Shalini Chhetri (X03742863), BUA 633

Statistical Model to Predict Equity Price of JPMorgan Chase & Co.

Abstract

This project studies the impact of financial indicators on equity prices of JPMorgan Chase & Co. Using a multivariate regression model, this project finds that BVPS has a significant impact on the equity prices. The project also suggests ways to further strengthen and improve the results.

I. Introduction

This research develops an econometric model to predict the equity price for JPMorgan Chase & Co. using yearly financial data from 2000-2022. This research is relevant for investors, bankers, portfolio managers, owners and stakeholders to get a clear picture of the determinants of equity prices of a leading financial institution.

II. Previous Research

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

III. Methodology

This research is composed of secondary time-series data for JPMorgan Chase & Co. 23 annual observations will be studied. Graphical techniques such as histograms, time-series plots, scatterplots and a statistical analysis composed of descriptive statistics for scalable variables, correlation and regression will be used. The research is developed using R.

$$\text{Eqn 1: Price} = f(\overset{+}{EPS} + \overset{+}{BVPS} + \overset{+}{OBS}) \quad \text{(Functional Specification)}$$

$$\text{Eqn 2: Price} = \alpha + \beta_{eps} * EPS + \beta_{bvps} * BVPS + \beta_{obs} * OBS \quad \text{(Population Regression Equation)}$$

$$\text{Eqn 3: Price} = a + b_{eps} * EPS + b_{bvps} * BVPS + b_{obs} * OBS \quad \text{(Sample Regression Equation)}$$

As Earnings per share increase, profits for stockholders increase, thus leading to a price increase. As Book value per share increases, the profitability ratio goes up, hence increasing the price.

IV. Results

Figure 1, 2 and 3 displays the histogram of Price, Earnings Per Share and Book Value Per Share respectively.

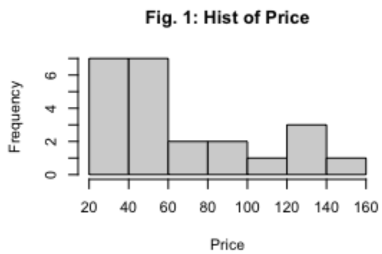


Fig. 1: Hist of Price

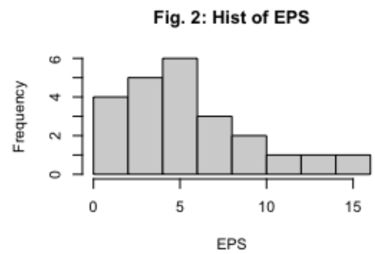


Fig. 2: Hist of EPS

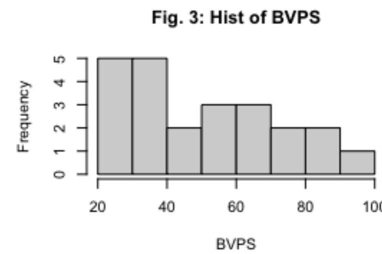


Fig. 3: Hist of BVPS

The histograms for Price and EPS are positively skewed. The histogram for BVPS is also positively skewed albeit to a lesser degree.

Figure 4, 5 and 6 displays the time series plot of Price, Earnings Per Share and Book Value Per Share respectively.

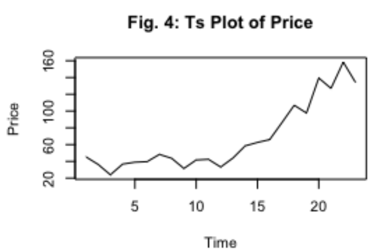


Fig. 4: Ts Plot of Price

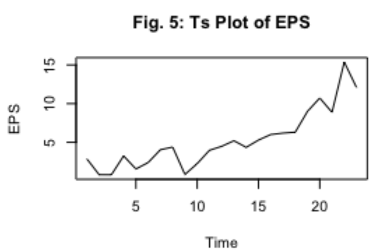


Fig. 5: Ts Plot of EPS

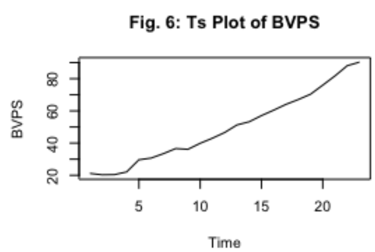


Fig. 6: Ts Plot of BVPS

In Fig 4 and 5, Price and EPS, respectively, are sightly volatile and trending upwards over time. In Fig 8, BVPS is stable and trending upwards over time.

