

A Statistical Model to Predict Equity Price in the S&P Semiconductors & Semiconductor Equipment Industry

Amritpal Singh

St. John's University, NY, USA

A Statistical Model to Predict Equity Price in the S&P Semiconductors & Semiconductor Equipment Industry

I. Introduction

Semiconductors and Semiconductor Equipment is one of the six industries in the Information Technology Sector of the S&P1500. This research develops a statistical model to predict the equity price of stocks in this industry. Financial advisors, investors and portfolio managers would be interested in the conclusions of this research.

II. Previous Research

There has been a lot of previous research, but none right on point with this.

III. Methodology

This research will analyze cross-section data with 50 observations obtained from FactSet. The research incorporates Graphical Techniques including histograms and scatterplots. It also uses Analytical Methods, namely descriptive statistics, correlation and regression. R is used to create and execute the statistical program.

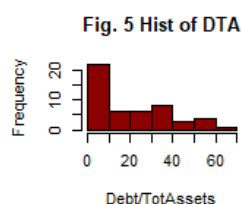
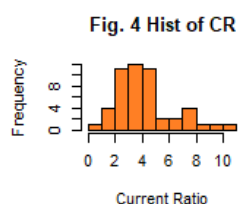
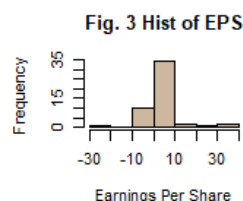
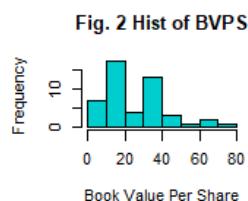
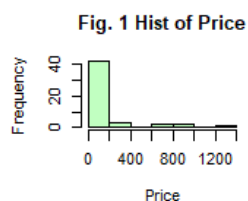
+ + + +

$$\text{Eqn. 1} \quad \text{price} = f(\text{bvps}, \text{eps}, \text{cr}, \text{dta})$$

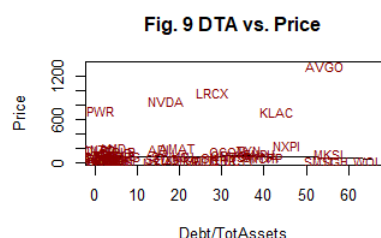
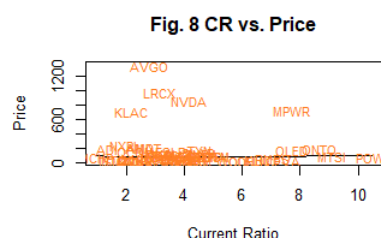
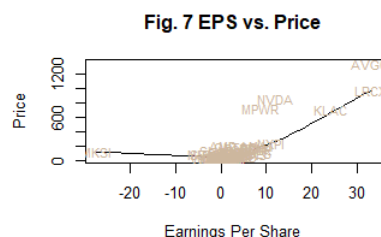
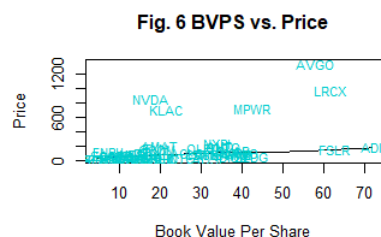
$$\text{Eqn. 2} \quad \text{price} = \alpha + \beta_{\text{bvps}}\text{bvps} + \beta_{\text{eps}}\text{eps} + \beta_{\text{cr}}\text{cr} + \beta_{\text{dta}}\text{dta} + \epsilon$$

$$\text{Eqn. 3} \quad \text{price} = a + b_{\text{bvps}}\text{bvps} + b_{\text{eps}}\text{eps} + b_{\text{cr}}\text{cr} + b_{\text{dta}}\text{dta} + e$$

IV. Results



a. Histograms of all variables are displayed below in Figs. 1 to 5. All histograms are skewed to the right except Fig. 3 which is somewhat symmetric. No outliers are present, but there are 2 null values in BVPS, whose data will be omitted for regression.



- a. Scatterplots are in Figs. 6 to 9. BVPS and EPS have a strong positive correlation, BVPS is linear while EPS is a B-Spline. CR and DTA have weak linear correlations. Notice that most scatterplots contain minor outliers in y-space.

- b. Table 1 displays the descriptive statistics.

