FOCUS

August/September 2006
Vol. 2, No. 5

Newsletter of the AICPA Business Valuation and Forensic & Litigation Services Section

What's Inside

The Application of Regression Analysis to the Direct Market Data Method

By James A. DiGabriele, D.P.S., CPA/ABV, CFE, CFSA, DABFA, Cr.FA, CVA, and Mark G. Filler, CPA/ABV, CBA, AM, CVA

What is regression analysis, and where have you seen it before?

Technically speaking, curve fitting, also termed regression analysis or regression, is a generic term for all methods of quantifying the relationship between two groups of variables by constructing a model that fits the data. The model may then be used either to merely describe the relationship between the two groups of variables or to predict new values.

The two groups of variables involved in regression are usually denoted as x and y, and the purpose of regression is to build a model $y = f(x)$. Such a model tries to explain or predict the variations in the y-variable, or dependent variable, from the variations in the x-variable, or independent variable. The link between x and y is achieved by applying the model to a set of data that includes both x- and y-values; for example, an economist may collect data in order to evaluate price increases based on either demand or changes in the money supply, or changes in inflation or interest rates.

In business valuation (BV), for example, it is commonly known that, *ceteris peribus*, value is a function of cash flow. Therefore, various databases such as Bizcomps, Pratt's Stats, Done Deals, and IBA, have collected data sets from market transactions that include for each transaction, among other items of interest, selling price and Seller's Discretionary Earnings (SDE). Regression is then used to relate selling price, the y-variable, to SDE, the x-variable. Once you have built a regression model, you can predict the selling price for your subject company, using the known SDEs from the database as the predictors.

Although regression is a technique that has not yet been accorded widespread use in BV, there are some valuation applications that are familiar. For instance, your personal residence is appraised for assessment purposes using a form of regression known as multiple regression, wherein the selling prices of all homes in the municipality over a given time period are regressed against such x-variables as square footage, age, number of bathrooms, lot size, and the like. Plugging the x-variables of your home into the model produces assessed value.

In BV, the pioneering work of Jay Abrams relating equity value and the subject company's discount rate to Ibbotson's 10 deciles of stock market returns was accomplished using regression, as was a similar study carried out by Grabowski and King using, instead of 10 deciles, 25 percentiles to determine the discount rate based on various proxies for size. Grabowski and King went on to regress the equity risk premium against annual average operating margin, the coefficient of variation of annual operating margin, and the coefficient of variation of annual returns to shareholders' equity. Each of these examples provided practitioners with the model's output, the x-variable coefficients that, when multiplied by the appropriate x-variable, produced the subject company's value or discount rate or equity risk premium, with no additional knowledge or work. However, all of these BV regression applications are confined to the income approach.

Continued on page [illegible]

FOCUS,

August/September 2006, Volume 2, Number 5. Published by the American Institute of Certified Public Accountants. Harborside Financial Center, 201 Plaza Three, Jersey City, NJ 07311-3881. Printed in the U.S.A.

Continued from page 1

In this series of articles, the authors will attempt to introduce and encourage the use of regression analysis as it applies in general to the market approach, and more specifically to the direct market data method using the Bizcomps database.

A Brief Introduction to Simple Linear Regression

If you plot two variables against each other in a scatterplot, or scatter graph, the values usually do not fall in a perfectly straight line. If you perform a linear regression analysis, you attempt to find the line that best estimates the relationship between the two variables (the *y*-, or dependent, variable, and the *x*-, or independent, variable). The line you find is called the fitted regression line, and the equation that specifies the line is called the regression equation.

If the data in a scatterplot fall approximately in a straight line, you can use linear regression to find an equation for the regression line drawn over the data. Usually, you will not be able to fit the data perfectly, so some points will lie above and some below the fitted regression line.

The regression line that Excel fits will have an equation of the form $y = a + bx$. Once again, *y* is the dependent variable, the one you are trying to predict, and *x* is the independent, or predictor, variable, the one that is doing the predicting. Finally, *a* and *b* are called coefficients. Figure 1 on page 3 shows a line with $a = 10$ and $b = 2$. The short vertical line segments represent the errors, also called residuals, which are gaps between the line and the points. The residuals are the differences between the observed dependent values and the predicted values. Because *a* is the point at which the line intercepts the vertical axis, *a* is sometimes called the intercept or constant term in the model. Because *b* shows the steepness of the line, *b* is called the slope. The slope gives the ratio, known as rise over run, between the vertical change and the horizontal change along the line. In figure 1, *y* increases from 10 to 30 when *x* increases from 0 to 10, so the slope is b = vertical change/horizontal change = $(30\text{-}10)/(10\text{-}0) = 2$.

Suppose that *x* is years on the job and *y* is salary. Then the y-intercept ($x = 0$) is the salary for a person with zero years' experience, the starting salary. The slope is the change in salary per year of service. A person with a salary above the line would have a positive residual, and a person with a salary below the line would have a negative residual.

If the line trends downward so that *y* decreases when *x* increases, then the slope is negative. For example, if *x* is age, and *y* is the price of used cars, then the slope gives the drop in price per year of age. In this example, the intercept is the price when new, and the residuals represent the difference between the actual price and the predicted price. All other things being equal, if the straight line is the correct model, a positive residual means a car is selling for more than it should, and a negative residual means a car is selling for less than it should (that is, it's a bargain).

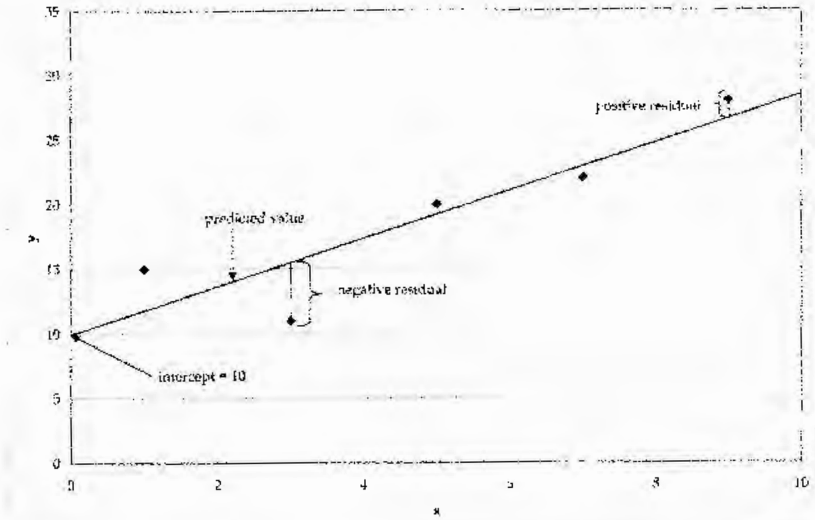
I Get Good Results with Average or Median Ratios — Why Should I Switch to Regression Analysis?

There are multiple reasons for using regression analysis, but, as they say, a picture is worth a thousand words. But first, let us rely on basic intuition. Figure 2 on page 4 is a schedule showing a truncated set of market transactions, along with three sets of predicted values derived from the average Price/SDE ratio, the median Price/SDE ratio, and a regression equation. Included in the schedule are various metrics, including means that were derived by using Excel's AVERAGE function; standard deviations (the average amount of dispersions around the mean) derived with STDEV; and the median, derived with MEDIAN. Also shown is the COV, or coefficient of variation, which is obtained by dividing the standard deviation by the mean, and which places all the outputs on a standardized footing for comparative purposes. The COD, or coefficient of dispersion is the average of the absolute deviations (AAD) from the median divided by the median. This too is used for comparative purposes. Although the standard deviation measures the dispersion about the average of a single column of numbers, root mean squared error (RMSE), or the standard error, measures the size of the average deviation of an observed value from the regression line; or the average deviation between two columns of numbers, the observed and the predicted. The Excel formula for RMSE for the average and median outputs is:

=SQRT(SUMXMY2(B4:B13,H4:H13)/(COUNT (B4:B13)-1)),

Continued on next page

Figure 1



Mark G. Filler, CPA/ABV, CBA, AM, CVA

of Filler & Associates, P.A., Portland, ME 04101, can be contacted at (207) 772-0153 and mfiller@filler.com

James A. DiGabriele, D.P.S., CPA/ABV, CFE, CFSA, DABFA, Cr.FA, CVA

of DiGabriele, McNulty & Co. LLC, West Orange, NJ 07052 can be contacted at (973) 243-2600 and jim@dmcpa.com. He is also Assistant Accounting Professor at Montclair State University School of Business and can be contacted there at 973-655-7288.

where B4:B13 are the observed values of *y* for this data set, H4:H13 are the predicted values of *y* for this data set, and 1 represents a loss of one degree of freedom. For regression, the function is STEYX, with the *x*-variables used in conjunction with the observed *y*-variables. The last two items to be addressed are the Excel functions for computing the slope and
the intercept, which are, surprise, SLOPE and INTERCEPT.

It is obvious from the schedule and the demonstrated metrics, namely, COV, COD, and RMSE, that when it comes to reducing dispersion and fitting the data, the average ratio is outperformed by the median ratio, which is outperformed by the regression equation. In fact, as shown in Figure 3 on page 4, using the average and median ratios increases the dispersion because those methods produce predicted values both greater and less than the actual selling prices. This is not true of the regression model, whose predicted values are greater than the smallest selling price, and less than the largest selling price, thereby reducing dispersion by 10%. If value truly is a range, then you want that range to be as narrow as possible. As we now know, regression analysis will narrow that range considerably more then any average or median of ratios.

Now, for that picture. Figure 4 on page 5 shows the original selling prices and the predicted selling prices derived from the three prediction methods. No comments are necessary, except to say that this is one example of how the use of average or median ratios can lead you far astray from a reasonable conclusion of value.

Other reasons for substituting regression analysis for ratios are the following:

1. A fundamental axiom of finance and BV is that cash flows, and by proxy, revenues, drive value. Don't we, as business valuators, want to establish a model that explores the relationship between selling price and cash flow or revenues? An average or median ratio cannot model that relationship, cannot determine the magnitude of the relationships between the two variables, and cannot be used to make accurate predictions. We have yet to see a finance text that encourages the use of average ratios for valuing publicly traded companies in lieu of a regression model. Nor have we seen any articles in peer reviewed research journals that explore areas of interest using average or median ratios. Regression analysis is the tool of choice for exploring relationships and making predictions among financial experts and scientists. So should it be for business valuators.
2. Most of the SIC Code data sets are not linear in the relationship between price and SDE or revenue. They are curvilinear, and as such, value predictions based on averages or medians will always incorrectly value a company that has a value driver in the upper ranges of the distribution. Regression can very easily be modified to address this problem.
3. Assume two companies, each with sales of $1.0MM, but one has SDE of $180.0M, and the other has SDE of $260.0M. *Ceteris paribus*, the company with the higher SDE

Continued on next page

Figure 2

	Y	X		Outputs		
	Selling Price	SDE	Ratio Y/X	Avg Ratio Y/X	Median Y/X	Linear Regression
	40,000	20,000	2.00	27,222	25,000	41,333
	43,333	20,000	2.17	27,222	25,000	41,333
	43,333	30,000	1.44	40,833	37,500	45,167
	46,667	30,000	1.56	40,833	37,500	45,167
	50,000	40,000	1.25	54,444	50,000	49,000
	50,000	40,000	1.25	54,444	50,000	49,000
	46,667	50,000	0.93	68,056	62,500	52,833
	53,333	50,000	1.07	68,056	62,500	52,833
	56,667	60,000	0.94	81,667	75,000	56,667
	60,000	60,000	1.00	81,667	75,000	56,667
Mean =	**49,000**		**1.36**	**54,444**		**49,000**
Std Dev =	**6,295**		**0.43**			**5,714**
COV (Std Dev) =			**31.9%**			
COV (RMSE) =				**29.2%**		**5.7%**
Median =			**1.25**		**50,000**	
AAD =			**0.32**		**5,000**	
COD =			**25.8%**		**10.0%**	
RMSE =				**15,886**	**13,229**	**2,801**

Elements of the Linear Equation:

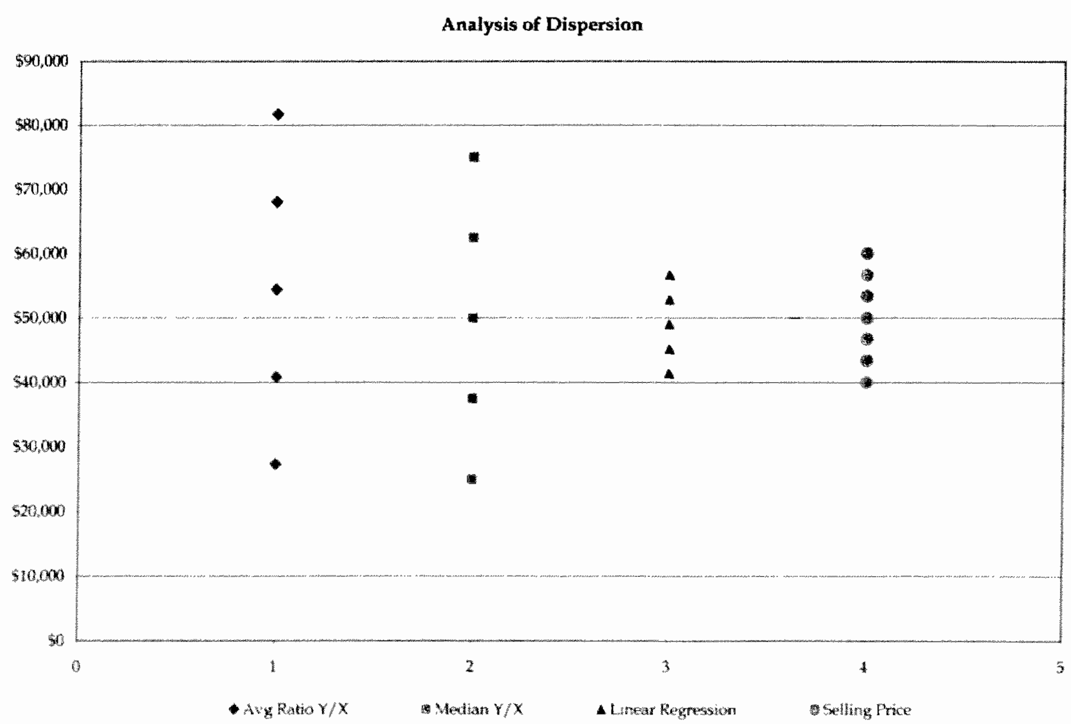
Intercept = a = **33,667**

Slope = b = **0.3833**

Equation ($y = a + bx$):

