# Use of Location in Modeling

- First, as an aside, the project was able to conclude that there were statistically significant differences among AVM methodologies. Therefore, it is false to believe the choice between models does not matter.

- In addition, results of the project actually conclude that the GAWR model, using attribute weighting in addition to distance weighting, is a statistically significantly better model than all of the models without parcel coordinates. The GWR model barely misses this distinction, however, so, technically, parcel coordinates alone were not proven, with this dataset, to be statistically significantly able to improve performance.

# Use of Location in Modeling

- However, the test that we used was not as powerful as the one I now use for my solo work. With that more powerful test, GWR would have been statistically significantly able to improve performance, meaning that the addition of parcel coordinates could be said to improve performance.
- The results show that GAWR, with or without using the RCN, has the best model COD, except it can't be said to be statistically significantly better, under our test, than GWR.
- On the next slide, we see GAWR performing best in comparison to a host of other models, produced by both practitioners and academics, using the large Fairfax Test Dataset.

# Use of Location in Modeling

*Equity Findings with the Fairfax County Data (January 1972 to June 1991)*

| | | Absolute Value Percentage Error | | |
|---|---|---|---|---|
| Model | Mean | 25th Percentile | Median | 75th Percentile |
| *Myers GAWR* | *10.6* | *3.0* | *6.4* | *11.8* |
| Gloudemans | 11.8 | 3.7 | 7.8 | 14.1 |
| Case | 11.8 | 3.7 | 8.0 | 14.1 |
| OLS (MRA) | 12.6 | 4.0 | 8.4 | 15.8 |

# Use of Location in Modeling

- A more in-depth presentation of the Moore and Myers research is available on my website.
- Conceptual explanation of GWR and GAWR.
- One Caveat:  no model performs best with every dataset, so different models must be tested and compared in each case.

# How do I Tell which Model is Better?

- Usually, people only compare CODs of competing models to see which is better, but does this make sense?
- No, for several reasons: model coefficient values for the competing models should also be compared to see if they are equally interpretable, a statistical test for the equality of the CODs is needed to tell if the CODs for the competing models are statistically significantly different (remember the problem with point estimators?), a more detailed comparison of the different types of inequity in the competing models must be undertaken, the competing models must be evaluated to see how equitably they apply to non-sales, etc.

# How do I Tell which Model is Better?

- It is important to compare model results under realistic conditions, namely, the forward prediction of sale prices, because that is what we do in mass appraisal. For this, you will need to establish a prediction sales set that is sequentially after the model-building sales set.

- It is also important to compare models using the exact same conditions (same model-building sales set, same prediction sales set, same rules)

# Problems with Land Modeling

- There is usually a lack of land sales.
- Not much data is available.
- Software programs don't always have great land modeling tools built in.

# Some Other Uses of Statistical Modeling

- Time trends for ratio studies.  It is best to get time trends by location and property type.
- Depreciation factors for cost models.

# Using Subjective Factors in Modeling

- Quality appraisers are critical to the application of subjective factors to real property.
- It is important to have consistent applications of subjective factors in your offices because they can lessen the effectiveness of any mass appraisal model, especially AVMs.
- In modeling, subjective factors are usually qualitative variables and often times have to be converted to a set of binary variables or linearized as a continuous variable.

# Questions

- Ask me anything - it can be related or unrelated to something we said today.

- Thank you for attending!!!

# How to Get a PDF Copy of this Presentation

- Go to my website: www.joshmyersvaluationsolutions.com
- This web address is also available on my business card (available up front).