Figure 7,8 and 9 displays scatterplot of Price, Earnings Per Share and Book Value Per Share respectively.

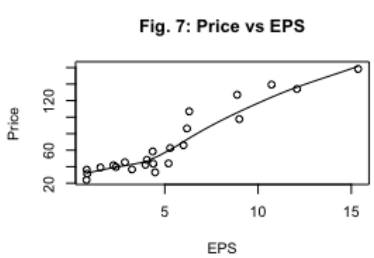


Fig. 7: Price vs EPS

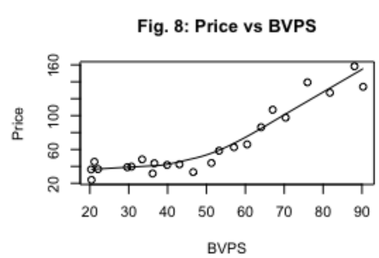


Fig. 8: Price vs BVPS

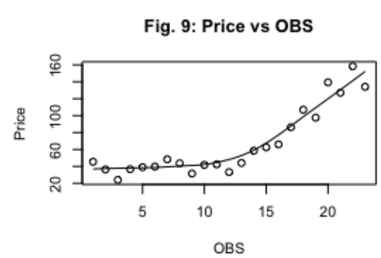


Fig. 9: Price vs OBS

From the scatterplots, one can infer that the relationship between EPS and price is strong and positive. Similar to EPS, BVPS also has a strong and positive relationship with price. The line of best fit is flat between OBS and price for earlier observations but strong and positive for the latter observations.

Table 1: Descriptive Statistics

	n	Mean	Median	Std. Dev	Skewness	Kurtosis
Price	23	67.08	45.44	40.15	0.95	2.64
EPS	23	5.26	4.38	3.78	0.99	3.60
BVPS	23	49.55	46.59	22.14	0.31	1.97
OBS	23	12.00	12.00	6.78	0.00	1.80

The variables Price and EPS are moderately skewed as displayed by the skewness value (between 0.5-1.5) from Table 1. The distribution of BVPS is fairly symmetrical. The kurtosis scores of Price and BVPS implies that these variables have a flatter distribution for the variables compared to a normal distribution. In the case of EPS, the kurtosis score suggests a sharper peak than normal distribution.

Table 2: Correlation Matrix

	Price	*EPS*	*BVPS*	*OBS*
Price	1.000			
EPS	0.940	1.000		
BVPS	0.913	0.916	1.000	
OBS	0.862	0.872	0.990	1.000

Table 2 shows that EPS, BVPS and OBS have a strong positive correlation with Price. It is consistent with the hypothesis. The correlation between EPS and BVPS is positive and high suggesting that there may be a concern for multicollinearity.

Table 3: Regression Analysis

Sample Regression Equation

Eqn 4: Price = -32.23 + 3.54*(EPS) + 3.17*(BVPS) - 8.64*(OBS)

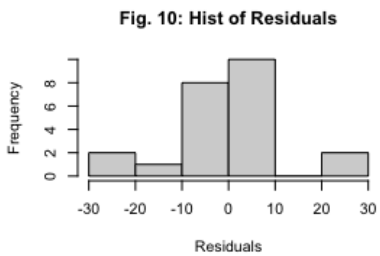
	Intercept	EPS	BVPS	OBS
t Stat	(2.18)**	1.727	(3.089)***	(2.677)**
P-value	0.042	0.100	0.006	0.015
r(corr)	0.000	0.940	0.910	0.860

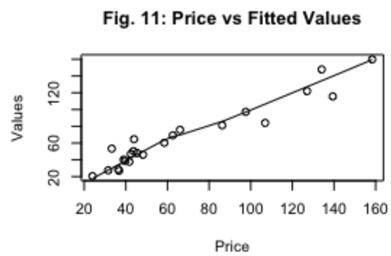
n = 23 r-sq = 0.93 F = 80.88*** SE = 11.64

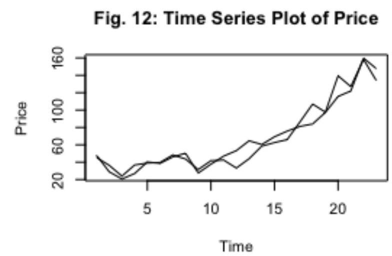
* - signifcant at 10% level
** - signifcant at 5% level
*** - signifcant at 1% level

The F score from Table 3 implies that the results of the regression model are statistically significant. The R-squared suggests that more than 92% of the variation in price is explained by variation in EPS, BVPS and OBS. The result also shows that BVPS and OBS are statistically significant predictors of price. On average, ceteris paribus, a $1 increase in BVPS leads to a $3.17 increase in price.