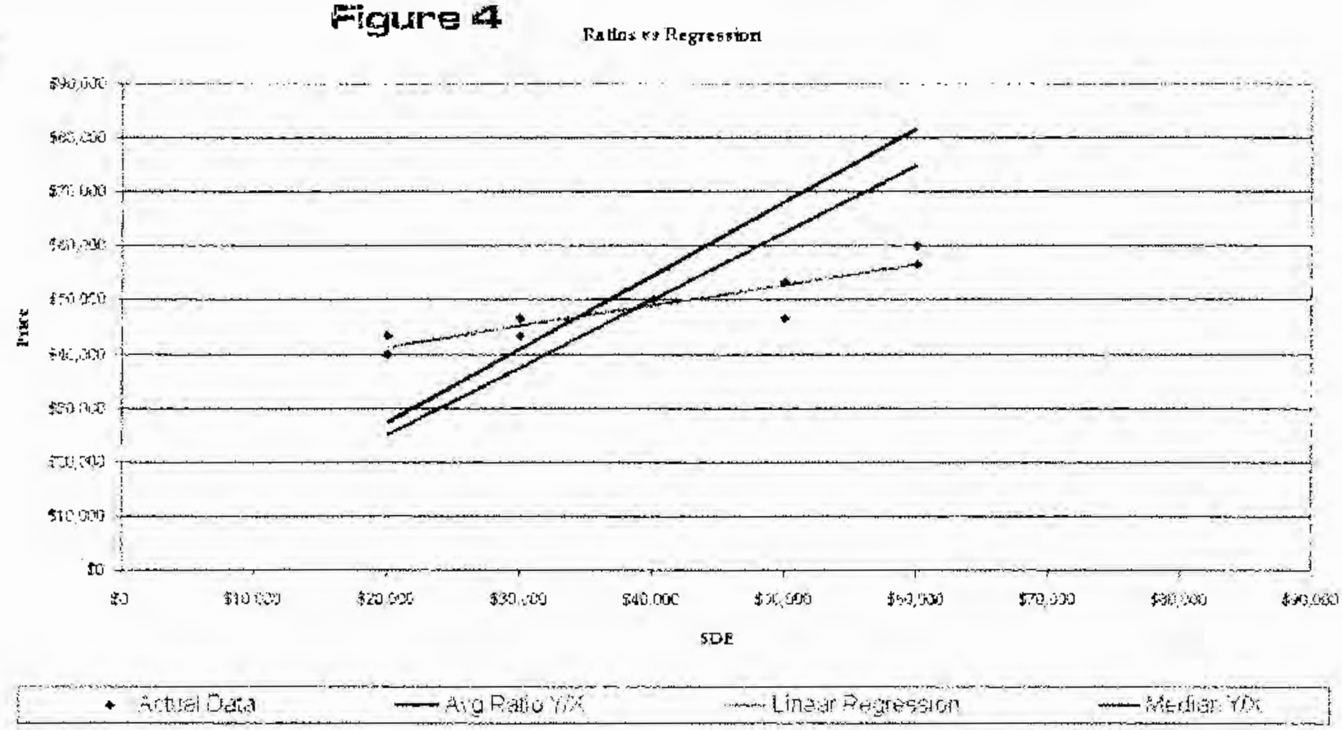
Linear Regression = Predicted Selling Price = 33,667 + .3833(SDE)

Figure 3



Continued on next page

Figure 4



should sell for more. Using just a multiple of sales, however, will not give this result, and in fact, will overvalue one company and undervalue the other. Regression analysis allows one to predict an entity's value based on revenue while also controlling for (i.e., accounting for or taking into consideration) operating margin.

4. The distributions of both selling prices and SDE or revenue are rarely normal, or even symmetrical (they are skewed to the right). This makes the predictions derived from averages and median ratios unreliable because the variation in predicted prices is not due to unbiased measurement error. That is, since the data are not symmetrical and normal, one cannot explain the variation in the data as just residual noise; it may be something else. Therefore, one needs to transform the data, both the *x*-variable and/or the *y*-variable, into normal and symmetrical distributions. Regression analysis more easily handles this task than does a univariate transformation process.

5. And finally, regression analysis provides for stronger courtroom testimony because it allows the business valuator to speak to the relationship between the value drivers, SDE or revenue, and the item at issue, value, in more definitive terms than average or median ratios allow.

Future articles in this series will address the following topics:

1. How does one perform a regression analysis using Microsoft's Excel?
2. How does one apply it to the Bizcomps database.
3. Why does simple linear regression or ordinary least squares (OLS) rarely give us the right answer, and what can we do about it?
4. Can one handle SDE and annual revenue as value drivers in the same manner, or must one use different procedures for each?
5. How do you know if your regression model is giving you the best answer, or how to interpret the summary output data that Excel provides?

Letters to the Editor

Focus encourages readers to write letters on business valuation, forensic, and litigation consulting services issues and on published articles. Please remember to include your name and telephone and fax numbers. Send your letters by e-mail to wmoran@aicpa.org.

FOCUS

OCTOBER/NOVEMBER/ DECEMBER 2006
VOL. 2, NO. 6

Newsletter of the AICPA Business Valuation and Forensic & Litigation Services Section

What's Inside

Focusing on Fraud: High- and Low-Stakes Gambles

In late September, the AICPA National Conference on Fraud and Litigation Services was held in Las Vegas. The following article focuses on highlights of some of the presentations in the fraud track. Some of the presentations in the litigation track will be covered in another article. The sessions summarized in this article made clear that success in preventing and detecting fraud in most organizations hinges on meeting several needs, namely, to assess risk, establish effective controls, set the right tone at the top, and exercise a healthy skepticism about how well the organization's culture and operations are being maintained.

The Bellagio in Las Vegas was the site of the AICPA National Conference on Fraud and Litigation Services. The increasing appeal of this annual conference was attested to by the increased number of practitioners attending, this year totaling about 525. Whether a practitioner's focus was fraud or other litigation services or both, the 42 sessions offered something for everyone.

Glenn Newman, conference steering committee chair, opened the conference with some introductory remarks and then introduced the keynote speaker Dick Thornburgh, whose long and distinguished career in public service includes having served as Governor of Pennsylvania and U.S. Attorney General under two presidents. More recently, Mr. Thornburgh served as court-appointed Examiner in the World.Com bankruptcy proceedings and cochaired the independent investigation into the alleged use of false documents by CBS News' "60 Minutes Wednesday" to report on President George W. Bush's service in the Texas Air National Guard.

In his remarks, Mr. Thornburgh focused primarily on the lessons learned through the World.Com and Arthur Andersen experiences. His message to practitioners, especially auditors, was: Have professional skepticism. In discussing Andersen's experience with Enron, he commented that red flags were abundant and possible risks of misstatement were missed. He believes, moreover, that Andersen lacked a forensic type of analysis and relied on management explanations. Although taking management explanations at face value may appear to serve the client, Mr. Thornburgh believes that failing to challenge management assertions and be skeptical does not in the end serve the interest of the client. Granted that the pressure on auditors remains high, he said, and includes the risk of alienating the client, but the loss of public trust is the biggest loss in these situations.

Scandal by the Sea

An investigation as dramatic and notorious as the high-profile investigations in which Governor Thornburgh participated is the investigation that was the subject of a concurrent session entitled "Government Fraud & Corruption—Investigation of the City of San Diego and Its Pension System." This much-publicized instance of government fraud and corruption illustrates the consequences of the failure to challenge management and instead to acquiesce to a culture of corruption. The session presenter was Troy Dahlberg, JD, CPA/ABV, a managing director and the national practice leader for Kroll's Forensic Accounting and Litigation Consulting Practice. Dahlberg, along with Arthur Levitt, Jr., and Lynn E. Turner, served on the Audit Committee formed to investigate the San Diego City Employees' Retirement System and the city's sewer rate structure. In its report, the

Continued on page 2

The Application of Regression Analysis to the Direct Market Data Method

Part 2: Performing a regression analysis using Microsoft Excel

By Mark G. Filler, CPA/ABV, CBA, AM, CVA, and James A. DiGabriele, D.P.S., CPA/ABV, CFE, CFSA, DABFA, Cr.FA, CVA

Like all Microsoft Office products, there are at least two ways to do anything in Excel, including regression analysis (RA). Rather than develop a tutorial that demonstrates all the possible ways Excel's RA features can be put to use, the authors will focus on instructing you in the use of the functions they use daily in their business valuation (BV) practices.

As we showed in Part 1 of this series, a picture is worth a thousand words, so let's start there. Figure 1 represents a sample of 15 sales transactions drawn from the Bizcomps database, without correcting for the fact that some of the transactions include seller financing with below-market rates of interest, an infirmity we will address later in this article. For ease of instruction, we are showing only those columns of information provided by Bizcomps that are pertinent to the task at hand. Please recreate Figure 1 in Excel on your own computer, or at a minimum, just fill in columns F for SDE and H for Selling Price, save the worksheet, and then follow the instructions below.

First, select the range F3:F17, then hold down the control key and select the range H3:H17. Click on the Chart function button, click XY (scatter), click next, click next again, remove the legend by right-clicking and selecting clear, select the Titles tab, enter Price to SDE as the chart title, enter SDE ($) as the X axis value and Price ($) as the Y axis value, click next, and place the chart in a new sheet. Your chart should look like Figure 2. Now, right click on any one of the data points, choose add trendline, select Linear type, click on the Options tab and select Display equation and Display R-squared. Click OK and save the workbook. Your chart should now look like Figure 3.

You now have a visual presentation of the relationship between the x-variable, SDE and the y-variable (the selling price), along with the equation for predicting selling prices, as well as a measure of goodness of fit, the equation's r-squared value. The chart is dynamic, not static, which means that if we change any of the data in Figure 1, the chart will automatically update. Don't mind the low R^2 and the outlying data points; we'll deal with those in a later article. For now, let's focus on learning about Excel's RA functions.

Analysis ToolPak

A static presentation of RA, useful for reports, can be found in Excel's Analysis ToolPak. If you don't already have the ToolPak loaded into Excel, go to Tools, Add-ins, and select Analysis ToolPak and Analysis ToolPak-VBA, and click OK. This will load the ToolPak for you. To use the ToolPak, go to Tools, Data Analysis, scroll down and select Regression, and click OK. This will bring up the regression analysis tool. The input Y range is H2:H17, and the Input X range is F2:F17. Select Labels, and for output, select New Worksheet Ply, and then click OK, and save the workbook. Your output will look like Figure 4 after you have deleted columns H and I, have selected the whole output section A1:G18, have clicked on Format, selected columns, and have chosen AutoFit Selection. Notice that R square is the same number as R^2 in Figure 3, and that the coefficients for the Intercept and SDE are the same numbers as in the equation in Figure 3. We will explain the purpose of the additional information contained in the Summary Output later in this series of articles.

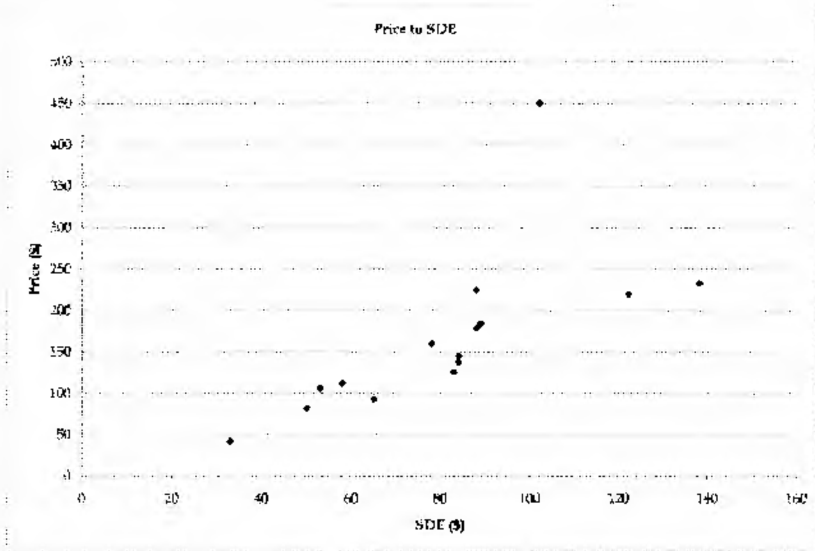
Another way to do an RA that contains almost as much information as the static regression analysis tool output is to use Excel's array formula in conjunction with one of its statistical functions. Beneath the columns for SDE and Selling Price in Figure 1 that you previously created, select and highlight with the cursor an area 2 columns wide and 5 rows deep, say the range H23:I27. Click on the Paste Function button, on the left side select the Statistical function category, and on the right side, select LINEST and click OK. For Known Y's, select

Figure 1

BIZCOMPS DATA

Data No.	SIC CODE #	Business Type	Annual Revenue	SDE	Sales Date	Selling Price	Per Cent Down	Terms	Area	Days on Market
1	2396	Silk Screen Printing	205	50	8/31/1993	82	70	2 Yrs @ 8%	Baton Rouge, LA	
2	2396	Silk Screen Printing	248	33	8/13/1999	42	100	N/A	Midwest	120
3	2396	Silk Screen Printing	283	58	9/23/1998	112	28	4 Yrs @ 8%	Ohio	201
4	2396	Silk Screen Printing	299	89	9/30/1998	185	21	6 Mos @ 10%	Tampa, FL	110
5	2396	Silk Screen Printing	346	83	6/30/1994	126	39	5 Yrs @ 9%	Central Florida	
6	2396	Silk Screen Printing	350	122	12/7/2001	220	45	4 Yrs @ 10%	Florida	118
7	2396	Silk Screen Printing	376	88	6/12/2001	179	100	N/A	Spokane, WA	120
8	2396	Silk Screen Printing	379	78	10/22/2002	160	100	N/A	San Diego, CA	87
9	2396	Silk Screen Printing	401	84	10/1/1998	145	33	10 Yrs @ 8%	Spokane, WA	350
10	2396	Silk Screen Printing	403	53	5/31/2002	106	76	10 Yrs @ 7	Tulsa, OK	90
11	2396	Silk Screen Printing	406	84	4/26/2002	138	50	3 Yrs	Colorado	166
12	2396	Silk Screen Printing	412	88	4/16/2002	225	100	N/A	San Francisco	236
13	2396	Silk Screen Printing	416	65	9/12/2002	93	100	N/A	Florida	54
14	2396	Silk Screen Printing	436	102	11/30/2000	450	100	N/A	Denver, CO	
15	2396	Silk Screen Printing	448	138	1/20/2000	233	20	10 Yrs @ Pr+2.3	Stockton, CA	170

Figure 2



H3:H17; for Known X's, select F3:F17 and enter TRUE for both Const" and "Stats. Do not click OK. Instead, hold down Control and Shift at the same time and simultaneously hit Enter. Save the workbook. Your output should look like the "Summary Output" in Figure 5. We have added a title and explanatory phrases to describe the output. This output, with some additional minor calculations, provides the same information as the regression analysis tool with the added benefit of being dynamic.

