# A Statistical Model to Estimate Equity Price in the Healthcare Providers and Services Industry

By: Laura Adeniyi

## I.Introduction

The Health Care Providers & Services industry has 44 companies in the S&P 1500. The purpose of this research is to develop a statistical model that will predict the equity price in the Health Care Providers & Services industry. This model will be of high use to portfolio managers, analysts, investment bankers and stockholders.

## III. Methodology

This research utilizes cross-sectional data specifically for the year 2022, obtained from FactSet.com, comprising a total of 44 observations. To explore and analyze the data, graphical techniques such as histograms and scatter plots are employed. Furthermore, analytical techniques including descriptive statistics for scalable variables, a correlation matrix, and a multivariate regression are utilized for further analysis. This research uses R to perform the statistical analysis.

Eqn. 1: Functional specification

$$\text{Eqn. 1} \qquad \text{Price} = f(\text{EPS}, \text{BVPS}) \quad (+, +)$$

Eqn. 2: Population regression equation

$$\text{Eqn. 2} \qquad \text{Price} = \alpha + \beta_{eps}*\text{EPS} + \beta_{bvps}*\text{BVPS} + \varepsilon$$

Eqn. 3: Sample regression equation

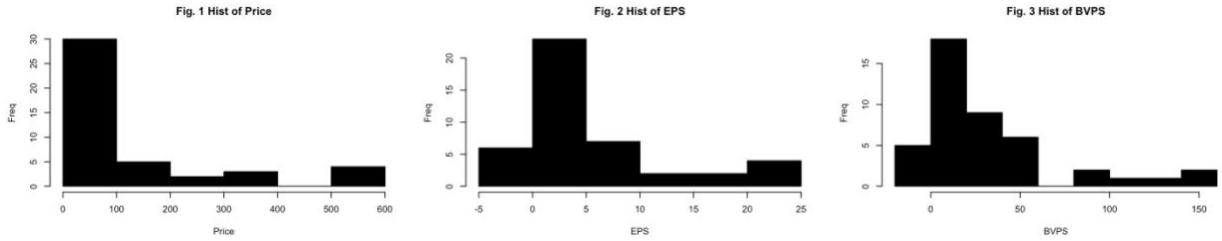
$$\text{Eqn. 3} \qquad \text{Price} = a + b_{eps}*\text{EPS} + b_{bvps}*\text{BVPS} + e$$

Price is hypothesized to be a positive function of earnings per share (EPS) and book value per share (BVPS).

## IV. Results

Fig. 1 – Fig. 3 depict the histograms of Price, EPS, and BVPS.

***Fig. 1 – Fig. 3: Histograms of Price, EPS and BVPS***




All three of the variables are skewed to the right and show the presence of outliers.

Fig. 4 – Fig. 5 show the relationship between BVPS and Price as well as EPS and Price.

***Fig. 4 – Fig. 5: Scatterplots of Each Dependent Variables (BVPS, EPS) with the Independent Variable (Price)***

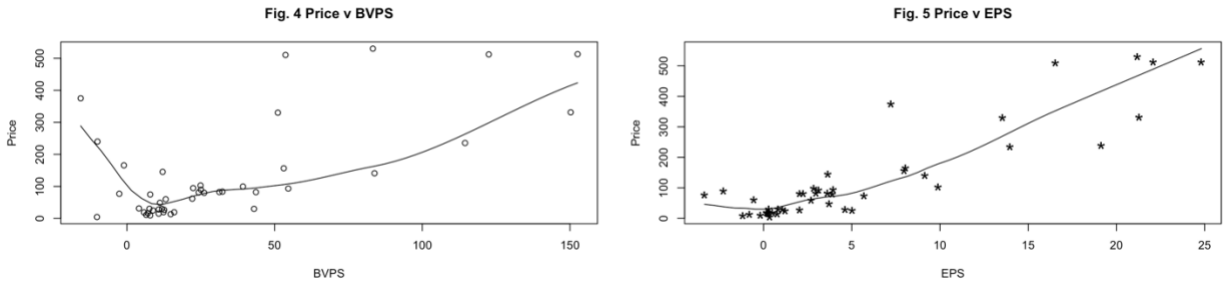



Fig. 4 displays a moderate, positive, non-linear relationship between price and book value per share with several outliers and no heteroscedasticity. Fig. 5 displays a strong, positive, non-linear relationship between price and earnings per share with few outliers and signs of heteroscedasticity.