Figure 10, 11 and 12 displays the histogram of residuals, scatterplot of residuals i.e. Price vs Fitted Values and time series plot of residuals i.e. Time vs Price and Predicted Price.







As shown in Fig. 10, residuals are roughly centered around 0. The points lie in close proximity to the fitted line in Fig. 11, implying that the model predicts the price variable very well. In Fig. 12, the comparison of predicted prices against actual prices also indicates that the model is a good fit.

V. Conclusion

The research project was successful. Almost 92% of the variation in prices are explained by the variation in Earnings per share (EPS), Book value per share (BVPS) and Observations (OBS). The hypothesis for this research was also correct since BVPS has a strong positive correlation with prices. JPMorgan Chase & Co. and financial institutions in general should therefore focus on increasing their BVPS for a higher share price. This research can be improved by adding further explanatory variables and tests for better model selection. More granulated data in terms of time series that explains various economic cycles would help improve the validity of the results. Use of log variables would also improve the model by removing the skewness of the variables.

VI. Bibliography

Excel Book- Dataset from FactSet provided by Professor Manuel Russon.

VII. Appendix I

Tkr	Name	Sector	Date	Price	EPS	BVPS
JPM	JPMorgan Chase & Co	Financials	12/31/2000	45.4375	2.86	21.16578
JPM	JPMorgan Chase & Co	Financials	12/31/2001	36.35	0.81	20.31536
JPM	JPMorgan Chase & Co	Financials	12/31/2002	24	0.8	20.40803
JPM	JPMorgan Chase & Co	Financials	12/31/2003	36.73	3.24	22.10152
JPM	JPMorgan Chase & Co	Financials	12/31/2004	39.01	1.55	29.61427
JPM	JPMorgan Chase & Co	Financials	12/31/2005	39.69	2.38	30.70879
JPM	JPMorgan Chase & Co	Financials	12/31/2006	48.3	4.04	33.44904
JPM	JPMorgan Chase & Co	Financials	12/31/2007	43.65	4.38	36.59251
JPM	JPMorgan Chase & Co	Financials	12/31/2008	31.53	0.8391	36.15114
JPM	JPMorgan Chase & Co	Financials	12/31/2009	41.67	2.2418	39.88195
JPM	JPMorgan Chase & Co	Financials	12/31/2010	42.42	3.9638	43.04178
JPM	JPMorgan Chase & Co	Financials	12/31/2011	33.25	4.4813	46.59089
JPM	JPMorgan Chase & Co	Financials	12/31/2012	43.9691	5.2004	51.26537
JPM	JPMorgan Chase & Co	Financials	12/31/2013	58.48	4.3495	53.25193
JPM	JPMorgan Chase & Co	Financials	12/31/2014	62.58	5.2869	57.06972
JPM	JPMorgan Chase & Co	Financials	12/31/2015	66.03	6.0025	60.46309
JPM	JPMorgan Chase & Co	Financials	12/31/2016	86.29	6.1881	64.0578
JPM	JPMorgan Chase & Co	Financials	12/31/2017	106.94	6.3093	67.03795
JPM	JPMorgan Chase & Co	Financials	12/31/2018	97.62	8.995	70.34904
JPM	JPMorgan Chase & Co	Financials	12/31/2019	139.4	10.7238	75.98423
JPM	JPMorgan Chase & Co	Financials	12/31/2020	127.07	8.878	81.74991
JPM	JPMorgan Chase & Co	Financials	12/31/2021	158.35	15.3648	88.06925
JPM	JPMorgan Chase & Co	Financials	12/31/2022	134.1	12.0848	90.28791