In addition to the three ways described above to simultaneously create all the elements of the regression equation, we also saw in Part 1 of this series that we can create the elements individually by use of the SLOPE and INTERCEPT functions. Now that we know how to develop the RA equation, let's explore two of the options Excel gives us to put it to use.

Those options consist of TREND, a function that implements the equation in one step, and second, the creation of a formula that draws on the intercept and SDE coefficients from the array formula summary output. Somewhere to the right of Figure 1, say starting at column O, please enter in row 2 the labels Trend and Array Formula Output in columns O and P. Select cell O3; click on the Paste Function button; on the left side, select the Statistical function category, and on the right side, select TREND and click OK. For Known Y's, select H3:H17 and hit the F4 function key to make the range reference absolute; for Known X's, select F3:F17 and hit the F4 key; and for X, select F3 and enter TRUE for Const. Then click OK. Cell O3 should present 91.60 as the predicted value.

Select P3 and enter the following formula: =+$I$23+$H$23*F3. This is the slope and intercept formula that we used in Part 1 of this series but with the difference that the coefficients have already been determined by another function, rather than using the SLOPE and INTERCEPT functions directly in the formula.

Cell P3 should also present 91.60 as the predicted value. Next, copy cells O3 and P3 down to row 17 and save the workbook. If each row does not contain the same numbers across the columns as shown in the Summary Output in Figure 5, you did not succeed in making the range references absolute in row 3 and you should try that step again.

Let's perform two more calculations to set up the worksheet for use in the next article, and then we'll finish by predicting the value of a sample subject company.

These two calculations are automatically performed for you in the regression analysis tool, and can be part of the output if you select "residuals" and "standardized residuals" in the regression command. However, because the regression tool is static, its use is inappropriate for the type of exploratory analysis we will be doing. In cells Q2 and R2 of what was originally Figure 1 but what is now Figure 5, place the labels Residuals and Standardized Residuals. In cell Q3, enter the formula: =+H3-O3, and copy it down to row 17.

This number is the difference between the actual selling price value and the value that the regression equation predicted for each individual selling price (the regression line). In cell R3, enter the following equation:

=STANDARDIZE(R3,AVERAGE($R$3:$R$17),STDEV($R$3:$R$17))

and copy it down to row 17. This formula in effect divides each residual by the standard deviation of the residuals. The result shows how many standard deviations each residual is from the average, which makes it easy to identify outliers, a topic we will explore in the next article. From the values shown in the Residual column of Figure 5, you can see that there is one residual that seems larger than the others. It is Data No. 14, found in row 16 and which has a standardized residual value of 3.326. You'll want to keep an eye on this observation as we continue to explore this regression model. As we'll show you in a later article, the residuals play an important role in determining the appropriateness of any regression model.

Mark G. Filler, CPA/ABV, CBA, AM, CVA,

of Filler & Associates, P.A., Portland, ME; Phone: 207) 772-0153; Fax: (207) 761-4013; *Email: mfiller@filler.com*

James A. DiGabriele, D.P.S., CPA/ABV, CFE, CFSA, DABFA, Cr.FA, CVA,

of DiGabriele, McNulty & Co. LLC, West Orange, NJ; Phone: (973) 243-2600; *Email: jim@dmcpa.com.* He is also Assistant Accounting Professor at Montclair State University School of Business; (973) 655-7288.

Figure 3

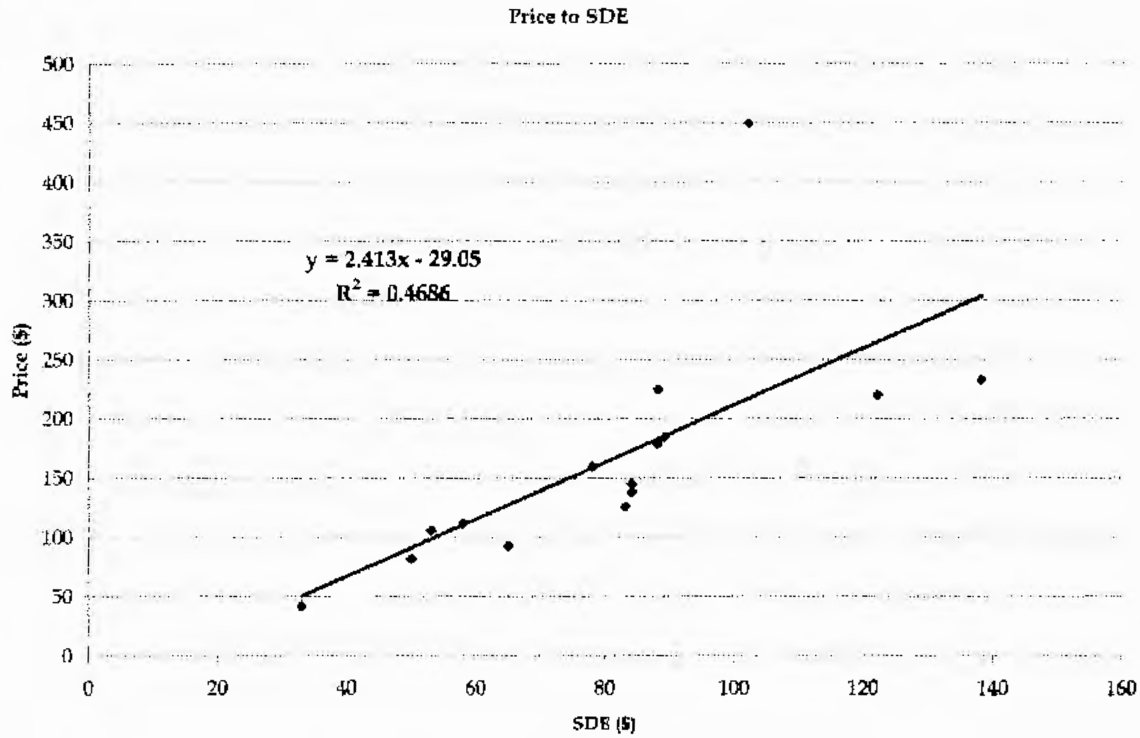


Figure 4

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.6845
R Square	0.4686
Adjusted R Square	0.4277
Standard Error	72.6775
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	60541.268	60541.268	11.462	0.005
Residual	13	68666.332	5282.026		
Total	14	129207.600			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-29.050	60.7044	-0.4786	0.6402	-160.1939	102.0938
SDE	2.413	0.7127	3.3855	0.0049	0.8732	3.9527

Figure 5

BIZCOMPS DATA

Data No.	SIC CODE #	Business Type	Annual Revenue	SDE	Sales Date	Selling Price	Per Cent Down	Terms	Area	Days on Market	Trend	Array Formula Output	Residual	Standardized Residual
1	2396	Silk Screen Printing	305	50	3/31/1998	82	70	2 Yrs @ 8%	Baton Rouge, LA		91.60	91.60	-9.60	-0.157
2	2396	Silk Screen Printing	268	33	8/13/1999	42	100	N/A	Midwest	120	50.58	50.58	-8.58	-0.222
3	2396	Silk Screen Printing	285	58	9/23/1998	112	28	4 Yrs @ 8%	Ohio	201	110.90	110.90	1.10	0.018
4	2396	Silk Screen Printing	299	80	9/30/1998	165	21	6 Mos @ 10%	Tampa, FL	110	165.70	165.70	-0.70	-0.010
5	2396	Silk Screen Printing	340	83	6/30/1999	126	29	5 Yrs @ 9%	Central Florida		171.28	171.28	-45.28	-0.746
6	2396	Silk Screen Printing	350	122	12/7/2001	220	45	4 Yrs @ 10%	Florida	118	265.33	265.33	-45.33	-0.747
7	2396	Silk Screen Printing	376	88	5/12/2001	179	100	N/A	Spokane, WA	128	183.29	183.29	-4.29	-0.071
8	2396	Silk Screen Printing	379	78	10/22/2002	160	100	N/A	San Diego, CA	87	159.16	159.16	0.84	0.014
9	2396	Silk Screen Printing	401	84	10/1/1998	145	35	10 Yrs @ 8%	Spokane, WA	350	173.64	173.64	-28.64	-0.469
10	2396	Silk Screen Printing	403	53	5/31/2002	106	78	10 Yrs @ 7	Tulsa, OK	90	98.84	98.84	7.16	0.118
11	2396	Silk Screen Printing	406	84	4/24/2002	138	41	3 Yrs	Colorado	164	173.64	173.64	-35.64	-0.585
12	2396	Silk Screen Printing	412	88	4/16/2002	225	100	N/A	San Francisco	236	183.29	183.29	41.71	0.686
13	2396	Silk Screen Printing	416	65	9/12/2002	93	100	N/A	Florida	34	127.79	127.79	-34.79	-0.567
14	2396	Silk Screen Printing	424	102	11/30/2000	450	100	N/A	Denver, CO		217.07	217.07	232.93	3.826
15	2396	Silk Screen Printing	448	138	1/29/2000	250	30	10 Yrs @ Pr+2.3	Stockton, CA	170	303.94	303.94	-53.94	-1.013
				81							166.40	166.40		

SUMMARY OUTPUT

Coefficient - SDE	2.413	-29.050	Coefficient - Intercept
Standard Error - SDE	0.713	61.706	Standard Error - Intercept
R Square	0.469	72.670	Standard Error
F stat	11.462	13	Residual df
Regression Sum of Squares	60541.268	68656.332	Residual Sum of Squares

Now let's predict the value of our sample subject company.

Predicting Value

In cell F20 of Figure 5, enter the number 81 that will represent the SDE of our subject company. We wish to predict the selling price, or value, of certain of its assets using the Direct Market Data method. That is, based on the relationship between value and SDE of other silk screen-printing companies that have been sold, what is the predicted value of our sample subject company's assets? Copy cells O17:P17 down to O20:P20, skipping over rows 18 and 19. Save the workbook. Your answer should be 166.40, and it should appear in both cells. Since this number represents only the value of the sample subject company's intangible and fixed assets, in a later article, we'll show you what needs to be added to and subtracted from this number to arrive at a value for a company's equity for both S and C corporation modes.

Seller Financing

We'd like to return to the topic of seller financing, referred to at the beginning of this part of the series. We all know that seller financing almost always carries a below-market rate of interest that results in the selling price being overstated. To prove this point, divide your data set into two segments, one consisting of all cash transactions, and the other consisting of seller-financed transactions. You will find that the one that consists of all cash transactions (9 count) has a Price/SDE average ratio of 1.78, and the other consisting of those that had some seller financing involved (6 count) have an average ratio of 2.29. This overstatement, which typically runs between 9% and 13% of the selling price, can be relieved by following Toby Tatum's procedure as outlined in his seminal text, *Transaction Patterns*. You can convert the six transactions that were supported by seller financing into all-cash equivalent selling prices by use of present-value techniques, which should be done so that there will be comparability among all the data, both all-cash and seller-financed transactions.

The discount rate used to determine the present value of the seller-financed sales is derived from a formula developed by Toby Tatum in Chapter 3 of *Transaction Patterns*. Essentially, it starts with 14% and adds 1% for each 1/10th of the selling price that is seller-financed. So, if a transaction is 70% seller-financed, the discount rate is 21%. This makes sense for two reasons, namely, (1) it's the formula that reduces the average Price/SDE multiple for seller-financed transactions down to the average Price/SDE multiple for all-cash transactions in the Bizcomps database, and (2) seller paper is usually behind the bank, is not collateralized, and will not be recovered upon a default, etc.; it is essentially a very low-grade junk bond and not a publicly traded junk bond either. Once revised, the selling prices would then be substituted back into the Bizcomps worksheet for further analysis.

We haven't demonstrated this technique because we already have enough topics to show you, and we think Tatum's book is something you should have in your library if you are going to apply RA to the Bizcomps database.

Next time we'll answer more questions: Why don't we stop right here and bring this methodology into our BV practices? Why does simple linear regression, otherwise known as ordinary least squares, that we have shown you here in Part 2, rarely give us the right answer when applied to the Bizcomps database in the simple manner demonstrated here, and what can we do about it?

FOCUS

March/April 2007
Vol. 3, No. 2

Newsletter of the AICPA Business Valuation and Forensic & Litigation Services Section

What's Inside

AICPA

Building a Boutique Firm

The following article provides an example of how a newly formed CPA firm that chose to focus on a relatively narrow range of services managed its successful start-up. It also describes how in the start-up process, the CPA firm addressed many of the issues faced by most firms today.

Recently, *Web*CPA surveyed accounting industry leaders, asking them to share their vision of the profession in five years: "What will be its major concerns? Its challenges? The hot new service areas? What shape will the firm landscape have taken?"

