# COMMENTS REGARDING VALUATION METHODOLOGY

The purpose of this memo is to lay out what I believe to be John Ryan's valuation methodology. Although I am fairly confident, I may not be exactly right. Even if I am not exactly right, I think I am close enough to develop specific questions that will get the full story. I hope Mr. Ryan is prepared to explain his methodology in a complete and transparent way.

**Background.**
John Ryan told Nanette Albanese in a June 15, 2016, email that, "The final values used in this revaluation are determined via an ALGORITHMIC process that anyone can replicate…. [T]he process is completely transparent."

Well, perhaps "anyone" can replicate it they know what it is, but Mr. Ryan has consistently failed to explain it.

During the November 17, 2015, meeting of the Committee of the Whole, I had a dialogue with Mr. Ryan. I told him,

> I just [wonder] what you are doing. With Tyler's model, once I had the sales data, I could reproduce the model exactly using Excel, using the standard multiple linear regression. I would like to be able to do that once your model is complete. So, I'm anticipating, I'm trying to understand, what your model is, or your models, and what the inputs are, and what the mathematical techniques…

He replied, "That's a fair request, and when the time is appropriate, I'm, we'll be more than happy to share that information."

So far, nothing. On July 18, I submitted a FOIL that requested, in part

> Documents that show explicitly the derivation of all "coefficients" and "multipliers" used in the valuation of single family residences, as presented in the **Scarsdale Valuation Sheets 2016 - For Web.pdf** document that is posted on the website….
>
> This request is not limited to general descriptions of the process. I am requesting spreadsheets and/or other documents that show explicitly and mathematically how the "coefficients", "multipliers" and land "amounts" were derived from basic inputs.

Village Hall responded to most of the rest of my FOIL, but the response here was that the request "is for data Ryan has not provided the Village, therefore, no such record exists."

The Uniform Standards of Professional Appraisal Practice ("USPAP") Standard 6, Rule 6-8 states in part that,

> Each written report of a mass appraisal must:
>
> …

(k) describe and justify the model specification(s) considered, data requirements, and the model(s) chosen;

> Comment
> …The report must include a discussion of the rationale for each model, the calibration techniques to be used, and the performance measures to be used.

….

(m) describe calibration methods considered and chosen, including the mathematical form of the final model(s); describe how value conclusions were reviewed; and, if necessary, describe the availability of individual value conclusions; …

**My Best Guess**

While I hope Mr. Ryan soon decides to honor his commitments and obligations, I have decided to present my best guess as to his methodology. To cut to the chase, I believe Mr. Ryan used the "Hybrid Appraisal Models" procedure of a statistical software product known as NCSS[1]. An NCSS document that explains this procedure can be found here, and I have provided a copy as Exhibit A. I believe this is a roadmap to Mr. Ryan's methodology.

I will explain in this section why I think this is the software that he used. In following sections, I will discuss the methodology built into the software and its implications.

First, Mr. Ryan referred positively to NCSS statistical software in a September 19, 2014 email to Michael Thrapp (an IT person) with cc to Ms. Albanese.[2] This was even before he had the contract with Scarsdale.

> NCSS, the statistical software that provides the capability to generate values, has two procedures that allow one to "force" otherwise statistically insignificant variables into a valuation model.

This sounds good – magically getting significance out of insignificant variables.

Second, the email includes a complicated formula that fairly well resembles a complicated formula that appears in the NCSS document. You do not have to understand these formulas to see the resemblance. This is the formula that Mr. Ryan's email provides as an "example of what a simple valuation model looks like".

| | |
|---|---|
| Overall | Date^(B1) *(B2)^(Neighborhood=4) *(B3)^(Neighborhood=6) |
| Land | LotAdjusted^(B8) *((B9) *LotSize) |
| Building Grade | Linear^(B4) *((B5) *SqFt1stFlr +(B6) *SqFtOthFlr +(B7) *Baths) |
| Garage | (B10) *GarageSqFt |

[1] The acronym "NCSS" refers to "Number Cruncher Statistical Software".

[2] A copy is provided as Exhibit B. I have converted the email to pdf for electronic versions of this analysis.

