Final Paper

Due: December 16th, 2023

A statistical project as to the stock price of…
Baker Hughes Company Class A

Brianna L. Palmisano (X03625986)

St. John's University, The Peter J. Tobin College of Business

BUA 602: Business Analytics (Online, Asynchronous)

Professor Manual G. Russon, PhD, CFA

Fall 2023

TABLE OF CONTENTS

I. Introduction

This paper examines the linear relationship between the stock price and both the earnings per share (EPS) and book value per share (BVPS) of the company Baker Hughes Company Class A. Baker Hughes is an energy technology company that specializes in the development and deployment of advanced technology for energy and industrial companies. The company focuses on creating long-term solutions that are both reliable and efficient for both the environment and global energy needs.[1] Information regarding the stock price, EPS and BVPS of Baker Hughes could be adventitious for investors in energy technology or innovation, as well as stock traders following the Energy sector of the investing.

II. Previous Research

There has been previous research on this company and others in its market, though none directly in-line with the objective of this paper.

III. Methodology

The input data here has been collected from the Energy Equipment & Services industry, in the Energy sector of the S&P500. The S&P500 is a stock market index of the top 500 leading publicly traded companies in the United States.[2] This is secondary data, as the data had been reports, collected, calculated, and adjusted for inflation (via CPI).[3]

The earnings per share (EPS) for the company is calculated by dividing the company's total profits by the number of outstanding shares of its common stock are being publicly traded, and therefore serves as a measure of corporate value.[4] Similarly, the book value per share (BVPS) is calculated by dividing the ratio of the company's equity available to common shareholders (the amount of equity held by each shareholder) by the number of outstanding shares of its common stock that are being publicly traded. BVPS is the main measure used by stock investors to evaluate a company's stock price.[5]

Here we have a total of 23 observations, all reported on the last day of the year (December 31st), once per year, each year from 2000 to 2022. Graphical techniques including histograms, time series plots and scatterplots are used in this analysis. This data has been

[1] https://www.bakerhughes.com
[2] https://www.investopedia.com/terms/s/sp500.asp
[3] https://www.macrotrends.net/2324/sp-500-historical-chart-data#:~:text=Historical%20data%20is%20inflation%2Dadjusted,December%2006%2C%202023%20is%204%2C549.34.
[4] https://www.investopedia.com/terms/e/eps.asp
[5] https://www.investopedia.com/terms/b/bvps.asp#:~:text=Book%20value%20per%20share%20(BVPS)%20is%20the%20ratio%20of%20equity,on%20a%20per%2Dshare%20basis.

analyzed using descriptive statistics (for scalable variables), as well as correlation and regression statistical analysis via Microsoft Excel.

The dependent variable here is the price per share, and the independent variables are the EPS and BVPS. The relationship between these variables is that the price of a share is influenced by the amount of return on equity a shareholder has in that company's profits (EPS), as well as the amount investors are willing to pay for that amount of profits (in comparison to the price of those shares). It can be assumed that the level of EPS drives the BVPS because the calculation of BVPS is inclusive of the cost of the profits that are EPS.

Listed below are the equations defining the functional specification (Eqn. 1), population regression equation (Eqn. 2) and sample regression equation (Eqn. 3).

Eqn. 1 $\quad$ Price = $\quad$ f(EPS, BVPS)

Eqn. 2 $\quad$ Price = $\quad \alpha + \beta_{eps}\,EPS + \beta_{bvps}\,BVPS$

Eqn. 3 $\quad$ Price = $\quad a + b_{eps}\,EPS + b_{bvps}\,BVPS$

When the BVPS is above the price per share, that stock is considered undervalued or an underpriced stock and vice versa. When the EPS is increased, the price of that stock is increased as well, because investors are now willing to pay more for that stock.[6] This could be for a list of reasons pertaining to company growth. The objective of this model is to evaluate the dependency or lack thereof that the stock price of Baker Hughes has on its EPS and BVPS, or both.

IV. Results

In the histogram of all variables being analyzed, *Figure 1*, there is an observably stronger pattern between Price and BVPS, as there are more similarities in the fluctuations of Price vs. BVPS than with EPS. It can be initially determined that EPS is less correlated than this relationship. The data is roughly left, or positively skewed until recent years, therefore considerably irregular.