Although the responses are quite varied (the responses can be viewed at http://www.webcpa.com/article.cfm?articleid=22377&print=yes), several leaders underscored certain issues that have been foremost in the minds of practitioners in recent years, such as staff recruitment and retention, succession planning, and staff development. What is noteworthy about these particular issues is in that the firm owners' transition to retirement will depend on attracting and retaining quality employees and their retention may depend on the opportunity they're given for development.

Among the hot service areas mentioned, many are not really new. Those of interest to *Focus* readers included forensic and investigatory accounting, fraud prevention, business valuation, fair value accounting, and auditing.

The Growth of Service Boutiques

Also of interest to *Focus* readers and other practitioners who have built or are building niches in business valuation, forensic, and litigation services is the prediction of Gale Crosley, CPA, president of Crosley & Co., who consults with CPA firms: "Many smaller firms that make the choice to remain independent and invest in leadership will become service boutiques, as they discover that the dynamics of standards-setting and a multiprovider environment will enable them to grow and leverage talent better if they focus on a narrow complement of specialized services."

The trend predicted by Crosley is well under way. An illustration of how such a service boutique can be developed is Melinda Harper's founding of Harper Lutz Zuber Potenza & Associates in Denver, Colorado. Melinda went out on her own after the expiration of her noncompete agreement with a national firm, one of the "roll up" firms, with offices across the United States, and several thousand professionals. She had been with the firm and all of its predecessors for about 15 years. Prior to that, she had been with her "first firm" for 13 years.

"I think it is probably a little unusual that, at my age, I wanted to start a new firm. But I had some good reasons," Melinda said. "I wanted to be able to implement management and marketing ideas that I believed would be very successful but were difficult to get support for in a firm that was focused on traditional work and cross-selling services and products. When I joined the firm, it was focused on consulting or project work, and so my practice fit in well, but that changed with the roll up."

Starting her own firm also provided Melinda with the opportunity to put together a team that had both depth and breadth in the litigation/valuation area (complex commercial damages; valuation (both litigation and nonlitigation); lost earnings; family law). Not only would this enable the firm to respond to most litigation support/valuation needs, but also that depth and breadth would differentiate the firm from others in its market.

Continued on page 2

time, the smaller things such as which phone system to get (traditional or VoIP) consumed lots of time and energy. I don't think any of us anticipated all the details we would need to deal with and how much time it would take. Even though we divided up the projects, we all got pretty tired and cranky, but there was a big commitment to seeing it through and to maintaining and valuing our relationship—and we did."

Despite the obstacles, the business began with a running start. Melinda had built a solid foundation: "Since most of our client relationships come through referral sources, I was fortunate in that I had a very loyal referral base, and people found me almost immediately after I left. In order to facilitate that, I made an extra effort to be out in the community and talking to people about my plans during the transition. Many of my referral sources actually assisted with letting people know where I was!" She added further, "Once the firm was formed, two of my partners also had a client base, so we opened the doors with a significant book of business."

Marketing Services

The firm markets primarily to attorneys because they are its largest source of work. Consequently, most marketing dollars go to building awareness of the firm's presence and capabilities. One of the first formal steps following informal contacts was to send an announcement about the new firm to all referral sources. The announcement was coordinated with ads in various legal publications, with press releases, and with two open houses, two nights in a row. "We also focused on attending events where we could let people know where we all were and about the firm's capabilities and on connecting through emails, phone calls, and lunches. We continued the personal contacts with participation in various organizations, regular ads and regular mailings."

The Benefits of Outsourcing

Commenting on the final success of starting up, Melinda said, "I think an important factor in what feels like a big success is that we did not try to do everything ourselves, but instead made sure we had great resources. Our goal was to focus on freeing the technical staff to get and help clients. For instance, we outsourced all of our telephone and computer technology, we found an extremely creative designer for our marketing and advertising materials, including branding, we found a PR person to handle our publicity, and we outsourced our business management and accounting to a fabulous woman that we all knew. We also hired a coach for the partners, and we have worked through many relationship and compensation issues with her, and continue to keep her involved."

The Application of Regression Analysis to the Direct Market Data Method–Part 3

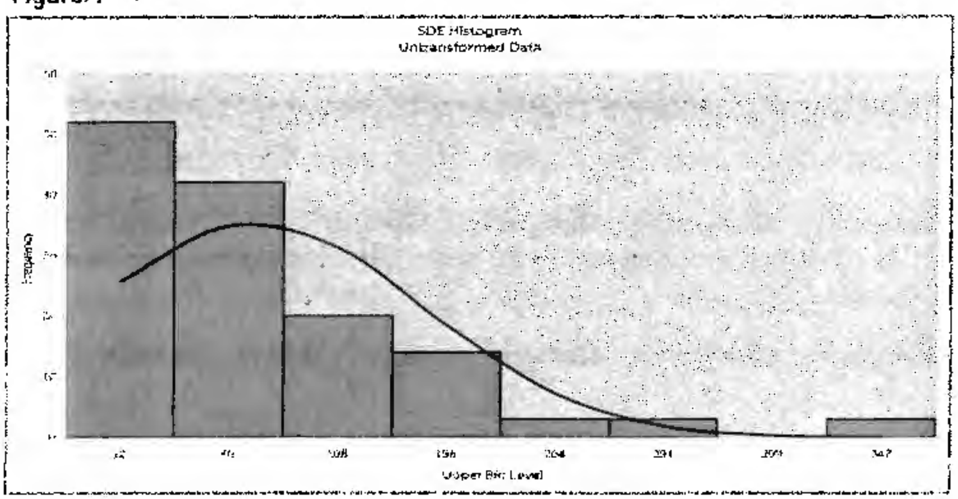
By Mark G. Filler, CPA/ABV, CBA, AM, CVA, and James A. DiGabriele, D.P.S., CPA/ABV, CFE, CFSA, DABFA, Cr.FA, CVA

Why does simple linear regression rarely give us the right answer, and what can we do about it?

The data sets that Bizcomps makes available to us by way of Standard Industrial Classification (SIC) Code Numbers and North American Industry Classification System (NAICS) Code Numbers are rarely distributed in such a manner that the application of simple linear regression will give us a relevant and reliable answer. This is because the individual databases are (1) hardly ever linear, (2) infrequently homogeneous as to variance (the larger the X variable, the greater, or smaller, the dispersion about the regression line), and (3) not often normal, or even symmetrical. If the data is linear, we can proceed to use simple linear regression without having to resort to more complex models, that is, we can stick with the tools that Excel provides us. The reasons that homogeneity and normality, or at least, symmetry are good things are beyond the scope of this series of articles, but suffice it to say that without these qualities, standard statistical tests and confidence intervals will not be reliable, nor will you be able to explain away the variation in your data as noise, or ordinary and expected random error. Simple tests will make it apparent that your model is deficient.

Fortunately, to fix these three problems we need only one procedure, and that is transformation of either or both the X and Y variables. This is so because data that is not normally distributed is also often neither linear nor homogeneous. Thus, transformation provides a simple way both to fix statistical problems (nonsymmetrical and heterogeneous distributions) and to fit curves to data (curvilinear regression). For example, using 143 transactions from SIC Code No. 2752, Figure 1 shows the distribution of the raw-data form of the X variable, SOE, being skewed positively to the right and

Figure 1



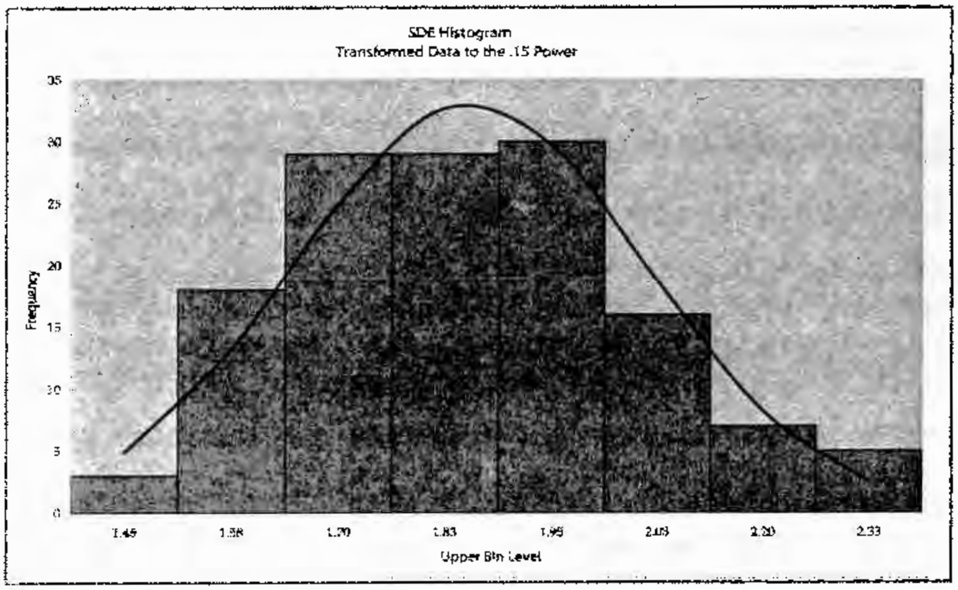
Continued on page 4

Continued from page 3

non-normal in its shape. A super-imposed normal distribution curve points out the discrepancy in shape between the two distributions. Since powers less than 1 can pull in the upper tail of a distribution and help make a skewed distribution more symmetrical, we applied this technique, with the results shown on Figure 2, in which the transformed data's histogram's outline now resembles that of the normal curve.

Transformation of variables is not new to business valuation, as shown by Jay Abrams in his work with the Ibbotson database and Roger Grabowski in his work with the Duff & Phelps database. In both instances, the X variable, market size, was transformed logarithmically to straighten out the curved distribution that is generated when discount rates are plotted against market size. However, transformation by logarithms does not work that well with the Bizcomps data sets as does transformation by exponents, because we can select the exponent that works best in the situation, while the logarithm of any number is fixed. Therefore, because of the flexibility afforded our transformation process by exponents, that will be the transforming process we demonstrate in this article. So, let's set up our worksheet so that we can transform our data and at the same time efficiently identify and remove outliers from the data set. We'll explain later in the article why removing outliers is not only permitted, but in the circumstances, often required.

Figure 2

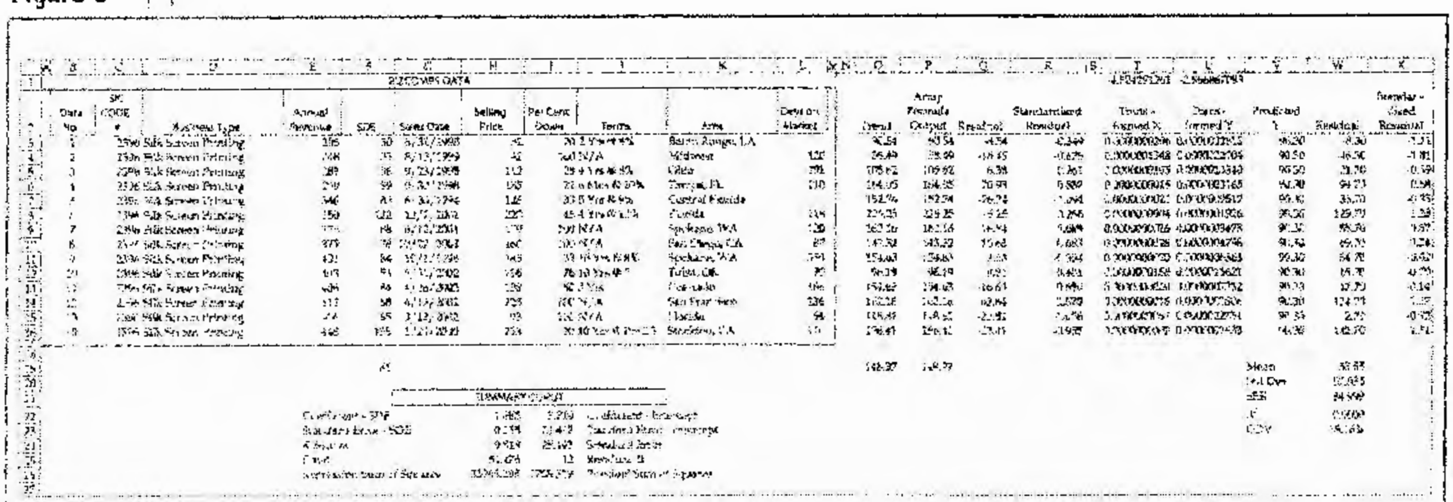


Returning to the last worksheet you created, the last column of figures you have should be titled "Standardized Residual" in column R. Starting in cell T2 and continuing to cell Y2, enter the labels: "Transformed X," "Transformed Y," "Predicted Y," 'Residual," "Standardized Residual," and "Delete if X" as shown in Figure 3. In cell T1 enter, as placeholder amounts, .1, in cell U1 enter .1, and in cell Y1 enter 2.5 as the standard deviation cut-off point. Next, we will transform the X and Y variables using the placeholder amounts in cells T1 and U1.

In cell T3, enter the formula: =F3 ^ $T$1, and in cell U3 enter: =H3 ^ $U$1. This transforms the variables by raising each to the power of .1. Copy cells T3 and U3 down to cells T17 and U17. Next, we will compute the predicted value for Y, using the transformed X and Y variables and then we will back-transform the result right in the formula itself using the reciprocal of the Y transforming exponent.

In cell V3 enter: =TREND($U$3:$U$17, $T$3:$T$17,T3,TRUE) ^ (1/$U$1) – raising the predicted value of Y to the power of the reciprocal of the transforming exponent translates that value back into the original state that Y was expressed in. Just as the square root of 9 can be expressed in Excel as: 9 ^ .5 = 3, then back-transforming makes: 3 ^ (1/.5) = 9. Rather than doing this in two steps, that is, predict Y in its transformed state, and then, in another cell, back-transform it into its original language, we have elected to do it one step. In cell W3 enter: +H3-V3. Copy cells V3 and W3 down

Figure 3



to cells V17 and W17. In cell W20 enter: =AVERAGE(W3:W17); in cell W21 enter: =STDEV(W3:W17); and in cell W22 enter: =SQRT(SUMXMY2(V3:V17,H3:H17)/(COUNT(H3:H17)-2)), which last formula gives us the standard error of the estimate (SEE), that is, the standard deviation about the regression line. Later, we will use the SEE to select the best exponents for the X and Y variables in cells T1 and U1. Also, place the labels "Mean," "Std Dev," "SEE," and "R^2," in cells V19 through V22.

