Inter-appraiser variability for written appraisals

James Komen

Abstract

The purpose of this experiment was to eliminate personal observation error by taking away the opportunity for the appraising arborist to make observations. Participants in the experiment were given a written description of four fictional trees and were asked to appraise the cost solution using the CTLA Trunk Formula Method (TFM). They were instructed to only use the information that was provided to them and to assume that no other defects or conditions were observable on the trees. The personal value opinions of each appraising arborist were independently recorded for each of the three depreciation ratings used in the TFM: species, location, and condition. The factor that was found to have the highest standard deviation was the location rating. This location rating variability is due exclusively to personal value error because there was no physical nexus for the appraisers to observe. The participants valued the locations differently and therefore assigned percentage ratings accordingly.

Introduction

The Trunk Formula Method (TFM) as outlined in the 9th Edition of the Guide for Plant Appraisal (CTLA 2000) has five basic component parts: trunk area, unit cost, species rating, location rating, and condition rating. Each of these components has been shown experimentally to contribute to inter-appraiser variability (Komen and Hodel 2015; Watson 2002). Inter-appraiser variability has four component parts identified in Komen and Hodel (2015): personal value error, personal observation error, measurement error, and systematic error. The two components of error relevant to the depreciation ratings were personal value error and personal observation error.

Both Watson (2002) and Komen and Hodel (2015) confounded personal observation error and personal value error in their experimental design. Appraisers were allowed to independently examine the subject trees and assign ratings, but there was no way to differentiate between differences of opinion regarding the same observation and differences of observed tree defects between appraisers. Some appraisers may have seen tree defects or location attributes that were not observed by others. They may have accordingly assigned lower location and or condition ratings.

Watson's research provided trunk area, unit cost, and species ratings to the appraisers as given and only measured variability of location and condition ratings (2002). Komen and Hodel's research provided unit cost and existing species ratings to the appraisers, but allowed for variability of the species rating within the TFM guidelines (2015).

The goal of this study was to eliminate personal observation error by eliminating the need to make field observations. If all of the relevant observations were listed as text and provided to the study participants, then the variability between their appraised answers would be solely due to personal value error. The appraising arborists would be coming to different conclusions from the same observations.

Materials and methods

25 arborists attended an appraisal class workshop held in Arcadia, California. As a class exercise, the class participants were given a written assignment. The data from the written assignment was used as a basis for this research.

Participants were given a packet of written descriptions of four fictional trees. These descriptions contained data such as the trunk diameter, the height, the spread, location characteristics, site characteristics, and condition of various parts of the tree (**Fig. 1**). Participants were given an appraisal data spreadsheet and were instructed to assign values to each of the respective depreciation ratings for the subject tree and to then appraise the trees using the trunk formula method.

Of the 25 participants in the class, 22 of them signed data release waivers allowing their data to be used for this experiment anonymously. One of the participants was not a certified arborist and so that participant's data was excluded. The 21 remaining appraisal sheets were inputted into an Excel spreadsheet for analysis. Only the arborist opinion values were used as inputs to eliminate the possibility of introducing mathematical error from the appraisers. Data was analyzed by calculating standard deviations of each of the component variables in the trunk formula method. Then the standard deviations were compared and ranked.

Per the Guide for Plant Appraisal (CTLA 2000), participants were instructed to only modify the species rating by a maximum of 10%.

Results

The independent variable with the highest standard deviation was the location rating (**Fig. 2**). These standard deviations were consistently and significantly higher than either the species or condition ratings. The

Tree #2

Pinus coulteri

- Height: 60 feet, Spread: 20 feet, DBH: 12 inches
- WCISA Data:
  - o Species rating: 70%
  - o Replacement size: 12.56 sq in
  - o Unit cost $118/sq in
  - o Installed tree cost: $1482
- Eastern-most tree of a tight grove on the western edge of a 2-acre vacant lot in the San Gabriel mountains, near a mountain resort.
- Site clearing has begun for development of an upscale vacation home.
- Subject tree was recently exposed to additional wind loads from site clearing.
- Neighboring lots are vacant, but proposed for future development.
- Tree has very few low branches. Most of the foliage is in the upper 20% of its crown. It has grown in a grove all of its life. It's neighbors are all the same age.
- The canopy is biased to the east, away from grove.
- Sap is exuding from several small scattered sites up and down the entire trunk. There appears to be recent boring activity.
- The foliage is green.
- The grade is steeply sloped downward to the south at 10 degrees from horizontal.
- Soil is covered in layer of pine needle mulch.
- There is no irrigation for the site.