We can also assume that the company lost profits going into 2016, 2021 and the current period. All variables were reasonably constant leading up to around 2007 or 2008. It can also be assumed that this company did better during the financial crisis of 2008 than during the similar economic shock experienced during the COVID-19 pandemic. It is possible that the struggles this company and industry inhibited prior to the pandemic were the initial factors weighing on their loss of profits in 2021.

The company seems to break even most years of the relative time period, though the Price and BVPS does not seem to be affected by the dips in levels of Baker Hughes profitability. BVPS could have potentially held the company's EPS at a break-even by not being adjusted entirely relative to these decreases. It is interesting to see that the Price has fallen below the

[6] https://www.investopedia.com/terms/b/bvps.asp

BVPS since 2018, insisting that this stock is undervalued compared to its company growth and profitability. This could be in anticipation of future loss of profits or plans of company degrowth.

In *Figures 2 –4*, sequence charts have been created for each variable, both independent and the dependent, over 23 years. It can be observed that when placed individually with the year, the dependent variable reacts to changes in the independent variable EPS more so than BVPS, and the independent variables shows even less of a visible pattern.

Figure 2 and *Figure 3* can be considered right or positively-skewed, whereas *Figure 4* is left or negatively-skewed, supporting the assumption that Price is more related to EPS than BVPS. *Figure 2* has more peaks than *Figures 3 & 4*, possibly due to outliers, or outside variables other than our independent variables (measurable and unmeasurable) affecting the price of shares. When there are outliers, it should also be assumed that there is room for error in the input of the dataset or the company's reporting of this data.

Should you follow the intervals of increases and decreases in these time series charts, there seems to be a delayed reaction between changes in price and EPS, relative to the BVPS; the BVPS changes similarly to the Price and EPS but within a year's time, opposed to directly. This does show that there is a slight pattern between all variables, but more so a ripple effect in the variable's changes from the Price and EPS to BVPS.

All variables are considerably volatile (or non-cyclical) for their lack of repeated pattern, but Price (Figure 2) is the most volatile and the least consistent variable over time. *Figures 3 & 4* (EPS and BVPS) can be loosely considered linear, though *Figure 2* (Price) cannot be considered linear as it does not run along the displayed linear trendline consistently, whatsoever.

The relationships visible via scatterplot *Figures 5 & 6* is the independent variables each relative to the Price or dependent variable. In both figures, the relationships is weak, but the data is a stronger or more positively correlated in *Figure 5*, *Price vs. EPS*. Price and EPS are more heteroscedastic, though there are a handful of outliers in both figures. There are more outliers seen in *Figure 6* (Price vs. BVPS). Price and BVPS are barely positively correlated showing that the Price is highly influenced outside that of BVPS.

Descriptive Statistics

In *Table 1*, are the descriptive statistics for all non-categorical variables or all variables in this model. The mean is greater than the median in both Price and BVPS, meaning that they are positively skewed, while EPS is negatively skewed. This is not in support of the initial assumption that Price is more closely related to EPS than BVPS.

The measure of kurtosis here is negative in Price vs. BVPS, and positive in the Price vs. EPS, in support of the observations drawn from symmetry. The peakedness here is rightward, and lower around EPS relative to Price. Both variables are light-tailed in terms of distribution.

Correlation Matrix

Using the regression correlation coefficient of the matrix in *Table 2*, it can be observed that the stock price of Baker Hughes is much more correlated with the EPS (at 0.524) than with

the BVPS (at 0.129). It should also be noted that the BVPS is more correlated with the Price (at 0.129) than it is with the EPS (at 0.089) and vice versa. This matrix further proves the initial hypothesis that Price is more effected by EPS or has a higher correlation between this independent variable than the latter. There are no signs of multicollinearity in this model.

Regression Results

Equation for Sample Regression Line.

			+	+	-
Eqn. 4	Price =	38.49	+ 2.07EPS	+	.12BVPS
	t-stat		(5.25)**	(2.71)***	(0.44)*
	p-value		(0)	(.01)	(.67)
	r (corr)		(2.07)	(.12)	

n = 23 r-sq. = .281 F = 3.91*** F-Prob = .037 SE = 14.865

Confidence Intervals.

*	Significant at the	10%	level of significance (90% Sure, or "are below")
**	Significant at the	5%	level of significance (95% Sure, or "are below")
***	Significant at the	1%	level of significance (99% Sure, or "are below")

Results of an F-test for the entire model.