In cell X3 enter: =STANDARDIZE (W3,$W$20,$W$21); and finally in cell Y3 enter: =IF (OR (X3>$Y$1,X3<-$Y$1),"X",""). Copy cells X3 and Y3 down to cells X17 and Y17. As we previously mentioned in Part 2 of this article, Data No. 14 exceeds 2.5 standard deviations from the mean, as indicated by the X in cell Y16. Since this is such a large outlier, whose residual has no chance of being reduced by the transformation, it should be removed from the data set at this time by deleting Row 16. With larger data sets, outliers discovered at this stage can be left in until Solver is set up and run at least once before they are removed. At this point we will set up Excel's powerful optimization feature called Solver Add-In which can calculate solutions to what-if scenarios based on adjustable cells and constraint cells. This allows us to minimize SEE and simultaneously uncover any other outliers that may exist in the data set.

Using Solver

Click on cell W22, select Tools, Solver (if you don't have Solver loaded, go to Tools, Add-ins, scroll down, find and check Solver Add-in, and click OK). Set Target Cell to: W22, set Equal To: Min. by Changing Cells: T1 and U1, then add the following constraint: W20=0. Click on Options, set Precision and Convergence to .000001, set Tolerance to 20%, choose Use Automatic Scaling, click OK and click Solve. Checking to see if there are any more outliers to be removed, we note that there are none. If there were, denoted by an "X" in column Y, then we would delete those rows, and run Solver again, this time by just clicking on Tools, Solver, Solve (Solver remembers your previous settings). We would continue to repeat this process until no more "Xs" showed up in column Y.

Figure 4

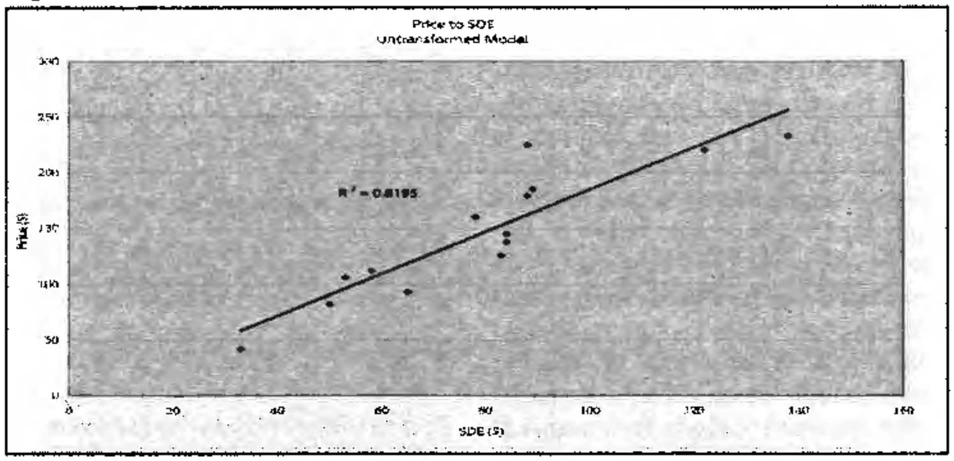
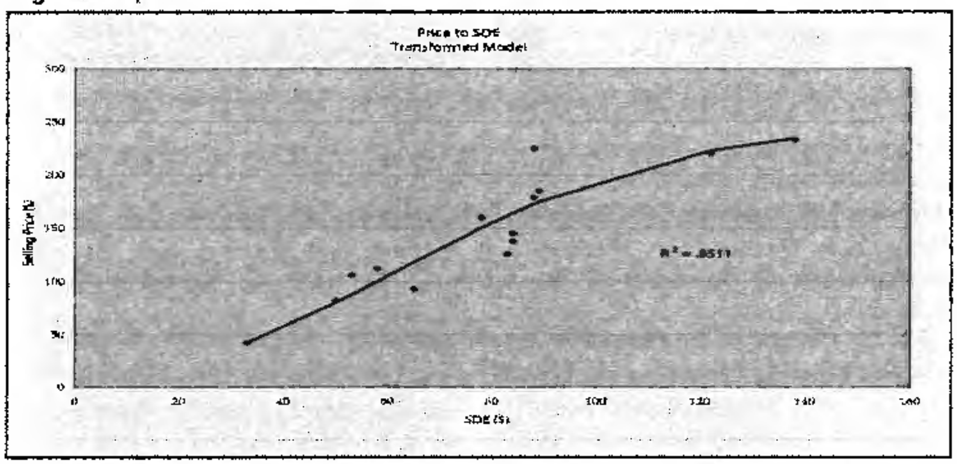


Figure 5



We have just accomplished a number of things, including having changed the values in cells T1 and U1 just enough so that the transformed variables used in the regression equation produce the lowest possible SEE, while at the same time producing the necessary outcome of a mean value of zero for the resulting residuals. In cell W22 enter the formula: =RSQ(V3:V16,H3:H16) so that we can compute R^2 for the transformed model. Checking the output metrics of SEE and R^2 for both the transformed model and the untransformed, or regular, model, we see that the regular model has an SEE and R^2 of 25.443 and .819 as shown in cells I24 and H24, respectively, while the transformed model's metrics are 23.142 and .8511 for SEE and R^2. Also, the regular model has a residual that is more than 2.5 standard deviations from the mean that would have to be removed if that was our model of choice. Therefore, the transformed model gives us a higher R^2, a lower SEE and at the same time allows us to minimize the Data Nos. that must be removed as outliers. Graphical presentations of these outcomes can be seen in Figures 4 and 5.

We can see from the scatterplot of Figure 5 that the data set is in fact curvilinear, but that by transforming the data we were able to fit a line to the data by using Excel's simple linear regression functions, without having to resort to more complex non-linear models. Also, by curving the regression line we were able to keep data no. 12 in the model by making it less than 2.5 standard deviations, and thus not converting it into an outlier that needed to be removed.

Continued on page 6

Continued from page 5

Outliers

As we have seen, outliers are extreme observations that for one reason or another do not belong with the other observations in our data sets. There are two ways that outliers can be introduced into the Bizcomps databases, the first of which results from incorrect recording, or especially, data entry errors that can put wild values into the data sets. The second cause of outliers is that data sets are not homogeneous to which a single regression model will apply, but rather a heterogeneous mix of two or more types of transactions, one of which is more frequent. The infrequent observations of the other types will appear as outliers.

What one does when outliers are identified in the data set is not without controversy. If the outlier is a result of a data entry error or is otherwise suspect in terms of its reliability or accuracy, then it should be clearly removed from the data set or repaired before any further analysis. But what should be done about outliers that are not clearly erroneous, such as those that lie between 2 and 4 standard deviations from the mean of the regression line? Somehow, leaving those observations in the data set has come to be viewed as the "honest" thing to do, and that removing them is viewed as "cherry-picking" or "cheating" or "making it work."

The issue of outlier removal is greatly influenced by what one is trying to accomplish. If you are performing basic science and trying to establish a relationship between, say the number of cigarettes smoked and the onset of lung cancer, then outliers will be important to your research as they will be counter-intuitive to what was expected and therefore will spark new research.

For our purposes, the relationship between SDE and selling price is a fundamental axiom of business valuation—it doesn't need to be established or proved. Therefore, outliers are not helpful sources of new research, but are anomalies. Outliers typically represent (1) input errors, (2) fools for buyers who have overpaid, (3) fools for sellers who have accepted less than fair market value, (4) distressed sellers, or (5) synergistic buyers. Items 2 through 5 violate the fair market value standard of value, and therefore do not belong in the data set. For that reason, it is necessary to delete them along with the obvious data input errors. If your data set contains 75 data points, and 65 of them are within 2.5 standard deviations of the mean, why do you need the other 10, and what helpful information do they contain?

If a data set is heterogeneous and contains all types of transactions, why wouldn't you want to exclude those that do not fit the fair market value standard of value? By definition, it is true that any transaction outside the mainstream does not conform to that standard, whatever the reason. For example, how can a sale that is 4.5 standard deviations from the mean be at fair market value? Mustn't it be at investment value—value to a particular buyer? Even if you make the heroic assumption that a sale at 4.5 standard deviations is truly a fair market value transaction, this question remains: why did it sell for such a high multiple? Perhaps it has, for example, the best location, the best management, superior service, or loyal customers. All these things tend to make its SDE far in excess of the average enterprise in its SIC Code No. Therefore, it sold at a premium; that is, not only was its SDE multiplied by the average multiple, but the buyer paid a premium for its superior performance, as well as the fact that its recipe for success has been systematized by management such that it will survive the closing.

Now, ask yourself whether your subject company enjoys such profits, or has such systems in place. If not, then how can the outlier company be similar and relevant to your valuation assignment? It cannot be, and therefore, it should be removed from the data set. So, remove the outliers because they either don't represent fair market value transactions, or remove them even if, in the extreme, they do. Do not fear that you are "making it work." The cutoff metric is set before you start to eliminate outliers, and it robotically makes the selections. Hence, you are not "cherry picking" the transactions that you keep in the data set; an algorithm decides what transactions fall outside the test metric you have set to determine fair market value.

The next article in this series will address the following topics:

- Should we always set the cut-off metric at 2.5 standard deviations?
- What is the coefficient of variation, and how does it tie in with the previous question?
- Can one handle SDE and Annual Revenue as value drivers in the same manner, or must one use different procedures for each?

Mark G. Filler, CPA/ABV, CBA, AM, CVA

***of Filler & Associates, P.A., Portland, ME 04101; Phone: (207) 772-0153; Fax: (207) 761-4013; Email:* mfiller@filler.com**

James A. DiGabriele, D.P.S., CPA/ABV, CFE, CFSA, DABFA, Cr.FA, CVA,

***is Assistant Professor in the Department of Accounting, Law & Taxation, School of Business, Montclair State University, Montclair, NJ 07042; Phone: (973) 243-2600; Fax: (973) 243-2646; Email:* jim@dmcpa.com**

FOCUS

May/June 2007
Vol. 3, No. 3

Newsletter of the AICPA Business Valuation and Forensic & Litigation Services Section

What's Inside

AICPA®

Another ADR Format

A divorcing couple can increase their chances of reaching a respectful and equitable resolution to their conflict by participating in the Collaborative Practice process. The following article describes the process and the roles of the divorcing couple and the professionals they may engage, including CPAs, to assist in the process.

Collaborative Practice is the term used by the International Academy of Collaborative Professionals (IACP) to refer to divorce proceedings and other settlement arrangements that take place outside the court system. The process for these proceedings has developed as an alternative to proceedings that usually are, at the very least, unpleasant if not lengthy, antagonistic, litigious ordeals that can drain the parties emotionally and financially. Helping divorcing couples to reach a more positive and productive resolution is one of the missions of Collaborative Practice.

Collaborative Practice is based on collaborative law, a process in which lawyers and their clients contractually agree to pursue nonadversarial means of resolving disputes and reaching agreement without going to court.

About Collaborative Practice, IACP president Sue Hansen said, ". . . the emphasis is on improving communication to help couples work through all the legal, financial, and emotional issues in a divorce, including the needs of children. Collaborative Practice gives clients control of decisions as well as access to the problem-solving skills of lawyers, financial specialists, divorce coaches, and child specialists—a full gamut of efficiency and expertise that one is not privy to in the court without the potential of great emotional and financial expense."

According to Hansen, this voluntary, private, out-of-court process often costs less than litigation. The Collaborative process allows couples to steer their divorce by pledging mutual respect and openness, determining the timetable, and working with the Collaborative team towards a settlement they determine together.

Lori Tricaro, a client of Collaborative Practice, cites the benefit of this approach in her case. "Working together with trained professionals," she said, "enabled us to pursue an amicable relationship and to walk away from each meeting without anger. I truly believe this was a positive alternative to moving from one stage of our lives to another that essentially sets the tone for the future of everyone involved."

In Collaborative Practice, a husband and wife are each represented by an attorney trained in the Collaborative Practice process. Attorneys and clients enter into a contract called a "participation agreement." According to the agreement, clients will disclose all information relative to their decision to divorce as well as all of their assets and liabilities. The goal of subsequent meetings between attorneys and clients is that each party understands his or her financial needs and the impact of the divorce on available finances, as well as the resolution of other issues, including parental responsibilities, before they reach a final agreement.

"A plan for the future" is how Steve Kaplan, CPA/ABV, MBA, describes the final agreement. The divorcing couple can formulate the agreement terms themselves. Kaplan, who is with Eisman, Zucker, Klein, & Ruttenberg, LLP, White Plains, NY, is a member of the Board of Directors and is treasurer of the New York Academy of Collaborative Professionals, which is the main group of collaborative professionals in the Metropolitan New York area. Kaplan describes how a team of professionals helps the couple to get to that point of final agreement. From the outset, they have the support of professionals in addition to their

Continued on page 2

The Application of Regression Analysis to the Direct Market Data Method   Part 4

By Mark G. Filler, CPA/ABV, CBA, AM, CVA, and James A. DiGabriele, D.P.S., CPA/ABV, CFE, CFSA, DABFA, Cr.FA, CVA

Should we treat the value driver Annual Revenue in the same manner as we treat Seller's Discretionary Earnings?

For as long as transaction databases have been available, the received wisdom has been that Annual Revenue (AR) is at least as good a predictor of value, if not better than, Seller's Discretionary Earnings (SDE). In this fourth of a series of articles, we will examine this assertion, and if the valuation analyst truly needs to include AR as part of the valuation equation, we will suggest a more appropriate model than merely regressing selling price against AR.

