# Valuation Strategies

JULY/AUGUST 2008


## Single Period Capitalization Method

Annuities

Discount Rates

# Valuation Strategies

JULY/AUGUST 2008
VOLUME 11, NUMBER 6



WG&L

# counterpoint

## REGRESSION ANALYSIS AND THE CLOSELY HELD COMPANY

MARK G. FILLER

I read with interest the recent article by Mary Ann Lerch concerning the application of regression analysis to restricted stock studies.[1] Kudos to her for pointing out that when it comes to regression models that have an r-squared ($r^2$) of less than .40, even if some or most of the independent variables are statistically significant at the 5% level, the emperor has no clothes. That is, these models, such as those provided by Silber, Hertzel & Smith, and by Bajaj, are useless in quotidian valuation work, as they have no predictive power because they explain so little.

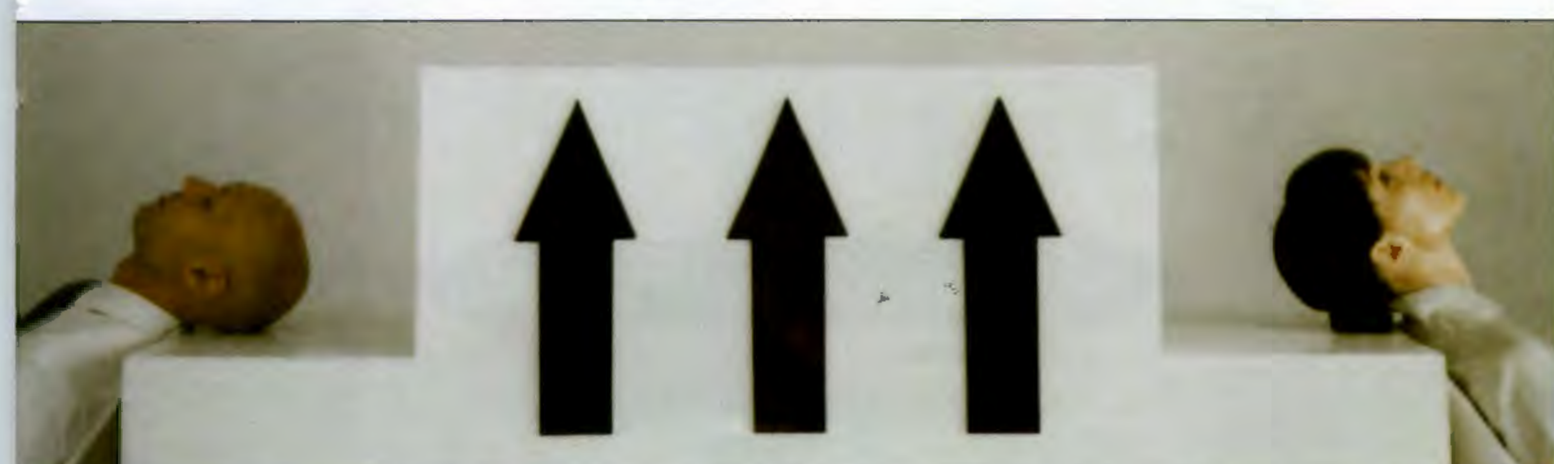
More kudos to her for pointing out that regression analysis is the best and proper tool for describing and analyzing the relationship between two variables, and that valuation analysts' insistence on the continuing use of means and medians serves only to demonstrate their unbecoming fear of statistics. Finally, kudos to her for applying Occam's razor to the whole issue of discounts for lack of marketability based on restricted stock studies. Rather than try and locate which factors drive the discount, she has focused on determining the size of the discount relative to the price of the freely traded stock. However, there are some definitions, applications, and arguments that I take exception to, and I feel that a response to those items is necessary.

First, when discussing the McConaughy, Cary, and Chen study,[2] she says that only three of the "response" variables were significant. The only variables that get tested for significance are the independent variables, often called the explanatory or predictor, variables (i.e., the $x$ variables). The $y$ variable is the response, often called the dependent or explained variable.

Next, she defines the correlation coefficient, also called the coefficient of correlation, as the "probability of the predicted value of $y$ being correct given that value for $x$." If only this were so! The coefficient of correlation, which can run from -1 to +1, measures only the strength of the linear relation between $x$ and $y$. There is such a thing as negative correlation—think of the relationship between high school grades and hours per week of TV viewing. How would Ms. Lerch interpret the negative correlation coefficient produced in that scenario? Would she conjure up the idea of negative probability? I believe she has conflated the idea of the total area under the normal curve representing 100% probability with the idea of positive correlation running from 0 to 1.