	Ho:	$b_{eps} = b_{bvps} = 0$	(Null Hypothesis)
* 5%	Ha:	at least 1 b_i not equal to 0 (63.91 > 4.99)	(Alternate Hypothesis)

To further interpret the regression statistics in *Table 3*, the standard deviation is less than the mean for both the Price and BVPS, meaning that the data for these variables is closely dispersed around the mean data or exhibit low dispersion compared to the EPS.

The above F-test includes the null (Ho) and alternative hypothesis (Ha). Because the F-Significance is greater than the p-value (0.0), we should assume that the model does effectively prove that the dependent variable is more correlated when analyzed relative to the independent variables of this model than if it weren't. The p-value also proves that this model is more significant when looking at Price vs. EPS (at 0.01) than BVPS, as it is greater than 0.05 in BVPS (at .067). The F-Statistic is above the

When looking at the results for the coefficient of determination, (R-Squared, equal to 0.281 or 28.1%) it can be concluded that that statistical model predicts that the association between the dependent and independent variables is weak; there is little variation in the dependent variable that can be associated with the independent variables. There is 28.1% of variation in the dependent variable to be associated with variations in the independent variables.

The high standard error value of the independent variable EPS (at 0.76) tells us that the EPS falls much further from the regression line than BVPS does (at 0.28). The greater Standard Error value seen in the intercept or price determines that the dependent variable is being affected by other independent variables, ones not included in this model.

The regression coefficient is much lower in the independent variables than the dependent or intercept. The amount of change required in the Price to create change in either independent variables is high, showing little relationship. The coefficient for EPS is much higher than it is for BVPS, suggesting that EPS is a better fit for this model or that changes in Price are more closely correlated with changes EPS than that of BVPS.

V. Conclusions

For this model, the null hypothesis is accepted, and the alternate hypothesis is rejected. The research presented here can be considered successful, though the dependent variable is only 28.1% associated with the independent variables in terms of variation. This model was not predictive, as the variables here do not influence one-another enough. The initial assumption that Price is more correlated with EPS is true, but it was violated when looking at some of the regression results in the context of dispersion and symmetry. It can be determined true that the dependent variable is better supported by the independent variables, than if it were not.

This model can be improved on by including additional variables that go into the price of a company's stock, such as price-to-earnings, price-to-earnings growth, or price-to-book ratio. These variables may have a higher correlation with pricing shares in terms of the supply and demand of that company. Increasing the number of observations would increase the validity of any model but it can be assumed that that would not change our overall results here, as the variables are too volatile. Should there be a need to break down the moments of closer associations, breaking this model up by time periods may add to the findings this model is capable of.

Public policy does not entirely pertain to driving a relationship between these variables, nor does it serve much prominence outside of regulating the terms of trade. It could be beneficial to investors to have more insight on the variables that influence the stock price of the company at hand. As, this model or an expanded version of this model could be beneficial to policy makers considering information allowance to investors of a company.

VI. Bibliography

"We Are Baker Hughes, an Energy Technology Company." *Baker Hughes*, 12 Dec. 2023, www.bakerhughes.com/.

VII. Appendix I: Input Data

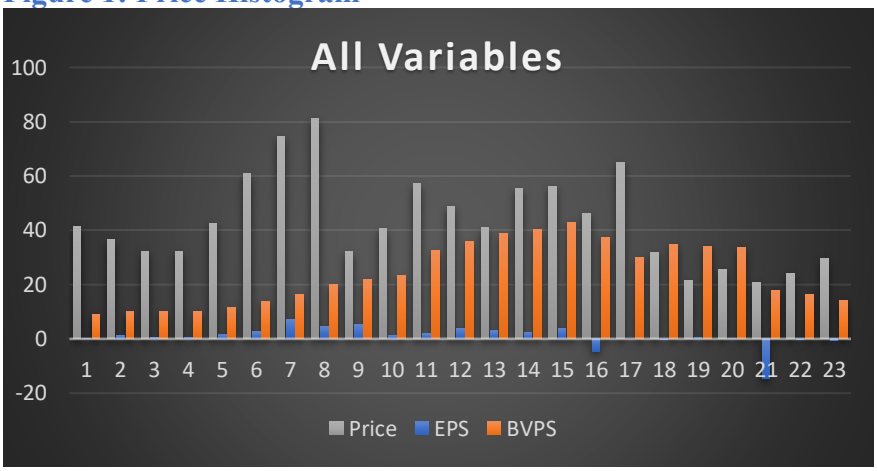
Dataset used in this model.