Table 1 shows the descriptive statistics.

***Table 1: Descriptive Statistics***

| | *name* | *obs* | *max* | *min* | *mean* | *median* | *std* | *skew* | *kurt* |
|---|---|---|---|---|---|---|---|---|---|
| ***1*** | price | 44 | 530.18 | 4.32 | 130.07 | 80.45 | 152.25 | 1.55 | 4.39 |
| ***2*** | eps | 44 | 24.81 | -3.34 | 5.56 | 3.06 | 7.23 | 1.25 | 3.61 |
| ***3*** | bvps | 44 | 152.58 | -15.67 | 31.72 | 15.40 | 39.84 | 1.63 | 5.28 |

The correlation matrix between all the variables is presented in Table 2.

***Table 2: Correlation Matrix***

| | *price* | *eps* | *bvps* |
|---|---|---|---|
| ***price*** | 1.0000 | 0.8939 | 0.6674 |
| ***eps*** | 0.8939 | 1.0000 | 0.7465 |
| ***bvps*** | 0.6674 | 0.7465 | 1.0000 |

There is a moderate correlation between Price and BVPS of 0.67, and a strong correlation between Price and EPS of 0.89. The demonstrated correlations agree with the original hypothesis that price is hypothesized to be a positive function of EPS and BVPS. Note, there is a slight potential of multicollinearity, with the correlation of 0.75 between EPS and BVPS.

Table 3 shows the regression results.

***Table 3: Regression Results***

***Coefficients:***

| | ***Estimate*** | ***Std. Error*** | ***t value*** | ***Pr(>\|t\|)*** |
|---|---|---|---|---|
| ***(Intercept)*** | 25.411 | 13.818 | 1.839 | 0.073 * |
| ***eps*** | 18.821 | 2.216 | 8.492 | 1.41e-10 *** |
| ***bvps*** | 0.001 | 0.402 | 0.002 | 0.998 |

n = 44 r-sq. = 0.799 F = 81.48 SE = 69.91

* - Significant at the 10% level of significance
** - Significant at the 5% level of significance
*** - Significant at the 1% level of significance

***Interpretation***

Eqn. 4: Estimated sample regression equation

Price = 25.411+18.821*EPS+0.001*BVPS

For every one dollar increase in EPS, price increases by $18.821. For every one dollar increase in BVPS, price increases by $0.001.

1. The F-statistic of 81.48 > 4.99, therefore, the entire equation is statistically significant.
2. As determined by r-squared, 79.9% of the variation in price is explained by the variation in EPS and BVPS. The explanatory power is high.
3. The regression coefficients for the intercept and EPS are statistically significant at least at the 95% confidence level (t-stat >1.66). The regression coefficient for BVPS is not statistically significant.
4. The standard error is 69.91

## V. Conclusion

The research was successful as it demonstrated a high degree of explanatory power. The results align with the original hypothesis that price is a positive function of earnings per share and book value per share. However, only EPS proved to be predictive. This research can be improved by increasing the number of observations, adding longitudinal data to compare the same sample at different points in time, and including additional independent variables

## VII. Appendix I: The Dataset

| tkr | price | eps | bvps | year | name | industry |
|---|---|---|---|---|---|---|
| ABC | 165.71 | 8.0433 | -1.02597 | 2022 | AmerisourceBergen Corporation | Health Care Providers & Services |
| ACHC | 82.32 | 2.9833 | 31.28253 | 2022 | Acadia Healthcare Company, Inc. | Health Care Providers & Services |
| ADUS | 99.49 | 2.8444 | 39.28199 | 2022 | Addus HomeCare Corporation | Health Care Providers & Services |
| AGTI | 16.31 | 0.2183 | 7.08135 | 2022 | Agiliti, Inc. | Health Care Providers & Services |
| AHCO | 19.22 | 0.4987 | 16.00145 | 2022 | AdaptHealth Corp. | Health Care Providers & Services |
| AMED | 83.54 | 3.6324 | 32.33791 | 2022 | Amedisys, Inc. | Health Care Providers & Services |
| AMEH | 29.59 | 1.0756 | 11.86932 | 2022 | Apollo Medical Holdings, Inc. | Health Care Providers & Services |
| AMN | 102.82 | 9.8964 | 24.92022 | 2022 | AMN Healthcare Services, Inc. | Health Care Providers & Services |
| CAH | 76.87 | -3.3441 | -2.59707 | 2022 | Cardinal Health, Inc. | Health Care Providers & Services |
| CCRN | 26.57 | 5.0208 | 12.59466 | 2022 | Cross Country Healthcare, Inc. | Health Care Providers & Services |
| CHE | 510.43 | 16.5325 | 53.69585 | 2022 | Chemed Corporation | Health Care Providers & Services |
| CI | 331.34 | 21.2991 1 | 50.23638 | 2022 | Cigna Group | Health Care Providers & Services |
| CNC | 82.01 | 2.0652 | 43.68012 | 2022 | Centene Corporation | Health Care Providers & Services |
| CRVL | 145.33 | 3.6636 | 12.08913 | 2022 | CorVel Corporation | Health Care Providers & Services |