Table 1 Descriptive Statistics

| | <i>obs</i> | <i>mean</i> | <i>median</i> | <i>std</i> | <i>skew</i> | <i>kurt</i> |
|--------------|------------|-------------|---------------|------------|-------------|-------------|
| <i>price</i> | 48 | 172.66 | 90.12 | 267.90 | 2.78 | 10.77 |
| <i>bvps</i> | 48 | 25.74 | 20.24 | 16.12 | 0.94 | 3.42 |
| <i>eps</i> | 48 | 3.53 | 1.64 | 8.81 | 0.87 | 9.24 |
| <i>cr</i> | 48 | 4.13 | 3.74 | 2.11 | 1.07 | 3.77 |
| <i>dta</i> | 48 | 20.24 | 16.42 | 18.50 | 0.67 | 2.36 |

- c. Table 2 displays the correlation matrix. EPS has a high correlation with price, BVPS has moderate correlation, while the rest have low correlation. The correlations agree with the initial hypothesis and there is no evidence of multicollinearity.

Table 2 Correlation Matrix

| | <i>price</i> | <i>bvps</i> | <i>eps</i> | <i>cr</i> | <i>dta</i> |
|--------------|--------------|-------------|------------|-----------|------------|
| <i>price</i> | 1.000 | 0.453 | 0.774 | 0.047 | 0.202 |
| <i>bvps</i> | 0.453 | 1.000 | 0.329 | (-0.110) | (-0.041) |
| <i>eps</i> | 0.774 | 0.329 | 1.000 | (-0.136) | 0.081 |
| <i>cr</i> | 0.047 | (-0.110) | (-0.136) | 1.000 | (-0.361) |
| <i>dta</i> | 0.202 | (-0.041) | 0.081 | (-0.361) | 1.000 |

- d. **Table 3 Linear Regression Results**

| | | | | | |
|-----------------|---|-----------|-----------|---------|----------|
| <i>Eqn. 4</i> | price=-150.22 +4.19bvps +21.14eps +19.55cr +3.08dta | | | | |
| <i>t-stat</i> | (-1.83)** | (2.73)*** | (7.63)*** | (1.62)* | (2.28)** |
| <i>p-value</i> | (.07) | (.01) | (.00) | (.11) | (.03) |
| <i>r (corr)</i> | | .45 | .77 | -.05 | .20 |

| Significance Legend |
|---------------------|
| * 10% level |
| ** 5% level |
| *** 1% level |

n = 48 $r^2 = .687$ F=23.54** F. Prob=.00 SE = 159.2

1. $H_0: \beta_{bvp} = \beta_{eps} = \beta_{cr} = \beta_{dta} = 0$
 H_a : at least 1 β_i not equal to 0 ($23.54 > 4.99$, H_0 is rejected)
2. R-squared- 68.7% variation in equity price can be explained by variation in the independent variables (bvps, eps, cr and tda)
3. Standard Error - There is a \$159.20 standard deviation for the residuals around the regression line
4. As bvps increases by \$1, price increases by \$4.19, on average, other things equal
 As eps increases by \$1, price increases by \$21.14, on average, other things equal
 As cr increases by 1 unit, price increases by \$19.55, on average, other things equal
 As dta increases by 1 unit, price increases by \$3.08, on average, other things equal

Fig. 10 Hist of LM Residuals

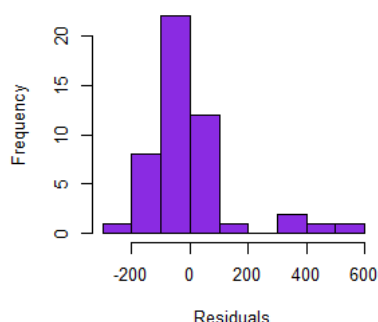
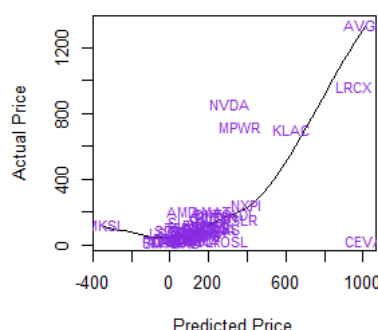


Fig. 11 Predicted vs Actual Price LM



Figs. 10 and 11 showcase the residuals and the predicted vs. actual values of the linear model. The residuals are moderately skewed to the right. The smooth scatterplot is positive and nonlinear with no evidence of heteroscedasticity.