This is the formula that appears on Page 4 of the NCSS document[3] as, "[a]n example of the form of a full hybrid appraisal model".

$$\begin{aligned} Sale\ Price = {} & b_1^{HERRICK} \times b_2^{SKYGLADE} \times & [Overall] \\ & (SoilQuality^{b_7} \times (b_8 \times LandSQFT + b_9 \times SLOPE) + & [Land] \\ & QUAL^{b_3} \times (b_8 \times SQFT + b_5 \times KitchenUpgrades + b_6 \times Age) + & [Building] \\ & ShedQuality^{b_{10}} \times (b_{11} \times ShedSQFT - b_{12} \times ShedAge)) & [Addition] \end{aligned}$$

I see many similarities.

- Mr. Ryan stacks "Overall", "Grade", "Building" and "Garage" on the left side. The NCSS example stacks "Overall", "Land", "Building" and "Addition" on the right side.
- Same structure.
- Same types of combinations of additions, multiplications and exponents.
- Both formulas are someone's logical idea of the components of the sale price of a real estate property.
- Mr. Ryan puts "Neighborhood = 4" and "Neighborhood = 6", where NCSS puts "HERRICK" and "SKYGLADE". Herrick and Skyglade seem like they could be examples of neighborhoods.
- Mr. Ryan uses familiar Scarsdale suburban terms like "SqFt1stFlr" and "Baths", where NCSS uses "SoilQuality" and "ShedSQFT", indicating a more agricultural situation. Either way, these are the known property characteristics
- Mr. Ryan uses B1, B2, etc. where NCSS uses $b_1$, $b_2$, etc. Either way, these are the unknowns – the numbers that you want to solve for.

I will more into the theory of this in the next section. *The basic point for now is simply that the two complicated formulas, each described as an "example", are similar in form and structure. It is unlikely that all these similarities are coincidental.*

Before getting to the third point, I will digress and explain the term "hybrid appraisal model." If you have traced through the derivation of your property's value in the "2016 Residential Valuation Detail Sheet", you will have seen that the value is built up through a fairly lengthy combination of additions and multiplications. This combination of additions and multiplications is a hybrid model. A hybrid model is more complicated than models that are just additive or just multiplicative. The *IAAO*[4] *Standard on Automated Valuation Models (AVMs)* (2003), provided

[3] All page numbers references are to the pdf pages. The page numbers that actually appears on the NCSS document are messed up.

[4] IAAO = International Association of Assessing Officers.

here and as Exhibit C, explains the three types of models on Pages 9-10, in the section titled, Direct Market Models.[5]

Thus, the third point is that everything here involves hybrid appraisal models. The formula embedded in our detail sheets is a hybrid model. Although his email does not use the term "hybrid", he attached a "Valuation Detail Example" from another project, provided as Exhibit D. This clearly resembles the Scarsdale sheets and a tracing through of the calculation shows that it also used a hybrid formula. And the NCSS document is titled "Hybrid Appraisal Models" and describes the how to use the software to fit these models.

Some additional points

- I have not been able to find any other vendor offering statistical software that supports hybrid models in mass appraisals.

- NCSS is an IAAO "Industry Partner".

**NCSS Methodology: How do you Fit Hybrid Models and Why do you need Special Software?**

The point here is not to get too technical about the NSCC methodology. I assume that any interested person with math aptitude can read the NCSS document themselves and will understand in detail how the approach differs from ordinary least square multiple regression, as was used by Tyler in 2014.

The point here is to describe the NCSS methodology in a more simple way, just deeply enough to convey the risks and issues associated the method.[6]

It is helpful to start with some basics. A "model" is simply a mathematical representation of reality. A simple model (a simple conceptualization) is that real estate sale prices are a multiple of the square footage plus a constant. As a formula, you would say

$$Sale\ Price = b_0 + (b_1 \times SquareFeet).$$

$b_1$ is the multiple of the square footage and $b_0$ is the constant. These can be referred to as the "parameters" of the model.

The next step is to use real data – real sales with real prices and real square footages – to derive numbers for $b_0$ and $b_1$. This is known as "fitting" the model.[7]

---

[5] Page 2 shows that John F. Ryan, CAE, served as a reviewer for the Standard.