| | | | | | | |
|---|---|---|---|---|---|---|
| CVS | 93.19 | 3.136 | 54.62692 | 2022 | CVS Health Corporation | Health Care Providers & Services |
| CYH | 4.32 | -0.3537 | 10.1482 | 2022 | Community Health Systems, Inc. | Health Care Providers & Services |
| DGX | 156.44 | 7.983 | 53.09009 | 2022 | Quest Diagnostics Incorporated | Health Care Providers & Services |
| DVA | 74.67 | 5.7072 | 7.87875 | 2022 | DaVita Inc. | Health Care Providers & Services |
| EHAB | 13.16 | -0.8064 | 14.80448 | 2022 | Enhabit, Inc | Health Care Providers & Services |
| EHC | 59.81 | 2.6992 | 13.1315 | 2022 | Encompass Health Corporation | Health Care Providers & Services |
| ELV | 512.97 | 24.8147 1 | 52.5773 | 2022 | Elevance Health, Inc. | Health Care Providers & Services |
| ENSG | 94.61 | 3.9507 | 22.40944 | 2022 | Ensign Group, Inc. | Health Care Providers & Services |
| FLGT | 29.78 | 4.6313 | 43.02891 | 2022 | Fulgent Genetics, Inc. | Health Care Providers & Services |
| HCA | 239.96 | 19.1505 | -9.97555 | 2022 | HCA Healthcare Inc | Health Care Providers & Services |
| HQY | 61.64 | -0.5327 | 22.11238 | 2022 | HealthEquity Inc | Health Care Providers & Services |
| HSIC | 79.87 | 3.9055 | 26.1471 | 2022 | Henry Schein, Inc. | Health Care Providers & Services |
| HUM | 512.19 | 22.0782 1 | 22.51285 | 2022 | Humana Inc. | Health Care Providers & Services |
| LH | 235.48 | 13.9640 1 | 14.47392 | 2022 | Laboratory Corporation of America Holdings | Health Care Providers & Services |
| MCK | 375.12 | -7.2291 | 15.66897 | 2022 | McKesson Corporation | Health Care Providers & Services |

| | | | | | | |
|---|---|---|---|---|---|---|
| MD | 14.86 | 0.7886 | 10.74942 | 2022 | Pediatrix Medical Group, Inc. | Health Care Providers & Services |
| MODV | 89.73 | -2.2619 | 25.04564 | 2022 | ModivCare Inc. | Health Care Providers & Services |
| MOH | 330.22 | 13.5385 | 51.10345 | 2022 | Molina Healthcare, Inc. | Health Care Providers & Services |
| NEO | 9.24 | -1.1613 | 7.86377 | 2022 | NeoGenomics, Inc. | Health Care Providers & Services |
| OMI | 19.53 | 0.2938 | 12.39665 | 2022 | Owens & Minor, Inc. | Health Care Providers & Services |
| OPCH | 30.09 | 0.8269 | 7.61772 | 2022 | Option Care Health Inc | Health Care Providers & Services |
| PDCO | 28.03 | 2.0628 | 10.76534 | 2022 | Patterson Companies Incorporated | Health Care Providers & Services |
| PGNY | 31.15 | 0.3037 | 4.06719 | 2022 | Progyny, Inc. | Health Care Providers & Services |
| RCM | 10.95 | -0.1635 | 6.53028 | 2022 | R1 RCM Inc | Health Care Providers & Services |
| RDNT | 18.83 | 0.1732 | 5.76883 | 2022 | RadNet, Inc. | Health Care Providers & Services |
| SEM | 24.83 | 1.2307 | 8.82195 | 2022 | Select Medical Holdings Corporation | Health Care Providers & Services |
| THC | 48.79 | 3.7189 | 11.16908 | 2022 | Tenet Healthcare Corporation | Health Care Providers & Services |
| UHS | 140.89 | 9.1405 | 83.82656 | 2022 | Universal Health Services, Inc. Class B | Health Care Providers & Services |
| UNH | 530.18 | 21.179 | 83.26767 | 2022 | UnitedHealth Group Incorporated | Health Care Providers & Services |
| USPH | 81.03 | 2.2535 | 24.2888 | 2022 | U.S. Physical Therapy, Inc. | Health Care Providers & Services |