e. Table 4 Robust Regression Results

| Eqn. 4 | price=-15.16 | +1.61bvps | +10.90eps | +7.48cr | +0.55dta |
|-----------------|--------------|-----------|-----------|---------|----------|
| <i>t-stat</i> | (-0.36) | (2.01)** | (2.84)*** | (1.30)* | (0.78) |
| <i>p-value</i> | (.72) | (.05) | (.01) | (.20) | (.44) |
| <i>r (corr)</i> | | .45 | .77 | -.05 | .20 |

| Significance Legend |
|---------------------|
| * 10% level |
| ** 5% level |
| *** 1% level |

$$r^2 = .300 \quad SE = 48.81$$

Fig. 12 Hist of Rob Residuals

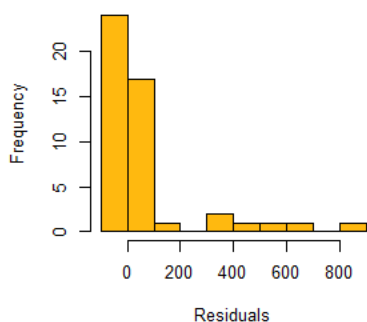
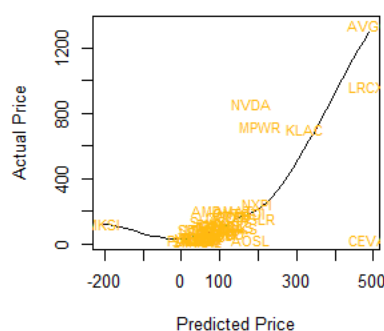
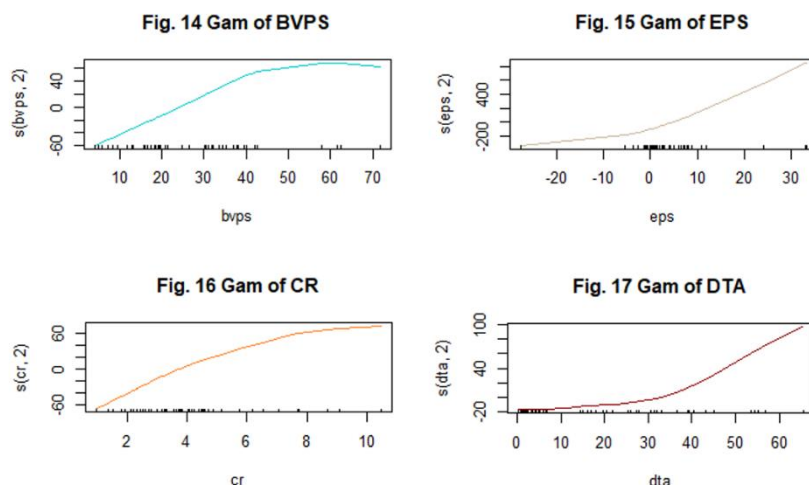


Fig. 13 Predicted vs Actual Price Rob



Compared to linear regression, robust regression had a lower explanatory power of 30%. Only bvps, eps and cr were predictive, while all independent variables were for linear regression. The residuals are much more heavily skewed to the right.

Figs. 14-17 GAM Regression



$$r^2 = .817$$

V. Conclusions

The research was extremely successful. The explanatory power using the linear model was high and all assumptions were not satisfied. All chosen independent variables were significant/predictive (bvps, eps, cr and tda). This research could be further improved by including other independent variables such as sales per share and cash flow per share.

VI. Bibliography

FactSet Data

VII. Appendix I

| <i>tkr</i> | <i>name</i> | <i>price</i> | <i>bvps</i> | <i>eps</i> | <i>cr</i> | <i>dta</i> | <i>sector</i> | <i>industry</i> |
|------------|--------------------------------|--------------|-------------|------------|-----------|------------|------------------------|--|
| NVDA | NVIDIA Corporation | 859.64 | 17.44 | 11.93 | 4.17 | 16.82 | Information Technology | Semiconductors & Semiconductor Equipment |
| AVGO | Broadcom Inc. | 1342.75 | 57.94 | 32.98 | 2.82 | 54.42 | Information Technology | Semiconductors & Semiconductor Equipment |
| AMD | Advanced Micro Devices, Inc. | 205.13 | 34.59 | 0.53 | 2.51 | 4.42 | Information Technology | Semiconductors & Semiconductor Equipment |
| INTC | Intel Corporation | 43.16 | 24.97 | 0.40 | 1.54 | 25.94 | Information Technology | Semiconductors & Semiconductor Equipment |
| QCOM | QUALCOMM Incorporated | 161.45 | 19.37 | 6.42 | 2.33 | 31.48 | Information Technology | Semiconductors & Semiconductor Equipment |
| AMAT | Applied Materials, Inc. | 207.39 | 19.63 | 8.11 | 2.60 | 19.52 | Information Technology | Semiconductors & Semiconductor Equipment |
| TXN | Texas Instruments Incorporated | 170.64 | 18.59 | 7.07 | 4.55 | 36.44 | Information Technology | Semiconductors & Semiconductor Equipment |
| LRCX | Lam Research Corporation | 963.1 | 61.59 | 33.21 | 3.16 | 27.95 | Information Technology | Semiconductors & Semiconductor Equipment |
| MU | Micron Technology, Inc. | 94.47 | 40.18 | -5.34 | 4.46 | 21.79 | Information Technology | Semiconductors & Semiconductor Equipment |