[6] None of this is intended as a criticism of NCSS. This appears to be fantastic software supporting many traditional and new statistical techniques applicable to many realms. NCSS is responsible about providing warnings, as discussed below.

[7] The appraisal industry refers to this as "calibrating" the model. USPAP Standard 6, Rule 6-4(c), Comment: "Calibration refers to the process of analyzing sets of property and market data to determine the specific parameters of a model."

More than 100 years ago, mathematicians developed a technique that could solve for $b_0$ and $b_1$ in our simple example. The answer might be something like this.

$$\textit{Sale Price} = \$200{,}000 + (\$300 \times \textit{SquareFeet})$$

We can refer to this technique as simple linear regression. Scarsdale High School has courses that explain how to do this. Excel has functions that will do this. Not only can you do this, but there are associated tests to determine how good the fit is, or how bad, and how much is not explained just by square footage, etc.

A more advanced technique (multiple linear regression) can fit a model, *i.e.,* a formula, that has more data inputs and more parameters. For, enhance our simple model by also considering the lot size. In addition to $b_0$ and $b_1$, you want to find $b_2$ in the following formula.

$$\textit{Sale Price} = b_0 + (b_1 \times \textit{SquareFeet}) + (b_2 \times \textit{Acres})$$

And based on your data, the result could be something like this.

$$\textit{Sale Price} = \$150{,}000 + (\$250 \times \textit{SquareFeet}) + (\$750{,}000 \times \textit{Acres})$$

This can be extended to many more inputs (age, number of bathrooms, condition of the house, neighborhood, etc.), solving for more unknowns ($b_3$, $b_4$, etc.). Tyler used this for its "model estimate." Tyler's model was more complicated than the simple examples above, but it could still be solved with the multiple linear regression function in Excel, quickly.

But at some point, peoples' imaginations as to a *really good* model/formula went beyond what the traditional methods can solve for. Even before I was alerted to the NCSS software and document, and before I saw the use of the term "hybrid" in this context, it was clear to me that Mr. Ryan's formula, as implied in the detail sheets, combines the additive and multiplicative elements in ways that the traditional techniques simply cannot handle.

The NCSS document explains a relatively new type of technique. This technique is basically an advanced, computer assisted form of trial and error. NCSS actually describes it as "(intelligent) trial and error."

> The coefficients … of a hybrid appraisal model are estimated from a (hopefully large) number of properties where the attribute values are known and the sale price is known. Whereas the coefficients in additive models (and some multiplicative models) may be estimated using multiple regression analysis (a closed form solution), the coefficients in hybrid models cannot. Instead, the coefficients must be estimated by nonlinear methods and (intelligent) trial and error. (Page 5.)

This is an approach that could not have been employed before the computer era, and it cannot be done with an Excel function. Here is the idea, greatly simplified.

- Tell the computer your goal, for example, "minimize the average of the absolute percent errors. These percent errors are the difference between the actual and predicted sale prices divided by the actual price." (Page 6.)

- Set some settings that control the algorithm. By my count, there are close to 20 of these. NCSS recommends defaults, but also advises that the defaults might not always work.

- Develop some "starting values".

- The computer tweaks starting the starting values and calculates the results. If the tweak improves the numbers it is kept. Otherwise rejected.

- The computer repeats the process over and over for all the items – tweak, calculate, accept or reject.

- Of course, everything is interdependent, so if the computer tweaks one item, and then multiple other items are tweaked, the first tweak might not be helping anymore.

- Continue repeating the process until the additional iterations do not get you closer to your goal. This is called "convergence".

Hopefully, everything eventually "converges", producing a stable result that meets the standards.

**What Can Go Wrong?**
The traditional methods use formulas and algorithms that will always get a solution – and will always get the same solution for the same inputs. The traditional methods also come with traditional tests of significance that tell a user when a result should not be relied upon. NCSS's modern, computer assisted trial and error does not have all these protections.

I am not an expert in this field, so I will just quote some of the warnings that appear in the NCSS document itself.