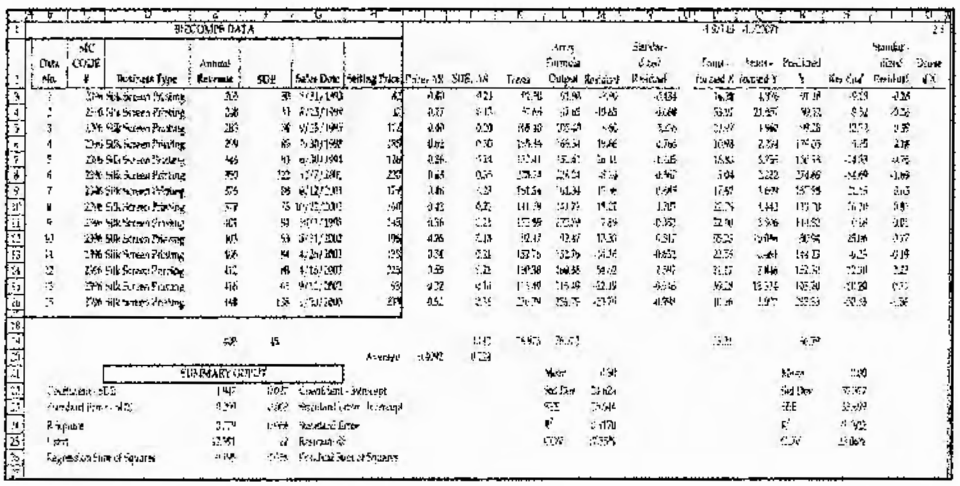
There are a number of reasons, some practical and some logical, for not using AR as the sole predictor of value. In the practice arena, if we use tha 14 data points remaining from our third article as shown in Figure 1, and simply regress selling price against AR, we get the graphic results shown in Figure 2. Notice how dispersed the data points are around the trend line. Many of the data points look like they might be outliers, but the degree of dispersion is so great that they are all within two standard errors of the trend line. This ocular conclusion is ratified by the very low R^2 of .29, indicating that AR only explains 29% of selling price. Not shown is the standard error of the estimate (SEE) of 50.38, an amount almost double that derived from using SDE as the X variable. This is a fairly typical result, and the authors have found that after performing the outlier removing process demonstrated in Part 3 of this series on scores of SIC Code No. databases, AR rarely has better metrics than the SDE of the same data set. (Part 3 was published in the March/April 2007 issue.)

A maxim of financial valuation is that investors buy cash flow. Therefore, when AR is the value driver, it is only serving as a proxy for cash flow, the underlying assumption being that the buyer can repair or reconstruct the company's cost structure so as to produce the necessary cash flow to justify the purchase price. The fact that some buyers will pay a seller a premium for the right to make the company (more) profitable might account for some of the outliers in the databases.

The final and most compelling reason not to use AR as the value driver in a regression equation can best be demonstrated with the following question: Should the assets of two companies sell for the same price when they both have AR of $1,000,000 each, but one of them has SDE of $350,000, and the other has SDE of $200,000? The answer is, *of course not!* Somehow, the selling price of each must reflect its own degree of profitability. Professor Aswath Damodaran, in his textbook, *Investment Valuation*, says that "the key determinant of a revenue multiple is the profit margin – the net margin for price-to-sales ratios and operating margins for value-to-sales ratios." He goes on to say that other "key determinants of the revenue multiple of a firm are its expected risk, payout ratios, and growth characteristics." Unfortunately, these last three determinants are not available to us through any of the transaction databases. But profit margin, the most important determinant, is available through the medium of SDE as found in Bizcomps.

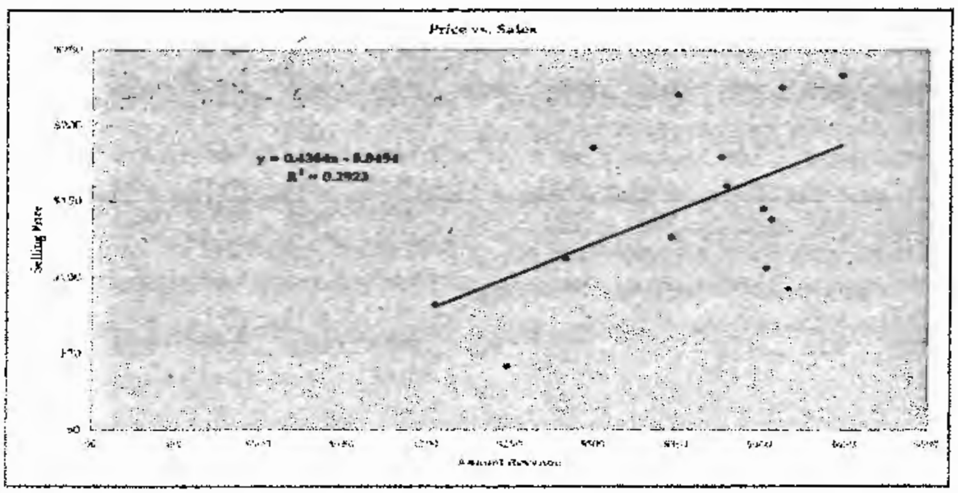
All this, of course, begs the question of why use AR in any case if it is inferior to SDE as a

Figure 1



Data No.	SIC CODE #	Business Type	Annual Revenue	SDE	Sales Date	Selling Price
1	2396	Silk Screen Printing	[illegible]	30	5/31/1993	42
2	2396	Silk Screen Printing	238	51	8/23/1999	[illegible]
3	2396	Silk Screen Printing	283	36	9/25/1995	174
4	2396	Silk Screen Printing	279	85	5/30/1998	[illegible]
5	2396	Silk Screen Printing	[illegible]	[illegible]	6/30/1994	186
6	2396	Silk Screen Printing	350	122	[illegible]	235
7	2396	Silk Screen Printing	[illegible]	86	6/12/2001	[illegible]
8	2396	Silk Screen Printing	377	75	10/22/2002	[illegible]
9	2396	Silk Screen Printing	403	59	[illegible]	145
10	2396	Silk Screen Printing	[illegible]	53	5/31/2002	[illegible]
11	2396	Silk Screen Printing	[illegible]	54	4/26/2003	[illegible]
12	2396	Silk Screen Printing	[illegible]	68	[illegible]	225
13	2396	Silk Screen Printing	416	[illegible]	[illegible]	[illegible]
15	2396	Silk Screen Printing	148	138	[illegible]	[illegible]

Figure 2



value driver? The answer is that there are some fact-specific situations in which the correct use of AR combined with SDE gives one the best answer available. For example, consider the situation in which the seller has expended great effort in developing sales, but for one reason or another, the company has a way below average profit margin. Valuing the company based on sales would certainly over-value it, while valuing it based on SDE alone would under-value it. Is there some way to value the company so that the seller is rewarded for building sales, but punished for not doing it profitably enough? There is, and the remainder of this article will be devoted to showing you how to account for low profitability coupled with AR by use of a formula that adjusts the price-to-sales ratio upwards or downwards based on the degree of profitability, measured as SDE/AR, of the subject company relative to its peers in the date set.

Once more, let's use the same data set that we left off with at the end of Part 3 of this series, the one with 14 data points as shown in Figure 1, data nos. 1-13 and 15, having eliminated data no. 14 as an outlier. First we'll do this as a linear regression, and then we'll do it a second time using the transformation techniques we learned in Part 3. As we are adding new columns to the worksheet, we removed enough columns to the right of the label "Selling Price" such that the label "Trend" winds up in column I. Put the cursor in column I and insert two columns to the left. Label column I "Price/AR", and label column J "SDE/AR." In cell I3, enter the formula +H3/E3, and in cell J3, enter the formula +F3/E3, and then copy cells I3 and J3 down to row 16. Figure 3 indicates that there is a definite linear relationship between the two variables. However, a linear relationship is not necessary for this model to work. In fact, the beauty of the model is that it will work even when R^2 drops to as low as .50.

First, let's make some more room for ourselves in the spreadsheet by moving the block of cells R19:S23 down two rows to R21:S25. Make the references to column H absolute in cells S23, S24, and S25, and then copy this block of cells to L21:M25. Put the cursor in row 18 and insert one row.

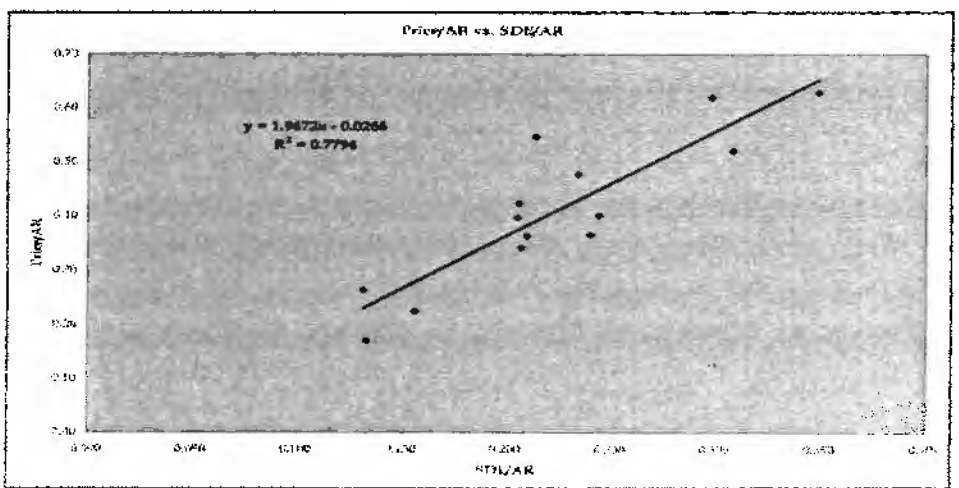
In cell K3, change the formula to read: =TREND($I$3:$I$16,$J$3:$J$16,J3,TRUE)*E3, and then copy cell K3 down to row 16. Change the array formula in cells E22:F26 to read: =LINEST(I3:I16,J3:J16,TRUE,TRUE). Remember to highlight all 10 cells and make your changes to the formula and then hit Control, Shift, and Enter simultaneously to alter the array. In cell L3, change the formula to read =(+$F$22+$E$22*J3)*E3 and then copy cell L3 down to row 16. Now let's create some variables for our subject company by entering 400 in cell E19 and 45 in cell F19. Then copy cells J16, K16, and L16 down to row 19 (skip rows 17 and 18). In cells I20 and J20, compute the averages of rows I3:I16 and J3:J16, respectively.

This valuation model produces a value of $76,973. If our subject company was deemed to have average profitability as measured by SDE/AR, then its value would have approximated $164,000, obtained by multiplying the AR of $400 by the average Price/AR ratio of .4092. But since our subject company's profitability is 50% of the average of those companies in the data set, its Price/AR ratio has been reduced by the regression model to .1924 (76.973/400) to reflect this low degree of profitability relative to sales. Also notice that with the use of a linear model, data no. 12 is an outlier. Rather than immediately removing this data number, let's try a transformation procedure as we did in Part 3 to see if we can keep this data number in the model, and at the same time, obtain superior metrics.

Reset both cells P1 and Q1 to .1. In cell P3, change the formula to read =J3 ^ $P$1. In cell Q3, change the formula to read =I3 ^ $Q$1. In cell R3, change the formula to read =TREND($Q$3:$Q$16,$P$3:$P$16,P3,TRUE) ^(1/$Q$1)*E3. Check to be sure that cell S3 contains the formula H3-X3 and that cell T3 contains the formula =STANDARDIZE (S3,$S$21,$S$22). Now copy cells P3:T3 down to Row 16 and then copy cells P16 and S16 down to cells P19 and S19 (skipping rows 17 and 18). Next, click on Tools, Solver, and click on Solve (again, Solver remembers your previous settings). Since Solver always searches for the perfect answer, it will frequently destabilize the model attempting to provide a solution. As this is probably what you have just experienced, we need to place some constraints on the model so that the best does not become the enemy of the good, and we get a meaningful solution. Click on Tools, Solver, Add, in "Cell reference" put P1:Q1, in the next box choose <=, and in "constraint" place 1. Repeat this process with the same cell references, choose >=, and make the constraint: -5. This limits how far Solver can roam in its search for a solution. Why did we choose these constraints? Trial and error. By substituting various values in cells P1 and Q1, we can estimate the points at which the model will destabilize and then place these estimates in the Solver function. While each data set will have its own set of constraints, the authors never set theirs higher than 5 or lower than –5, and very often, as in this case, one or the other constraint will be

Continued on page 6

Figure 3



Continued from page 5

Figure 4

	B	C	D	E	F	G	H	I	J	K	L	M	N	P	Q	R	S	T	U
1			BIZCOMPS DATA											1.43091	-1.373775				2.5
2	Data No.	SIC CODE #	Business Type	Annual Revenue	SDE	Sales Date	Selling Price	Price/AR	SDE/AR	Trend	Array Formula Output	Residual	Standardized Residual	Transformed X	Transformed Y	Predicted Y	Residual	Standardized Residual	Delete if X
3	1	2396	Silk Screen Printing	205	50	3/31/1999	83	0.40	0.24	89.56	89.56	-6.56	-0.384	7.52	3.531	89.04	-6.04	-0.28	
4	2	2396	Silk Screen Printing	268	30	8/13/1999	42	0.17	0.11	53.90	53.90	-11.90	-0.703	17.90	11.467	50.88	-8.88	-0.42	
5	3	2396	Silk Screen Printing	283	58	9/23/1998	112	0.40	0.20	101.80	101.80	10.20	0.590	9.65	3.573	97.14	14.86	0.70	
6	4	2396	Silk Screen Printing	299	89	9/30/1998	185	0.62	0.30	162.47	162.47	22.53	1.312	5.66	1.934	169.34	15.66	0.74	
7	5	2396	Silk Screen Printing	346	93	6/30/1994	126	0.36	0.34	148.41	148.41	-22.41	-1.318	7.71	4.006	145.37	-19.37	-0.91	
8	6	2396	Silk Screen Printing	350	122	12/7/2001	220	0.63	0.35	225.48	225.48	-5.48	-0.327	4.52	1.892	252.69	-32.69	-1.54	
9	7	2396	Silk Screen Printing	376	88	6/12/2001	179	0.48	0.23	156.93	156.93	22.07	1.285	7.96	2.772	152.94	26.06	1.23	
10	8	2396	Silk Screen Printing	379	78	10/22/2002	160	0.42	0.21	136.98	136.98	23.02	1.340	9.59	3.370	130.78	29.22	1.37	
11	9	2396	Silk Screen Printing	401	84	10/1/1998	145	0.36	0.21	147.85	147.85	-2.85	-0.173	7.36	4.045	141.48	3.52	0.17	
12	10	2396	Silk Screen Printing	405	53	5/21/2002	106	0.26	0.13	86.35	86.35	19.65	1.143	18.20	6.263	81.58	24.42	1.15	
13	11	2396	Silk Screen Printing	406	84	4/26/2002	138	0.34	0.21	147.61	147.61	-9.61	-0.569	9.52	4.404	142.03	-3.03	-0.14	
14	12	2396	Silk Screen Printing	416	65	3/12/2002	93	0.22	0.16	109.52	109.52	-16.52	-0.973	14.23	7.631	102.83	-9.83	-0.46	
15	13	2396	Silk Screen Printing	448	138	1/30/2000	233	0.52	0.31	252.64	252.64	-19.64	-1.156	5.29	2.456	266.89	-33.89	-1.59	
17																			
18				450	45				0.113	70.637	70.637			22.76		67.86			
19							Average	0.3987	0.225										
20			SUMMARY OUPUT								Mean	0.12				Mean	0.08		
21	Coefficient - SDE			1.981	-0.046	Coefficient - Intercept					Std Dev	17.091				Std Dev	21.254		
22	Standard Error - SDE			0.225	0.033	Standard Error - Intercept					SEE	17.851				SEE	22.199		
23	R Square			0.875	0.051	Standard Error					R^2	0.9057				R^2	0.6595		
24	F stat			77.236	11	Residual df					COV	12.74%				COV	15.85%		
25	Regression Sum of Squares			0.202	0.029	Residual Sum of Squares													
26																			