| | | | | | | | | |
|------|---|--------|-------|-------|-------|-------|------------------------|--|
| KLAC | KLA Corporation | 705.87 | 21.35 | 24.15 | 2.20 | 43.09 | Information Technology | Semiconductors & Semiconductor Equipment |
| ADI | Analog Devices, Inc. | 189.8 | 71.67 | 6.55 | 1.37 | 15.11 | Information Technology | Semiconductors & Semiconductor Equipment |
| NXPI | NXP Semiconductors NV | 248.46 | 33.61 | 10.83 | 1.91 | 45.19 | Information Technology | Semiconductors & Semiconductor Equipment |
| MCHP | Microchip Technology Incorporated | 85.64 | 11.94 | 4.02 | 0.98 | 40.32 | Information Technology | Semiconductors & Semiconductor Equipment |
| MPWR | Monolithic Power Systems, Inc. | 717.64 | 42.68 | 8.76 | 7.74 | 0.32 | Information Technology | Semiconductors & Semiconductor Equipment |
| ON | ON Semiconductor Corporation | 76.99 | 18.25 | 4.89 | 2.71 | 27.42 | Information Technology | Semiconductors & Semiconductor Equipment |
| ENPH | Enphase Energy, Inc. | 125.09 | 7.25 | 3.08 | 4.59 | 38.95 | Information Technology | Semiconductors & Semiconductor Equipment |
| SWKS | Skyworks Solutions, Inc. | 100.33 | 38.14 | 6.13 | 3.33 | 17.91 | Information Technology | Semiconductors & Semiconductor Equipment |
| FSLR | First Solar, Inc. | 159.06 | 62.59 | 7.74 | 3.55 | 6.02 | Information Technology | Semiconductors & Semiconductor Equipment |
| TER | Teradyne, Inc. | 103.76 | 16.54 | 2.73 | 3.28 | 2.37 | Information Technology | Semiconductors & Semiconductor Equipment |
| QRVO | Qorvo, Inc. | 113.29 | 39.56 | 1.00 | 3.65 | 31.93 | Information Technology | Semiconductors & Semiconductor Equipment |
| LSCC | Lattice Semiconductor Corporation | 77.29 | 5.04 | 1.85 | 3.78 | 1.94 | Information Technology | Semiconductors & Semiconductor Equipment |
| ONTO | Onto Innovation, Inc. | 187.83 | 35.38 | 2.46 | 8.69 | 1.02 | Information Technology | Semiconductors & Semiconductor Equipment |
| MKSI | MKS Instruments, Inc. | 124.18 | 36.95 | 27.56 | 3.18 | 55.10 | Information Technology | Semiconductors & Semiconductor Equipment |
| AMKR | Amkor Technology, Inc. | 32.79 | 16.11 | 1.46 | 2.31 | 20.45 | Information Technology | Semiconductors & Semiconductor Equipment |
| OLED | Universal Display Corporation | 169.92 | 30.55 | 4.24 | 7.72 | 1.57 | Information Technology | Semiconductors & Semiconductor Equipment |
| RMBS | Rambus Inc. | 61.4 | 9.63 | 3.01 | 7.08 | 2.44 | Information Technology | Semiconductors & Semiconductor Equipment |
| MTSI | MACOM Technology Solutions Holdings, Inc. | 90.9 | 13.35 | 1.28 | 9.11 | 33.00 | Information Technology | Semiconductors & Semiconductor Equipment |
| ALGM | Allegro MicroSystems, Inc. | 29.42 | 5.04 | 0.97 | 4.03 | 3.57 | Information Technology | Semiconductors & Semiconductor Equipment |
| CRUS | Cirrus Logic, Inc. | 89.34 | 30.10 | 3.09 | 4.70 | 6.84 | Information Technology | Semiconductors & Semiconductor Equipment |
| SLAB | Silicon Laboratories Inc. | 140 | 37.89 | -1.09 | 4.51 | 4.99 | Information Technology | Semiconductors & Semiconductor Equipment |
| POWI | Power Integrations, Inc. | 71.54 | 13.26 | 0.97 | 10.47 | 1.22 | Information Technology | Semiconductors & Semiconductor Equipment |
| SYNA | Synaptics Incorporated | 102.23 | 32.24 | 1.83 | 4.89 | 39.42 | Information Technology | Semiconductors & Semiconductor Equipment |
| SEDG | SolarEdge Technologies, Inc. | 67.61 | 42.22 | 0.60 | 3.70 | 16.03 | Information Technology | Semiconductors & Semiconductor Equipment |
| ACLS | Axcelis Technologies, Inc. | 111.41 | 26.46 | 7.43 | 3.79 | 5.92 | Information Technology | Semiconductors & Semiconductor Equipment |
| FORM | FormFactor, Inc. | 44.14 | 11.75 | 1.05 | 4.38 | 4.35 | Information Technology | Semiconductors & Semiconductor Equipment |
| WOLF | Wolfspeed Inc | 26.46 | 13.00 | -2.65 | 5.74 | 65.43 | Information Technology | Semiconductors & Semiconductor Equipment |
| DIOD | Diodes Incorporated | 66.3 | 37.89 | 4.91 | 3.02 | 4.15 | Information Technology | Semiconductors & Semiconductor Equipment |
| KLIC | Kulicke & Soffa Industries, Inc. | 48.68 | 20.86 | 0.99 | 6.55 | 3.23 | Information Technology | Semiconductors & Semiconductor Equipment |
| SITM | SiTime Corporation | 91.19 | 31.20 | -3.63 | 5.13 | 0.84 | Information Technology | Semiconductors & Semiconductor Equipment |
| VECO | Veeco Instruments Inc. | 35.99 | 11.93 | -0.56 | 3.24 | 25.26 | Information Technology | Semiconductors & Semiconductor Equipment |
| UCTT | Ultra Clean Holdings, Inc. | 45.36 | | -0.70 | 2.88 | 34.26 | Information Technology | Semiconductors & Semiconductor Equipment |
| PLAB | Photronics, Inc. | 28.46 | 15.90 | 2.03 | 4.24 | 2.01 | Information Technology | Semiconductors & Semiconductor Equipment |
| MXL | MaxLinear, Inc. | 20.03 | 8.39 | -0.91 | 2.20 | 14.55 | Information Technology | Semiconductors & Semiconductor Equipment |
| COHU | Cohu, Inc. | 31.37 | 19.22 | 0.59 | 6.18 | 5.12 | Information Technology | Semiconductors & Semiconductor Equipment |