VIII. Appendix II

```
## BUA 633 Term paper - Time Series (JPMorgan) ##
## Shalini Chhetri ##
# Load the necessary packages
library("dplyr")
library("ggplot2")
library("tidyverse")
library("readxl")
library("moments")
library("tidyr")
library("broom")
library("YRmisc")
library("robust")
library("gam")
library("quantmod")
library("robustbase")

# Importing the two datasets
#spData - contains financial data of companies
spdata <- read_excel("spData.xlsx")
#spInfo - contains the sector and industry of companies
spinfo <- read_excel("spInfo.xlsx")

# Check the type/class of data and convert to dataframe
#spdata
data.class(spdata)
spdata<-as.data.frame(spdata)
data.class(spdata)
#spinfo
data.class(spinfo)
spinfo<-as.data.frame(spinfo)
data.class(spinfo)

# Merging the two datasets using common variable "tkr"
names(spdata)
names(spinfo)
spdf<-merge(spdata,spinfo,by="tkr")
dim(spdf)
names(spdf)

# Replace mm/dd/yyyy date format with yyyy
spdf$date
spdf$date<-as.numeric(substring(spdf$date,7,10)) # modify date to only keep year
names(spdf)
names(spdf)[2]<-"year"  # rename "date" to "year"
```

```
names(spdf)

## Select "JPMorgan" - "JPM" as the company for analysis and create a new dataset
unique(spdf$tkr)
JPMorgan<-spdf[spdf$tkr=="JPM",c("tkr","price","eps","bvps","cr","dta",
                    "sales","year","name")]
JPMorgan<-df.sortcol(JPMorgan,"year",FALSE)
JPMorgan$obs<-1:23
names(JPMorgan)

## Histogram of the variables
par(mfrow=c(3,3))
hist(JPMorgan$price, main="Fig. 1: Hist of Price", xlab = "Price")
hist(JPMorgan$eps,main="Fig. 2: Hist of EPS", xlab = "EPS")
hist(JPMorgan$bvps,main="Fig. 3: Hist of BVPS", xlab = "BVPS")

## Generate Timeseries plots
par(mfrow=c(3,3))
ts.plot(JPMorgan$price, main = "Fig. 4: Ts Plot of Price", ylab = "Price", xlab = "Time")
ts.plot(JPMorgan$eps, main = "Fig. 5: Ts Plot of EPS", ylab = "EPS", xlab = "Time")
ts.plot(JPMorgan$bvps, main = "Fig. 6: Ts Plot of BVPS", ylab = "BVPS", xlab = "Time")

## Generate scatterplots of independent variables against the dependent variable
par(mfrow=c(3,3))
scatter.smooth(JPMorgan$eps,JPMorgan$price, main = "Fig. 7: Price vs EPS",
          xlab = "EPS", ylab = "Price")
scatter.smooth(JPMorgan$bvps,JPMorgan$price, main = "Fig. 8: Price vs BVPS",
          xlab = "BVPS", ylab = "Price")
scatter.smooth(JPMorgan$obs,JPMorgan$price, main = "Fig. 9: Price vs OBS",
          xlab = "OBS", ylab = "Price")

## Summary of data
names(JPMorgan)
JPMorgan<-na.omit(JPMorgan)
ds.summ(JPMorgan[,c("price","eps","bvps","obs")],2)
# Table of summary statistics
summary(JPMorgan[,c("price","eps","bvps","obs")])
# Table of correlation
round(cor(JPMorgan[,c("price","eps","bvps","obs")]),3)

## Regression Model
#OLS Model
fit<-lm(price ~ eps+bvps+obs,na.action=na.omit,data=JPMorgan); summary(fit)
cor(JPMorgan$price,predict(fit,JPMorgan))^2
summary(fit)
```

```
## Residual Plots
# Histogram
par(mfrow = c(3, 3))
hist(fit$residuals, main="Fig. 10: Hist of Residuals", xlab="Residuals")
# Scatterplot with smooth line
plot(JPMorgan$price, fit$fitted.values, main="Fig. 11: Price vs Fitted Values", xlab="Price",
ylab="Fitted
Values")
lines(lowess(JPMorgan$price, fit$fitted.values))
# Timeseries plot
plot.ts(JPMorgan$price, main="Fig. 12: Time Series Plot of Price", ylab="Price")
lines(fit$fitted.values)
```