Name	Sector	Industry	Date	EPS	BVPS	Price
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2000	0.31	9.13006	41.5625
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2001	1.3	9.90417	36.47
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2002	0.66	10.11674	32.19
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2003	0.54	10.09157	32.16
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2004	1.57	11.57279	42.67
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2005	2.5696	13.75637	60.78
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2006	7.2153	16.38918	74.66
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2007	4.73	19.99239	81.1
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2008	5.2913	22.02913	32.07
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2009	1.3537	23.34615	40.48
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2010	2.06	32.63889	57.17
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2011	3.97	36.03204	48.64
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2012	2.97	38.70522	40.8477
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2013	2.47	40.44064	55.26
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2014	3.92	42.91475	56.07
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2015	-4.49	37.29519	46.15
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2016	0	29.84906	64.97
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2017	-0.24	34.85545	31.64
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2018	0.45	34.04483	21.5
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2019	0.2298	33.73692	25.63
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2020	-14.7259	17.80801	20.85
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2021	-0.2658	16.31463	24.06
Baker Hughes Company Class A	Energy	Energy Equipment & Services	12/31/2022	-0.6089	14.30815	29.53

CLICK HERE: Link to Excel Attachment

VIII. Appendix II: Figure 1: Histograms *(All Variables)*

Figure 1: Price Histogram



IX. Appendix III: Figures 2 - 4: Timeseries *(All Variables)*

Figure 2: Price Timeseries

Figure 3: Earnings Per Share Timeseries

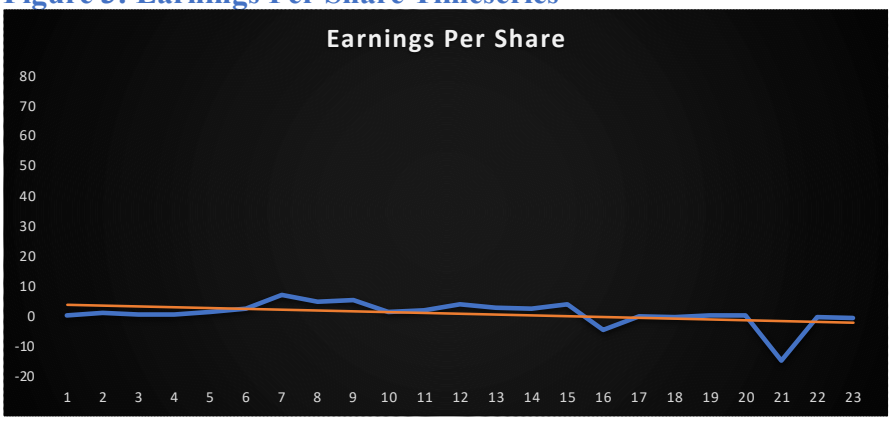
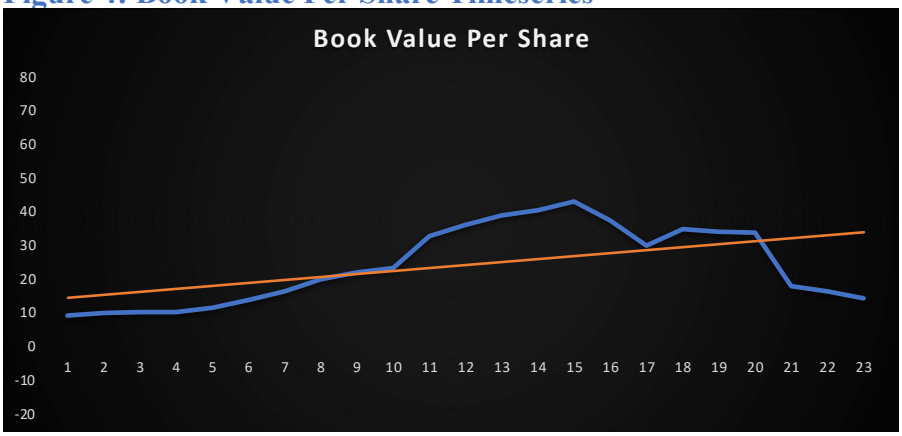