| | | | | | | | | |
|-------------|---------------------------------------|-------|-------|-------|------|-------|------------------------|--|
| <i>SMTC</i> | Semtech Corporation | 21.82 | 11.83 | 0.96 | 1.82 | 53.42 | Information Technology | Semiconductors & Semiconductor Equipment |
| <i>PDFS</i> | PDF Solutions, Inc. | 32.89 | 5.98 | 0.08 | 3.85 | 2.13 | Information Technology | Semiconductors & Semiconductor Equipment |
| <i>ICHR</i> | Ichor Holdings, Ltd. | 42.94 | 19.18 | -1.47 | 4.06 | 30.51 | Information Technology | Semiconductors & Semiconductor Equipment |
| <i>SGH</i> | SMART Global Holdings, Inc. | 23.14 | 4.29 | 0.00 | 2.13 | 56.90 | Information Technology | Semiconductors & Semiconductor Equipment |
| <i>AOSL</i> | Alpha and Omega Semiconductor Limited | 22.29 | 31.96 | 0.42 | 2.46 | 6.57 | Information Technology | Semiconductors & Semiconductor Equipment |
| <i>CEVA</i> | CEVA, Inc | 22.48 | | -0.51 | 7.43 | 2.12 | Information Technology | Semiconductors & Semiconductor Equipment |