## VIII. Appendix II: R Script

```
library("YRmisc")
library("readxl")

1. import spInfox excel file
spInfox <- read_excel("~/Desktop/BUA 633/Class 3 June 1/spInfox.xlsx")
spInfoSave<-spInfox
spInfox<-as.data.frame(spInfox)
dim(spInfox)
names(spInfox)

2. import spDatax  excel file
spDatax <- read_excel("~/Desktop/BUA 633/Class 3 June 1/spDatax.xlsx")
spDataSave<-spDatax
data.class(spDatax)
spDatax<-as.data.frame(spDatax)
dim(spDatax)
names(spDatax)

3. Merge 2 dataframe by "tkr" overlap
spmdf<-merge(spInfox,spDatax,by="tkr")
names(spmdf)
dim(spmdf)

4. Create new numeric variable date extracted from date
spmdf$date
spmdf$year<-as.numeric(substring(spmdf$date,7,10))
names(spmdf)

5. Cross Section Data Analysis
unique(spmdf$sector)
unique(spmdf$industry)
data.frame(table(spmdf$sector))

6. extract cross section data for year 2022
csdf<-spmdf[spmdf$year==2022 & spmdf$industry=="Health Care Providers &
Services",c("tkr","price","eps","bvps","year","name","industry")]
dim(csdf)
names(csdf)
csdf
csdf$price

7. Graphical
```

```
par(mfcol=c(3,3))  # par - partion
par(mfrow=c(3,3))  # par - partion

hist(csdf$price,xlab = "Price",ylab="Freq",main="Fig. 1 Hist of Price",col="black")
hist(csdf$eps,xlab = "EPS",ylab="Freq",main="Fig. 2 Hist of EPS", col="black")
hist(csdf$bvps,xlab = "BVPS",ylab="Freq",main="Fig. 3 Hist of BVPS", col="black")

8. Scatterplots
par(mfrow=c(2,2))
scatter.smooth(csdf$bvps,csdf$price,xlab="BVPS",ylab="Price",main="Fig. 4 Price v BVPS")
scatter.smooth(csdf$eps,csdf$price,xlab="EPS",ylab="Price",main="Fig. 5 Price v EPS",
    pch="*",cex=2)

9. Analytical
# des stats
ds.summ(csdf[,c("price","eps","bvps")],2)
ds.summ(csdf[,c("price","eps","bvps")],2)[,-c(7,8)]

# cor
round(cor(csdf[,c("price","eps","bvps")]),4)

#lm
dim(csdf)
csdf<-na.omit(csdf)
dim(csdf)
fit<-lm(price~eps+bvps,data=csdf,na.action=na.omit)
summary(fit)
names(fit)
coefficients(fit)
names(summary(fit))
predValues<-predict(fit,csdf)
residValues<-residuals(fit)

# Model Validation
vdf<-data.frame(tsdf,predValues,residValues)
par(mfrow=c(2,2))
 hist(vdf$residValues)
 plot(vdf$predValues,vdf$price,type="n")
 text(vdf$predValues,vdf$price,vdf$tkr)
 scatter.smooth(vdf$predValues,vdf$price,type="n")
 text(vdf$predValues,vdf$price,vdf$tkr)
 pl.2ts(vdf$price,vdf$predValues,"TSPlot Act v Pred")
```