The process might not converge, and if it does not converge you really have to know what you are doing to modify certain "options".

> When confronted with the series of … options in the procedure, the task of setting proper values may seem daunting. Ideally, the default set of options would always yield convergence and a 'best' estimated model. Unfortunately, in practice, convergence is sometimes not achieved with the default options. (Page 10.)

Even if it converges, there is no guarantee that the process creates the best estimate.

> Because hybrid models don't have a closed form solution, iterative methods must be used to determine the estimated coefficients of the models. While these methods allow for increased flexibility in the types of models that may be considered, convergence on a "best" model estimate is not guaranteed. (Page 8.)

What this also implies is that different goals, different setup options and different starting values can converge to different results. Even just changing the starting values could result in different results -- the algorithm might converge on different "local" optimums but not on the "global" optimum.

The data could be a problem, and could cause repeated runs to give different results.

> We have found that in some cases, the nature of the data does not give a stable solution, even though the algorithms converge. For this reason, we recommend that the analysis be run more than once, with the same settings, even when the run seems to complete normally. In the cases where repeated runs give different results (perhaps with substantially varying coefficient estimates), there may be problems in the dataset itself causing the issue. (Page 10.)

It is recognized that the process can result in poor estimates. The software actually produces a report of "Poorly Estimated Properties".

> This report shows those rows with a large (percentage) difference from the estimated sale price to the actual sale price. The percent error cutoff … is 30%. Each row in this report should be analyzed to determine if there is some underlying explanation as to why the estimation is so poor. In some cases it may be reasonable to try re-estimating the same model without these poorly estimated properties, to determine their influence. (Page 21.)

NCSS's license agreement (in another document) includes a clear disclaimer.

> NO WARRANTY OF PERFORMANCE. Dr. Jerry L. Hintze does not and cannot warrant the performance or results that may be obtained by using NCSS. Accordingly, NCSS and its documentation are licensed "as is" without warranty as to their performance, merchantability, or fitness for any particular purpose. ***The entire risk as to the results and performance of NCSS is assumed by you.*** Should NCSS prove defective, you (and not Dr. Jerry L. Hintze nor his dealers) assume the entire cost of all necessary servicing, repair, or correction. (Emphasis added.)

**An Interesting Evasion**

Assuming he used this methodology, Mr. Ryan's failure to explain it may have been because he was aware of the risks and complexities.

In retrospect, there was a specific occasion where Mr. Ryan avoided answering a straightforward question that should have revealed his use of the NCSS algorithm or any other similar software. At the April 21, 2016, Committee of the Whole, starting at about 192:19 on the recording posted on the village website, Ron Parlato asked a great question.

> Ron Parlato: Does your program have artificial intelligence programmed into it? Have you gotten up to that new technology or no?

John Ryan: The sales inform the model. It's all market based.

Ron Parlato: So, you don't have artificial intelligence here, okay.

Although I simplified the explanation of NCSS's trial and error algorithm, the algorithm definitely falls into the mainstream definition of "artificial intelligence".

> The method used in NCSS for making adjustments to the coefficients for each iteration is differential evolution….
>
> Differential evolution is one of a group of *genetic algorithms* (see for example, the recent book by Haupt (1998)). (Page 5.)

The Wikipedia article on Artificial Intelligence explicitly discusses "genetic algorithms" as well as "evolutionary computation" and "evolutionary algorithms".

So, Mr. Ryan just evaded Mr. Parlato's question, and left Mr. Parlato with the false impression that the assessment modeling was not dependent on this "new technology". I was at the meeting, and it also left me with that impression.

**The Software Generates Reports**

It is clear from the NCSS documentation that their software generates many reports. Even if he was using a different software product, it would have produced reports.

I would expect that any responsible person using software of this sort for anything important (such as the creation of a $10 billion assessment roll) would save the reports. This has to be standard operating procedure. Assuming he used this software, did Mr. Ryan not save the reports or did not provide them to the Assessor? Perhaps they are sitting in a village computer in Village Hall, but the Assessor does not know this.

Reports from prior runs would certainly help in understanding the path that Mr. Ryan took to get to the final run. Were there more sales? Were there too many "poorly estimated properties?