Figure 5

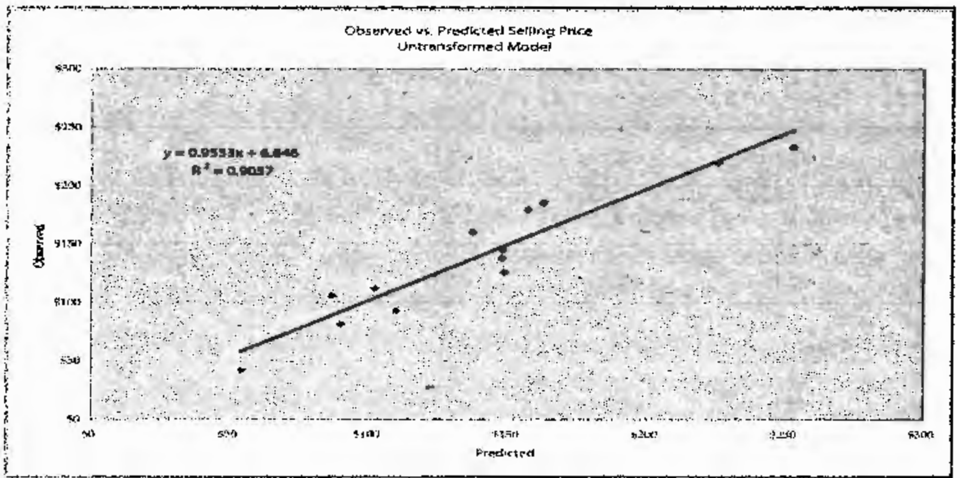
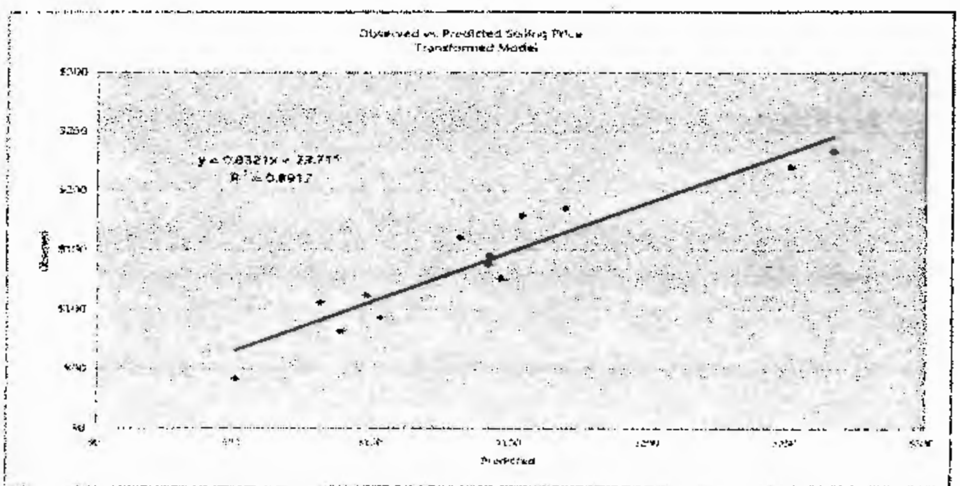


Figure 6



considerably closer to zero than either of these arbitrary maximums. Now click on Tools, Solver and Solve and repeat the process. Very often, especially in a complicated model such as this one, Solver needs two or more tries to optimize the model and produce usable results.

Let's compare the results of the two models, transformed and untransformed, to see which has the better metrics. While the predicted value for selling price is lower with the transformed model, and there is no standardized residual greater than 2.5 as there is in the untransformed model, the metrics for the transformed model are worse than those of the untransformed model. This just goes to show that in this area of business valuation, as in all others, often there are unexpected surprises, blind alleys, dead-ends, and cul-de-sacs. What course of action do the authors recommend at this point? As always, reasonableness, informed judgment, and common sense will come into play.

Save your file and then make a copy of the current worksheet and place it next to the worksheet we were just working on (giving it a different name). One possible solution to this conundrum is to remove data no. 12, as it is more than 2.5 standardized residuals from the mean in the untransformed model, and at 2.23 standardized residuals in the transformed model, it is close to the cut-off point. Place your cursor in Row 14 and delete that row and run Solver once more.

Once more, let's compare the results as shown in Figure 4.

Again, much to the surprise of the authors, the output metrics show the untransformed model still outperforming the transformed model. Comparing Figures 5 and 6 readily shows this. This is very unusual and may be just because of the truncated nature and narrow range of variables of this particular data set which was created for ease of demonstration, but please do not rely on this example as a reason to not transform your data sets. In the authors' experience, nine times out of ten, transforming the data sets produces superior results. However, in this case, the untransformed model gives superior results as demonstrated in Figure 7, which is a line chart comparing observed (actual) selling price with its predicted value per the linear equation. If R^2 were 1, rather than .9057, each set of data points would lie on top of each other.

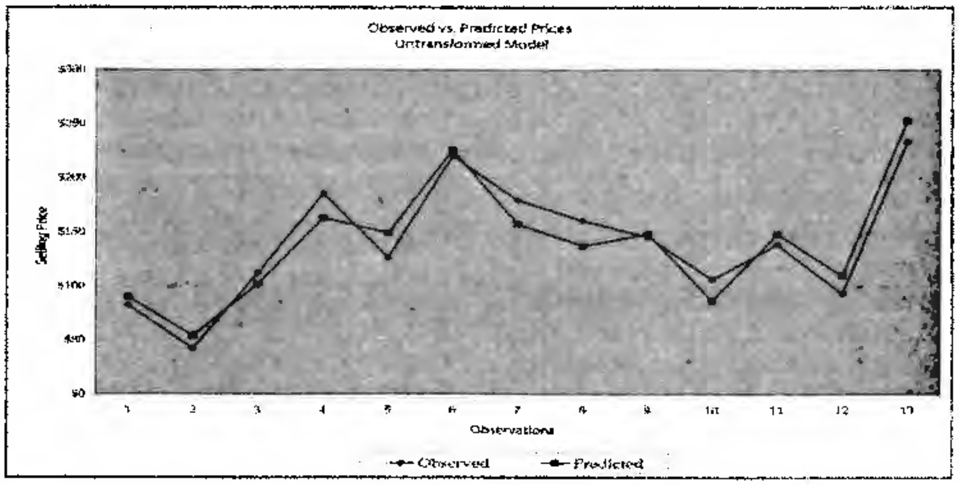
In conclusion, we can see that using Price/AR as a function of SDE/AR will produce a more realistic value when sales are relatively high and profits are relatively low, as opposed to the use of either AR or SDE alone as the sole value driver. With this particular data set, after removing two data nos. as outliers, the value results are as follows:

Label	X Variable	Predicted Selling Price
AR	$400	$159.476
SDE/AR	.113	$ 70.637

This table indicates that the use of AR alone will over-value the subject company's assets by a considerable amount, and that the use of either SDE alone, or in combination with AR as demonstrated in this article, will produce a more realistic value.

One more relevant topic is the question of how small should one make the outlier cutoff? The authors consistently use 2.5 standard deviations because experience has shown them that as we drop the cut-off to 2 standard deviations, thereby obtaining both lower SEEs and corresponding coefficients of variation (COVs), too many data points are given up to achieve this desired result.

Figure 7



The cutoff of 2.5 standardized residuals was chosen as a compromise between the textbook recommended 3 and the Toby Tatum suggested 2. One of the authors, starting with a data set of 137 observations and using lowest COV and observation count as his metrics, ran a transforming model with three different cutoff figures and came up with the results shown in the following table:

Residual Cutoff	Lowest COV	No. of Observations
2.0 Standard deviations	16.11%	90
2.5 Standard deviations	22.31%	118
3.0 Standard deviations	27.80%	128

The decrease from 3 to 2.5 standard deviations results in a decrease in the COV of 24.6% at a cost of an 8.4% decrease in the number of observations, for a ratio of 2.93 (24.6/8.4) to 1. On the other hand, a decrease from 3 to 2 standard deviations results in a decrease in the COV of 72% at a cost of a 42% decrease in the number of observations, for a ratio of 1.71 (72/42) to 1. More than a third of the observations are given up to get that highly desirable low COV of 16.11%. We think that this is too high a price to pay and recommend a cutoff of 2.5 standard deviations.

In the next and final article in this series, we will offer assistance in understanding, interpreting, and using Excel's summary output for regression analysis.

Mark G. Filler, CPA/ABV, CBA, AM, CVA of Filler & Associates, P.A.. Portland, ME 04101; Phone: (207) 772-0153; Fax: (207) 761-4013 ; Email: mfiller@filler.com

James A. DiGabriele, D.P.S., CPA/ABV, CFE, CFSA, DABFA, Cr.FA, CVA, is Assistant Professor in the Department of Accounting, Law & Taxation, School of Business, Montclair State University, Montclair, NJ 07042; Phone: (973) 243-2600; Fax: (973) 243-2646; Email: jim@dmcpa.com

Correction

There is an editing error in part 3 of the series of articles "The Application of Regression Analysis to the Direct Market Data Method," which appeared in the March/April 2007 issue of *Focus*.

In the second paragraph, which begins in the middle column on page 3, the second sentence should read "This is so because data that is ***not*** normally distributed is also often neither linear nor homogeneous." The word *not* was omitted in the article.

Our apologies for the error.

FOCUS

July/August 2007
Vol. 3, No. 4

Newsletter of the AICPA
Business Valuation and
Forensic & Litigation Services Section

What's Inside

Letters to the Editor

Focus encourages its readers to write letters on consulting services issues and on published articles. Please remember to include your name and telephone and fax numbers. Send your letters by e-mail to wmoran@aicpa.org.

AICPA

Implementing the AICPA Business Valuation Standards

By Randie Dial, CPA/ABV

The AICPA Consulting Services Executive Committee (CSEC) issued the long-awaited Statement on Standards for Valuation Services No. 1 (SSVS No. 1). SSVS No. 1 is entitled *Valuation of a Business, Business Ownership Interest, Security, or Intangible Asset.* SSVS No. 1 provides professional guidance to AICPA members (members) who provide client services to estimate the value of a business, business ownership interest, security, or intangible asset (a subject interest). Although early application is encouraged, SSVS No. 1 is effective for client engagements to estimate value accepted by a member after January 1, 2008.

The culmination of years of deliberation and debate, SSVS No. 1 provides professional guidance to members with regard to (1) valuation engagement acceptance and planning considerations, (2) the development of the valuation analysis, and (3) the reporting of the value conclusions. For some members, the application of SSVS No. 1 may simply mean the documentation of valuation engagement procedures already performed at the member's firm or practice. However, for other members, the application of SSVS No. 1 may involve implementing a new set of firm valuation engagement procedures and practices.

It is noteworthy that the SSVS No. 1 standard applies to members from all disciplines (including, for example, audit, tax, consulting, personal financial planning, and litigation services) who perform business valuation (BV) services. The following are 10 implementation recommendations for the application of SSVS No. 1 into a member's existing BV practice.

1. *Designate one BV practitioner as the firm's SSVS No. 1 BV standard expert.* Obviously, each practitioner who provides client valuation services should read and be familiar with SSVS No. 1. However, one practitioner should serve as the firm's "go to guy/gal" on SSVS No. 1 training, implementation, interpretation, and quality control issues.

2. *Each practitioner who performs valuation services should have a copy of SSVS No. 1.* Obviously, the firm should have a copy of SSVS No. 1 in its library, along with copies of all other AICPA professional standards. In addition, the firm should have a sign-off procedure confirming that each BV practitioner has received—and has read—a copy of SSVS No. 1.

3. *All firm BV practitioners should meet (in person, if possible) to review the new requirements of SSVS No. 1 and to discuss the firm's proposed implementation procedures for the application of SSVS No. 1.* This procedure should help ensure both communication and consistency among the firm's BV practitioners.

4. *The firm BV practitioners should inform all firm partners and staff of the issuance of SSVS No. 1.* The BV practitioners may offer internal training on SSVS No. 1 (on a summary level, if appropriate) to all firm members who are interested. This internal communication may reinforce the awareness of all partners and staff as to (1) the professionalism of the firm's BV practice and (2) the breadth of the firm's BV client services.