Figure 1. Example of one of the tree descriptions in the packet provided to the appraising arborist participants on December 12, 2015.

species ratings had a relatively low standard deviation, but this is likely due to the artificial restriction of the TFM that limits the assignment of species ratings to within 10% of the published rating. The condition ratings had approximately the same standard deviation as in the field trials by Komen and Hodel (2015).

Discussion

Because there were no field observations to be made, every appraising arborist used the same set of observations provided in the packet. Therefore, this experiment trial eliminated personal observation error. All that remained was personal value error for the respective values that arborists assigned to the subjective ratings of species, location, and condition.

The standard deviations of the condition ratings were understandably lower than in the field trial because all of the arborists were using the same set of observations. These standard deviations ranged from 8% to 14% on the four trees in this study. This reflects the differing personal values on rating the condition of the subject trees. In many cases, condition rating components differed by two or more points on a scale of 1 to 4, suggesting that strong personal value biases may have been incorporated into the opinions of the appraising arborist.

The location ratings were the most variable of the three depreciation ratings in the TFM. The variability of the location ratings in this experiment was much higher than in prior field-based research (Watson, 2002; Komen and Hodel 2015). This is likely because all of the trees in those prior experiments were located in an arboretum, an ideal example of a perfect location rating. Those ratings tended to be high and have little variability. There has not been much research on the location rating outside of a controlled arboretum, and this experiment is the first to highlight the issues with appraising location ratings on private property.

One source of variability is the starting value. Appraisers used different starting values and made appropriate adjustments relative to a base value. From discussions in class, participants described their differing thought processes:

Figure 2.Results from the data analysis. The standard deviations are gradient-shaded from highest variability (darker) to lowest variability (lighter). The attribute with the largest standard deviations was the location rating.

Tree #	Mean Loc	STDV Loc	Mean Spec	STDV Spec	Mean Cond	STDV Cond	Mean Cost	STDV Cost	STDV %
1	80 %	11 %	90 %	4 %	84 %	9 %	$ 7,103.51	$ 1,471.21	21 %
2	67 %	18 %	68 %	5 %	72 %	14 %	$ 4,379.61	$ 1,566.82	36 %
3	76 %	23 %	67 %	5 %	85 %	8 %	$ 4,940.05	$ 1,645.45	33 %
4	75 %	14 %	28 %	4 %	57 %	11 %	$ 2,081.34	$ 687.58	33 %

- Some appraising arborists began with a 100% rating and deducted points for defects as they observed them.
- Some began at a "middle" value of 50% percent, adding points for attributes that they believed to be positive and subtracting for attributes that they believed to be negative.
- Some arborists used 75% as an average starting value to mimic the grading system used in high schools and colleges.
- Another possible process was to use 0% as a starting value and add to it for positive attributes only. No participants in the class described using this method.

The condition ratings' standard deviations were slightly lower than in the field trial, suggesting that one of the possible components of the condition ratings' standard deviations was personal observation error. Performing this experiment as a written exercise instead of a field trial eliminated that component of error. Personal observation error appeared to be a relatively small component on the error of the overall condition rating – most of the inter-arborist variability must be due to personal value error.

To reduce inter-appraiser variability, there are two concurrent strategies that may be considered. To reduce personal observation error, appraising arborists would be trained to better identify tree attributes. This would be done with extensive field training and requirements for years of prior experience inspecting trees.

To reduce personal value error, appraising arborists could be calibrated together as a group. This could be done with extensive field training and requirements for a certain number of years of prior experience inspecting trees. This classroom-precision research suggests that more effort should be put towards the strategy of calibrating arborists to reduce personal value error because the personal value error component of variability was much larger than the personal observation error component. Whether such a calibration exercise is effective at reducing inter-appraiser variability could be the subject of future research.

Future research should attempt to isolate personal observation error to reconcile with the results of this experiment. If a future study finds that personal observation error is a small component of inter-appraiser variability, then it will further support the conclusions in this research.

Conclusion

Eliminating personal observation error did not eliminate or significantly reduce variability between appraising arborists. The personal value error of the appraising arborists was observed to be a large component of the inter-appraiser variability of the final appraised value. The location rating was observed to be the component of depreciation with the largest variability.

James Komen
Board Certified Master Arborist with a background in finance and accounting.

References

CTLA. 2000. Council of Tree & Landscape Appraisers. Guide for Plant Appraisal, ninth edition. International Society of Arboriculture, Champaign, Illinois, U.S.

Komen, J. and D. Hodel. 2015. An Analysis of the Field Precision of the CTLA Trunk Formula Method. Arboriculture & Urban Forestry 2015. 41(5): 279-285

Watson, G. 2002. Comparing formula methods of tree appraisal. Journal of Arboriculture 28(1):11-18.