VII. Appendix II

```

library("YRmisc")

#import data
library(readxl)

sp1500 <- read_excel("C:/Users/Amritpal/Desktop/BUA 633/Data/sp1500.xlsx",
                    sheet = "allToRStudio")

View(sp1500)
dim(sp1500)
names(sp1500)

#Analyze a Sector
idf<-sp1500[sp1500$industry=="Semiconductors & Semiconductor Equipment",]
idf<-as.data.frame(idf)
names(idf)
dim(idf)

#Histograms
par(mfrow=c(3,3))

hist(idf$price,xlab="Price",ylab="Frequency",main="Fig. 1 Hist of Price", col="darkseagreen1")
hist(idf$bvps,xlab="Book Value Per Share",ylab="Frequency",main="Fig. 2 Hist of BVPS", col="cyan3")
hist(idf$eps,xlab="Earnings Per Share",ylab="Frequency",main="Fig. 3 Hist of EPS", col="bisque3")
hist(idf$cr,xlab="Current Ratio",ylab="Frequency",main="Fig. 4 Hist of CR", col="chocolate1")
hist(idf$dta,xlab="Debt/TotAssets",ylab="Frequency",main="Fig. 5 Hist of DTA", col="darkred")

#Scatterplots
par(mfrow=c(2,2))

scatter.smooth(idf$bvps,idf$price,xlab="Book Value Per Share",ylab="Price",main="Fig. 6 BVPS vs. Price",type="n")
text(idf$bvps,idf$price,as.character(idf$tkr),cex=.8, col="cyan3")

scatter.smooth(idf$eps,idf$price,xlab="Earnings Per Share",ylab="Price",main="Fig. 7 EPS vs. Price",type="n")
text(idf$eps,idf$price,as.character(idf$tkr),cex=.8, col="bisque3") #overlay, cex- character expansion

scatter.smooth(idf$cr,idf$price,xlab="Current Ratio",ylab="Price",main="Fig. 8 CR vs. Price",type="n")

```

```

text(idf$cr,idf$price,as.character(idf$tkr),cex=.8, col="chocolate1")
scatter.smooth(idf$dta,idf$price,xlab="Debt/TotAssets",ylab="Price",main="Fig. 9 DTA vs. Price",type="n")
text(idf$dta,idf$price,as.character(idf$tkr),cex=.8, col="darkred")

#Descriptive Statistics
ds.summ(idf[,c("price","bvps","eps","cr","dta")],3)[-c(3,4,7,8)]

#Correlation Matrix
ds.corm(idf[,c("price","bvps","eps","cr","dta")],3)

#Regression - Linear Model (Parametric)
idf1<-na.omit(idf)
names(idf1)
dim(idf1)
fit<-lm(price~bvps+eps+cr+dta,na.action=na.omit,data=idf1)
summary(fit)

#Residual and Prediction Analysis
par(mfrow=c(2,2))
hist(fit$residuals, main = "Fig. 10 Hist of LM Residuals", xlab="Residuals", ylab="Frequency", col="blueviolet")
scatter.smooth(fit$fitted.values,idf1$price, main="Fig. 11 Predicted vs Actual Price LM", xlab="Predicted Price", ylab="Actual Price", type="n")
text(fit$fitted.values,idf$price,as.character(idf$tkr),cex=.8, col="blueviolet")

#Regression - robust (linear parametric with outlier mitigation)
library("robust")
fit<-lmRob(price~bvps+eps+cr+dta,na.action=na.omit,data=idf1)
summary(fit)
par(mfrow=c(2,2))
hist(fit$residuals, main = "Fig. 12 Hist of Rob Residuals", xlab="Residuals", ylab="Frequency", col="darkgoldenrod1")
scatter.smooth(fit$fitted.values,idf1$price, main="Fig. 13 Predicted vs Actual Price Rob", xlab="Predicted Price", ylab="Actual Price", type="n")
text(fit$fitted.values,idf$price,as.character(idf$tkr),cex=.8, col="darkgoldenrod1")

#Regression - General Additive Model (nonlinear nonparametric)
library("gam")
fit<-gam(price~s(bvps,2)+s(eps,2)+s(cr,2)+s(dta,2),na.action=na.omit,data=idf1)
plot(fit)
par(mfrow=c(2,2))
plot(fit, main="Fig. 14 Gam of BVPS", col="cyan3") #repeat for all 4 and combine into a final image
cor(idf1$price, fit$fitted.values)^2

```