In the same paragraph as the incorrect definition of the correlation coefficient, she asserts that the unexplained variance of 4.92% is excellent in spite of the data set containing only 53 transactions. While 4.92% is an excellent result, I find her reasoning odd as $r^2$, and consequently $1 - r^2$, is not determined or affected by the size of the data set. One calculates $r^2$ by dividing the sum of the squared differences between predicted $y$ and the average of $y$ by the sum of the squared differences between $y$ and the average of $y$. Reading from Excel's summary output produced by its regression data analysis tool, it would be Regression SS divided by Total SS. The reader will notice that there is noth-

*MARK G. FILLER, CPA/ABV, CBA, AM, CVA, is the founder and president of Filler & Associates, P.A., a CPA firm in Portland, Maine. His personal practice focuses on business valuations, measurement of damages for lost profits, and assistance in adjusting business interruption losses.*



ing in that formula that references the number of observations in the data set. Perhaps Ms. Lerch is referring to the standard error of the estimate, a measure of accuracy, but this too is not affected by the population or sample size. Only confidence intervals are so affected, but this concept is never mentioned in her article. Thus, it is not clear what she is referencing.

Next on the list is the idea that 1 - the *x* coefficient, or the slope of the line, is the actual discount itself. Ms. Lerch refers to the "discount popping out of the equation." This is true only if the regression model is of the type known as "regression through the origin," and that is true because the formula for a straight line is: $y = a + bx$, which is the same formula for simple linear regression, the type used by Ms. Lerch in her article. Therefore, if *y* in this case equals the trend-line value of the restricted stock price, and if the discount itself is the difference between the trend-line value and the freely traded value divided by the freely traded value, then by definition *y* must be computed using *a*, the constant or intercept. To use only *b* to compute the discount is to end up only with the rate of change in the discount between two different values of *x*, and not the discount itself.

For example, from Ms. Lerch's Exhibit 1, the regression equation is: $y = .8195x - .7348$. Applying the equation 1 - the *x* coefficient of .8195 to freely traded stock prices of $10 and $15 produces discounts of 18.03% for each, when in fact the actual discounts per the full regression equation are 25.4% and 22.9%, respectively. Regression through the origin applied to the 53 observations in the data set produces the equation: $y = .7729x$; 1 - the *x* coefficient will produce a uniform discount of 22.71% for each of the 53 restricted shares in comparison with their freely traded equivalents. This type of model is rarely suggested by the literature for use in a case like this, and should not be used just to put forth a simple expression of a discount percentage.

Ms. Lerch frets needlessly about the two "outliers" shown on her Exhibit 1 and the single outliers shown on her Exhibits 3 and 5. These data points are a type of outlier known as influential points, and in all three cases would not need to be removed, as they so obviously lie on the plane of the trend line. This is proven by that fact that when they were removed the *x* coefficients changed slightly from .8195 to .8054, from .7871 to .8075, and from .8635 to .845, respectively.

Discussing her Exhibit 5, Ms. Lerch makes the comment that "with only 10 data points (and eight degrees of freedom) . . . the relationship is modest." I do not know what this means, as the strength of the "relationship" is defined by r, $r^2$, the standard error of the estimate and the *t*-statistic, with no reference to the size of the data set. Fewer data points make it more difficult for the *t* statistic to be statistically significant when measured against the critical value of *t*, but this concept is not addressed anywhere in the article. It is not clear what she is referencing.

When discussing her Exhibit 7, Ms. Lerch states that the generic equation for a simple power curve is $y = ax^b$, which she claims results in a zero intercept. Power models are also known as log-log models because the previous equation can be re-expressed as $ln(y) = ln(a) + bln(x)$. Nowhere in either of these equations does the intercept disappear. These models are not forms of the "regression through the origin model," which is by definition a zero intercept model, but simply logarithmic transformations of a nonlinear data set so that it can be manipulated by the basic linear model.

If Ms. Lerch had subjected her Exhibit 1 model to further testing, she would have discovered that (1) there is a data point more than three standard deviations from the mean of the trend line, and (2) that the model's resulting residuals are heteroscedastic, i.e., the variance about the trend line is not constant. Transformations of *x* and/or *y* will fix both of these problems and produce a better fitting model as a result.

Finally, as interesting an exercise as this was, where does it leave us? While we now have a more precise way to calculate marketability discounts than just using means or medians, there is still the problem of employing this methodology with closely held companies. Simply applying the regression equation to the marketable minority value of the subject company will not allow one to derive the correct discount, as the equation supplies the correct discount only if the marketability minority value is that of a freely traded publicly held stock. The problem remains of how much more the discount ought to be for a closely held stock that has no possibility of ever being freely traded. ●

1 Lerch, "Quantification of Marketability Discounts Using Regression Analysis," 11 Val. Strat. 28 (March/April 2008).

2 McConaughy, Cary, and Chen, "Factors Affecting Discounts on Restricted Stocks," 4 Val. Strat. 14 (November/December 2000).