5. *BV practitioners should communicate the content and the intent of SSVS No. 1 (on a summary level, if appropriate) to the firm's recurring valuation clients and referral sources.* This communication

Continued on page 2

Continued from page 3

departments may not always understand how to best handle data subject to legal discovery. The volume, complexity, and expense associated with electronic discovery may present enormous challenges for IT departments with limited resources, training, or experience. Because of the fragile nature of electronic evidence, a company should engage expert assistance if the IT staff lacks the requisite equipment, time, training, and experience to perform a best practices collection. An expert may also be necessary if calling an IT person as a witness at trial is undesirable or if a conflict of interest might hurt the case.

9. **Neglecting to Carefully Choose an Electronic Evidence Expert.** If the project requires highly trained and sophisticated technologies, and it is necessary to engage an outside expert, choosing that expert is an extremely important decision. Failing to choose an expert with the proper training, tools, and expertise could cost law firms and their clients unnecessary time delays and added expenses. When helping counsel and clients select an electronic evidence expert, consider how long the expert has been in business, whether the expert outsources any of its services, the number of electronic evidence projects the expert handles on a yearly basis, the expert's capacity to process paper and electronic documents, whether the expert has a secure online review and hosting solution integrating paper and electronic documents, and if the expert maintains strict quality control measures and has a record of quality deliverables. In larger cases, it may be prudent to actually visit the expert's facility to do a full inspection of their capabilities and facility security.

10. **Failing to Use an Online Repository Tool for Paper and Electronic Document Review.** The days of conducting hardcopy document review page by page and box by box are nearly over. Instead, litigation support teams should capitalize on the advancements in the document discovery marketplace by reviewing both paper and electronic documents in an online repository tool. Leveraging the Internet and a database of discovery documents, electronic document review saves time and money because reviewers can search, categorize, and produce documents in an electronic format. After narrowing the universe of data, reviewers can print various collections, convert them into local litigation support database load files, or save them natively. By using technology to integrate paper and electronic documents, law firms likely will reduce the amount of time, effort, and cost spent on document review and production.

Although electronic discovery can seem like a daunting journey into an unknown place, a solid strategy for handling electronic data will put you and your clients in the best position for avoiding discovery sanctions and ensuring that the electronic evidence is admissible should the case proceed to trial. Those who develop a solid discovery plan, monitor preservation requirements, and address potential discovery problems long before they actually occur will set the stage for a comprehensive, efficient, and seamless discovery process–ultimately allowing them to scale the highest peak to gain the strategic edge in their cases.

Jonathan Sachs is a Legal Consultant for Kroll Ontrack based in New York. Mr. Sachs assists attorneys and corporations with discovery and investigations involving electronically stored data and emails. ***The author gratefully acknowledges the assistance of Charity Delich, a Kroll Ontrack Law Clerk.***

The Application of Regression Analysis to the Direct Market Data Method

Part 5–Conclusion: How to Read, Understand, and Interpret Excel's Regression Output

By James A. DiGabriele, D.P.S., CPA/ABV, CFE, CFSA, DABFA, Cr.FA, CVA, and Mark G. Filler, CPA/ABV, CBA, AM, CVA

We now have the results of the regression equation that we have been working with in four earlier parts of this series (Part 1, August/September 2006; Part 2, October/November/December 2006; Part 3, March/April 2007; and Part 4, May/June 2007). At this point, you may be asking yourself: now what? The good news is that there are specific metrics included in the Excel summary regression output that will further explicate the results of the model. For this explication we will be using as our demonstration model the summary output available through Excel's regression tool found in its Analysis ToolPak. Please refer to Part 2, Figure 4 of this series for a sample summary output, as well as the paragraph in Part 2 (October/November/December 2006) that explains how to use the regression tool.

The summary output will be discussed in two sections: regression statistics and the analysis of variance (ANOVA). The regression statistics section illustrates the summary statistics of the regression equation, which includes Multiple R, R Square, Adjusted R Square, Standard Error, and Observations. The ANOVA section includes the analysis of variance considerations, including the F-statistic and F-significance, as well as the regression coefficients and p-values. In the following sections, each part of the summary output will be discussed and its applicability to the valuation assignment duly articulated. Please note that the summary output that follows was derived from Figure 3 of Part 3 (March/April 2007) of this series by regressing selling price against seller's discretionary earnings (SDE) for the 14 remaining data points (Nos. 1–13 and No. 15).

Regression Statistics

Table 1

Regression Statistics	
Multiple R	0.9053
R Square	0.8195
Adjusted R Square	0.8044
Standard Error	25.4433
Observations	14

Multiple R, or the coefficient of correlation, is equal to the absolute correlation between the observed values of the dependent variable Y (selling price) and the values of the independent variable X (SDE). It measures the strength of a linear relationship. The value of Multiple R lies between -1 and +1, and the closer your result is to 1, the stronger the relationship. As a result, large values of Multiple R represent a greater correlation between SDE and selling price. For example, a Multiple R value of 1 represents a model that is perfectly linear where all the points in a scatterplot lie on a straight line. In our example, the Multiple R is .9053, which is very close to 1. This indicates that SDE and selling price are highly correlated. As a rule of thumb, coefficients of correlation that are below .70 are not useful in a valuation setting.

R Square is a goodness-of-fit measure for a regression model that ranges between 0 and 1 and is also called "the coefficient of determination." R Square is the proportion of variation in Y (dependent) variable (selling price) that is explained by changes in the X (independent) variable (SDE). The value of .8195 suggests that 81.95% of the selling price of a business can be explained by the independent variable SDE. The remaining 18.05% is presumed to be random variation in the data. As a rule of thumb, coefficients of determination that are less than .50 are not useful in a valuation setting.

Since the addition of extra X variables into the regression equation has the result of making R Square larger, Adjusted R Square has been introduced to penalize those models that have extra X variables with no additional explanatory value. Since all of your models will have only one X variable, Adjusted R Square is something you don't have to bother with.

The Standard Error, also known as "root mean square error," provides an estimate of the distribution of the prediction errors when predicting Y values from X values in the regression model. In other words, the standard error measures the size of a typical deviation of an observed value from the regression line. Think of the standard error as a way of averaging the size of the deviations from the regression line. The larger the value, the less well the regression model fits the data, and therefore, the model will not be as good at predicting the outcome as would be a lower standard error model. It has been said that for a successful regression model, the standard error of the estimate should be considerably smaller than the standard deviation of the dependent variable. In other words, the observations should vary less about the regression line than about the mean.

Table 2

ANOVA

	df	*SS*	*MS*	*F*	*Signifcance F*
Regression	1	35,265.3953	35,265.3953	54.4757	0.0000
Residual	12	7,768.3190	674.3599		
Total	13	43,033.7143			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-3.7101	21.4116	-0.1733	0.8653	-50.3621	42.9419
SDE	1.8849	0.2554	7.3808	0.0000	1.3285	2.4414

We have taken the task of calculating the standard deviation of our Y (dependent) variable. The standard deviation of this variable is 57.54. Since the standard error for our regression model is 25.44, it is fair to state that the model is very good at predicting selling prices based on SDE. The standard error can also be used to calculate the coefficient of variation (COV), which is the standard error divided by the average of Y (the dependent variable). In this case, the average selling price is $146.14, making the COV 17.4%, an excellent outcome when using the Bizcomps database because COVs as high as 25 – 30% are very common.

The observation value is the size of the sample used in the regression. In this case, the regression is based on the values from 14 market transactions.

ANOVA

The top half of the ANOVA table (Table 2) tells us if the overall regression model results in a significantly acceptable level of predictability for the outcome (dependent) variable. The bottom half of the ANOVA table informs us if the slope of the regression line is different from zero, and therefore, whether or not we have a statistically significant regression model. The top half of the ANOVA table analyzes the variability of the selling prices. The variability is divided into two parts: the first is the variability due to the regression line and the second is due to random variation. This is shown in the summary table by use of the various sums of squares (SS). Let's go through each of them.

As a strategy for predicting an outcome, for lack of a better estimate, one may choose to use the mean as a fairly good guess. By substituting the mean as a model, we can calculate the difference between the observed values and those values predicted by the mean.

The Regression row of the SS column refers to differences between the mean value of the outcome (dependent) variable Y and the regression line. If this value is large, then the regression model is different from the mean, which is our best guess as to the outcome. On the other hand, if this number is small, then using the regression model is just a little better than using the mean as an estimate.

The Residual row of the SS column explains the differences between the observed data and the regression line. This value represents the degree of error when the regression model is fitted to the data. A low number here relative to Regression SS indicates a model that fits the data well.

The Total represents the sum of squared differences about the mean. The figure indicates how good the mean is as a model of the observed data.

At this point, you may be asking yourself: why is this SS stuff useful? The first use of these numbers is that R Square can be calculated by dividing Regression SS by Total SS (35,265.3953/43,033.7143 = .8195). As we already know, R Square is the proportion of variation in the Y (dependent variable) that is explained by the X (independent) variable.

Let's move over to the column with the heading "MS." The numbers in this column can be defined as "the mean sum of squares for Regression and

Continued on page 6

Continued from page 5

Residual" and are easily calculated by dividing the SS numbers by the degrees of freedom (df) column. The practical use for these numbers is to calculate the F-ratio. In the ANOVA table, there is a column for the F-statistic, which is a measure of how much the model has improved our ability to predict the outcome compared with just using the mean of the dependent variable as a predictor. The calculation is Regression MS divided by Residual MS. If the model provides a good overall fit, we would expect the improvement in the prediction due to the model to be large. That is, the Regression MS would be large, and the difference between the model and observed data would be small (Residual MS). As a result, a good model should have a large F-ratio (greater than 1). In addition, the significance of the F-ratio is assessed using critical values (p-values). In our model, the F-ratio is statistically significant as the Significance F number is less than .05. However, in a model with a single X coefficient (SDE in this case), F and Significance F are redundant just like Adjusted R Square, because the t-statistic is the square root of F (or in reverse $7.3808^2 = 54.4757$).

Just as R Square can be derived from the ANOVA table, so can the Standard Error be derived by simply taking the square root of Residual MS (or in reverse, $25.44^2 = 647.36$).

In the bottom part of the ANOVA table the most important numbers are the coefficients for the intercept and SDE. These two numbers represent the point of interception on the Y axis and the slope of the least squares regression line, respectively. With these coefficients, our regression equation now becomes:

$$Y = 1.8849x + -3.7101.$$

Now let's assess the individual predictor (independent) variable, SDE. The t-statistic (which measures the number of standard deviations from zero that the SDE coefficient is, and is computed by dividing the Coefficient by its Standard Error) tests the null hypothesis that the value of this variable is zero. If the variable has a significant p-value (less than .05), we would accept that the value is significantly different from zero, and therefore the independent variable contributes significantly to our ability to predict the value of a selling price for any particular business. In our case, SDE is statistically significant, because its p-value is less than .05, and its t-statistic is greater than 2. So, it is safe to say that SDE contributes significantly to our model, that is, it is significantly greater than zero, and therefore the model is a better predictor of value than the average selling price of the 14 businesses in our database.

Just a quick note on the intercept. A t-statistic of -.173 (less than 2.0) and a p-value greater than .05 ($p = .8653$) indicates that the intercept does not differ from zero, and therefore the regression line goes through the origin (the point where the X and Y axes meet). The interpretation of the intercept is less important than that of the X variable. It is literally the predicted selling price when there is no SDE. However, none of the observations in our 14-market transaction sample had an SDE of zero. Therefore, in a situation like this, where the range of independent variables does not include zero, it is best to think of the intercept term as an "anchor" for the regression line that allows us to predict selling prices for the range of observed SDE values.

The remaining item to be explained in the regression output is the 95% limits. These limits allow us to report with 95% confidence that for each $1 increase in SDE, the selling price of any particular business increases between $1.33 and $2.44.

Conclusion

This series of articles was intended to introduce practitioners to the statistical method of regression analysis and to demonstrate how this procedure can improve their valuations, especially when used in combination with an Income Method. This technique has always been a popular tool of economists. Recently, however, regression analysis has also found its way into the courts as evidence of damages in contractual actions, torts, and antitrust cases. These developments should further emphasize the importance to practitioners of understanding this technique.

In this series of articles, we focused on bivariate simple regression analysis, and although many other forms of RA are available, the tools we provided in this series are all that you will need to competently apply RA in the use of the Direct Market Data Method and derive good valuation results. The authors feel so strongly that RA is the best way to get valuation results using the Direct Market Data Method that they are willing to answer your e-mail–submitted questions, at no charge, regarding the application of the theory and practice demonstrated in this series of articles.

James A. DiGabriele, D.P.S., CPA/ABV, CFE, CFSA, DABFA, Cr.FA, CVA, is Assistant Professor in the Department of Accounting, Law, & Taxation, School of Business, Montclair State University, Montclair, NJ 07042. Phone: (973) 243-2600; fax: (973) 243-2646; e-mail: jim@dmcpa.com

Mark G. Filler, CPA/ABV, CBA, AM, CVA, is founder of Filler & Associates, P.A. Portland, ME 04101. Phone: (207) 772-0153, x222; fax: (207) 761-4013; e-mail: mfiller@filler.com

Financial Statement Fraud: A Collaborative Effort

In instances of financial statement fraud, the number of organizations and individuals involved typically averages 7.2. This was "one of the main themes" that emerged from the study conducted by Robert Tillman and Michael Indergaard of St. John's University (Queens, NY) and reported in "Control Overrides in Financial Statement Fraud," which can be downloaded from the Web site of the Institute for Fraud Prevention (see "The Institute for Fraud Prevention" on page 7).

Another significant finding was that, of the organizations which were defendants and respondents in class action lawsuits or SEC actions, more than half were not the restating firms. Many were accounting firms and banks.

Tillman and Indergaard conclude from their case studies that the relationship between the restating firm's senior managers and their auditors cannot be characterized simply as "collusion or no collusion." More important in the relationships was "the extent to which external auditors resisted efforts by senior managers to engage in fraudulent financial reporting and whether that resistance was consistent or inconsistent."

They also conclude that the "reputational penalty" theory often fails to deter fraud. Under this theory, directors and auditors are unlikely to cooperate with senior managers to deceive shareholders. The reason is they fear tarnishing