All available reports should be disclosed immediately.

**Ratio of Estimates to Actuals**

One of the mathematical features of traditional linear regression (for example as used by Tyler) is that the sum of the actuals (sales prices) always equals the sum of the estimates (value estimates based on the derived formula). In other word, overall, the ratio of the estimates over the actuals is 100%.

As a genetic algorithm – implementing an advanced form of trial and error – the NCSS model does not inherently preserve this 100% relationship. And, in fact, as seen in his sales base and

his report, the ratio for the estimates over the actuals is in the range of 93-94%.[8] In retrospect, this low ratio should have been a red flag.

**Does the use of the NCSS Methodology Explain why Mr. Ryan Excluded Sales?**
This is an important question and I am not quite sure of the answer.

First, as a basic point, it is always possible that the underlying data simply does not strongly support any model.

- As an example, the data could contain internal inconsistencies -- situations where one property is clearly superior to another property based on the characteristics, but the superior property had the lower sale price. This could happen, and there might not be any objective basis for saying that one of them is wrong. The market is not that efficient, and perhaps the Scarsdale real estate market really was unsettled during the analysis period.

- As another example, there may not be enough information to say that specific fitted values are statistically significant. The analogy is that you cannot poll the election by sampling only ten people. Even if seven people say Clinton, you cannot say with any confidence that she gets 70% of the vote. Or, if you test a new drug by giving it to ten people and none of them get bad side effects, you cannot confidently say the drug is safe. With a small sample size and a large diversity of neighborhoods, grades, conditions, etc, many of the derived values could have huge "margins of error" and are just not credible.

These problems could affect any model, a simple traditional model as well as NCSS's advanced algorithm, but I can think of reasons why the latter could be more problematic.

- Mr. Ryan's hybrid model is just so ambitious, with so many refinements that did not exist in the Tyler model. It is a nice idea, but it just might not have been supportable.

- Traditional regression methods have traditional tests that warn when the results are not statistically significant. In other words, better or more familiar red flags.

- The NCSS methodology – an advanced trial and error – might have just taken too long to run. Long run times without convergence would be very frustrating. This may have limited Mr. Ryan's ability to modify the model to better adapt to the data.

I repeat that these are just some thoughts I have, not hard factual claims. The whole idea is that Mr. Ryan should provide hard factual explanations and documentation.

In any event, if you are having trouble fitting a model, it has to be tempting to discard inputs. This is a temptation that has to be avoided. I know others will ask about it, so I do not want to go further in making accusations.

[8] I have written elsewhere that the reported Uniform Percentage of Value on the Tentative Roll should have been 94%, not 100%.

But I will say what a responsible person should have done if in fact it was clear that the data did not support a significant result. A responsible person should have "put the pencil down" and should have informed the client that the assignment could not be completed. It would not have been a terrible thing if Scarsdale had to go another year with the Tyler-based values.

**Land**

I did not have had time in this analysis to explore whether NCSS software or similar software many have contributed to land anomalies.

I have previously submitted questions regarding the land situation, so I hope Mr. Ryan answers them in his initial presentation.

**Questions**

Based on the foregoing, here are some specific questions that I would ask Mr. Ryan. I will augment this by Wednesday. Obviously, if he totally denies using NCSS or any other AI-type software, the questions go in other directions.

- Exactly what software did you use to fit your model? Did you use the NCSS Hybrid Appraisal Models procedure?

- Why have you not previously disclosed your use of this technology?

- What is your understanding of the risks associated with this technology? Did you encounter problems consistent with those risks? What did you do to manage those risks/problems?

- In the course of working with the software, did you initially begin with more than 220 sales? How many? Why were some excluded?

- Was the software installed and used on a Scarsdale computer?

- Did the software produce documentation? Did you save it? Where is it?

- Did the Appraiser or anyone in her office ask about the software and the methodology? Did you ever explain it? Did you ever explain the risks? [Parallel questions for the Appraiser.]

- When did you start the calibration process? When did you complete it? How many hours/days did you spend on it?

Michael Levine
August 13, 2016