Figure 4: Book Value Per Share Timeseries



X. Appendix IV: Figures 5 & 6: Scatterplots *(Dependent Variable v. Each Independent Variable)*

Figure 5: Price vs. Earnings Per Share

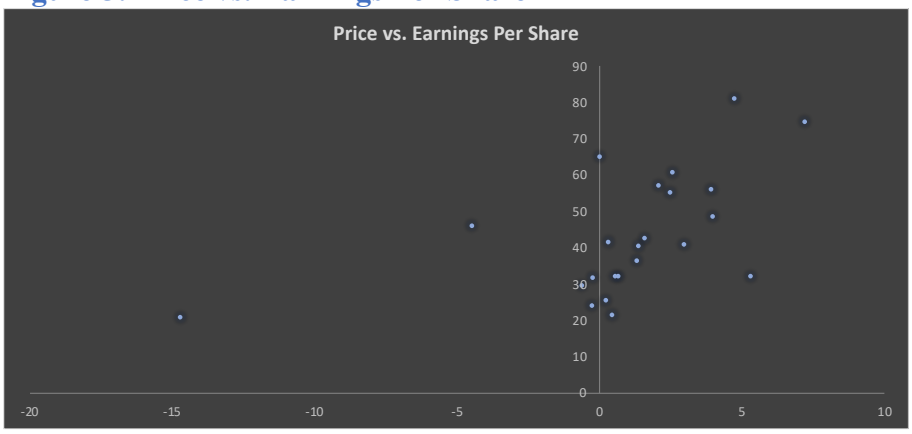
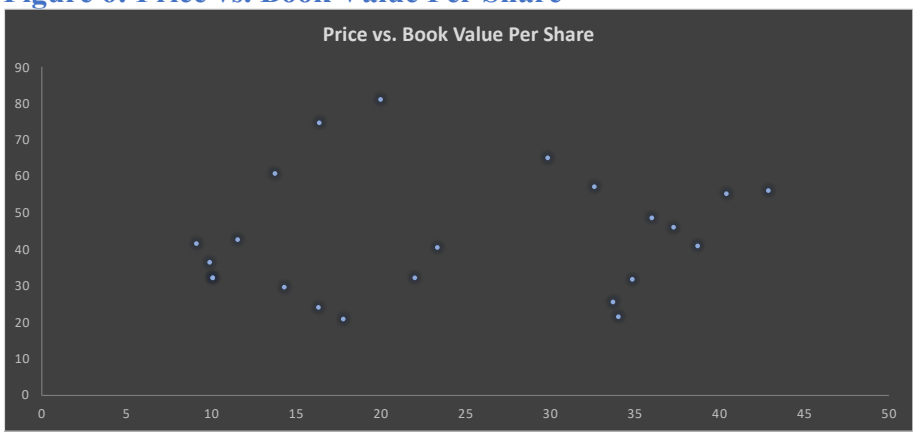


Figure 6: Price vs. Book Value Per Share



XI. Appendix V: Table 1: Descriptive Statistics *(Dependent Variable & Independent Variables)*

Table 1: Descriptive Statistics

	price	*eps*	*bvps*
Mean	43.32	0.93	24.14
Standard Error	3.49	0.87	2.40
Median	40.85	1.30	22.03
Standard Deviation	16.72	4.18	11.53
Sample Variance	279.43	17.46	132.85
Kurtosis	(0.23)	8.92	(1.59)
Skewness	0.68	(2.41)	0.15
Range	60.25	21.94	33.78
Minimum	20.85	(14.73)	9.13
Maximum	81.10	7.22	42.91
Sum	996.46	21.28	555.27
Count	23	23	23

XII. Appendix VI: Table 2: Correlation Matrix *(All Variables)*

Table 2: Correlation Matrix (Correlation Coefficients)

	eps	*bvps*	*price*
eps	1.000		
bvps	0.089	1.000	
price	0.524	0.129	1.000

XIII. Appendix VII: Table 3: Regression Results *(Dependent Variable & Independent Variables)*

Table 3: Regression Results

	df	*SS*	*MS*	*F*	*Significance F*			
Regression	2	1,728.104	864.052	3.910	0.037			
Residual	20	4,419.333	220.967					
Total	22	6,147.438						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	38.49	7.33	5.25	0.00	23.21	53.78	23.21	53.78
eps	2.07	0.76	2.71	0.01	0.48	3.65	0.48	3.65
bvps	0.12	0.28	0.44	0.67	(0.45)	0.70	(0.45)	0.70