DOES PRICE-TO-BOOK RATIO HAVE ANY EFFECT ON LONG-TERM STOCK RETURNS? AN ANALYSIS OF AMERICAN LARGE CAP STOCKS, 1990-2018

Introduction

Since the 2001 "Tech Bubble", the traditional market capitalization weighted indexe has fallen out of favour with many passive investors. The rationale behind this is that during bubbles, the most overpriced stocks will be overrepresented and underpriced stocks underrepresented. Factor-based passive strategies, "Smart Beta", has since been growing in popularity. Factor-based strategies involve creating an index that weighs stocks on factors other than market capitalization. These include variants of "value", momentum and volatility.

This paper seeks to examine value-weighted funds. Value funds often place emphasis on Price-to-Earnings (Price/Earnings), Price-to-Book (Price/Book), dividend yields and cash flow. The construction's aim is to exploit market inefficiencies with respect to these factors. Although the construction of these alternative indices will consist of a mix of these value factors, this paper will examine Price/Book and subsequent returns to verify whether there has historically been any relationship between a low Price/Book and subsequent returns. Specifically, this paper will look at large cap American stocks, specifically constituents of the Standard and Poor's 500 index.

Background

First identified in Benjamin Graham's *The Intelligent Investor* (1949), the Price-to-Book ratio is a metric that remains hugely relevant in equity valuation, especially among value investors and proponents of fundamental analysis. Price/Book is simply:

$$\frac{\text{Market Capitalization}}{\text{Book Value}}$$

A low Price/Book suggests that a company is trading at a price below its intrinsic value but may also incorporate other factors such as negative investor outlooks on future growth or earnings.

Critics of fundamental analysis methods such as Edward Thorp have placed very little weight on such metrics' predictive power, stating that 'many "undervalued" stocks remain bargains for years', suggesting that even if it were a good predictor in valuation, the intrinsic value may never be realized (Thorp and Kassouf, 1967). Value metrics have since been a key part of the debate against Eugene Fama's Efficient-Market Hypothesis (Lakonishok, Shleifer and Vishny, 1994). This debate began with Basu (1977) empirically demonstrating that a security's Price-to-Earnings ratio was a valid indicator for future returns (Basu, 1977). De Bondt and Thaler also found that there was a predictable overreaction with respect to "winner" stocks and that "loser" stocks outperformed market beta(Bondt and Thaler, 1985).

Rosenberg, Reid and Lanstein were the first to demonstrate a statistically significant strategy centred around Price/Book, concluding that there was pricing inefficiency for their examined period (Rosenberg, Reid and Lanstein, 1985). The Fama-French three-factor model, proposed in

1992, is an asset pricing model that factors in market risk, market capitalization and Book-to-Price. They found that a high Book-to-Price ratio (a low Price/Book) tended to outperform but argue that these stocks had a higher risk premium that justifies their higher returns (Fama and French, 1992;Fama and French, 1993).

This paper has examined large cap stocks exclusively from the Standard and Poor's 500 Index between 1990-2018 and looked at the relationship between Price/Book and subsequent 10 year annualised returns. The findings are consistent with Rosenberg et al. (1985) in that there is a statistically significant, but weak, negative correlation between Price/Book and returns: a lower Price/Book predicts an above average return. Equally, Fama and French's findings that low Price/Book stocks have a higher risk premium is consistent with this 30 year period.

Data

Data was taken from Thompson Reuters' Datastream and Eikon platforms. Data on constituents of the S&P 500 were recorded annually at the start of the year, 1990-2008. For each constituent the following values taken were: Name, Price/Earnings, Price/Book[1], Adjusted Price, Unadjusted Price, Stock Ticker, and a Return Index that spanned the succeeding 10 years[2]. The return index displays returns in nominal terms. The return index in year t was standardised to base 100 for cross sectional comparison. The return index rather than adjusted price or unadjusted price was the main metric used as it accounted for both stock splits and dividends (reinvested when paid). Stocks with missing values in either Price-to-Book or Price-to-Earnings were omitted. Stocks which have a 0% change in their Return Index between years were also omitted as this also constituted missing data. Recurring members of the S&P 500 were included multiple times with different base years as this comparison is focussed on the relationship between subsequent returns and Price/Book values at a point in time. In line with other analyses that focus on book equity, companies with a negative book value (1.1% of the sample) were removed[3].

Analysis

The return index was used to create a list of subsequent 10 year annualized return using

$$10\text{ Year Annualized Return} = \left(RI_{t+10}^{\frac{1}{10}}\right) - 1$$

Where RI_{t+10} is the return index at year 10. This was compared to Price/Book at year t. Below is a descriptive summary of the sample:

vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se

[1] Book Equity and Earnings are values as of a company's respective last reporting period

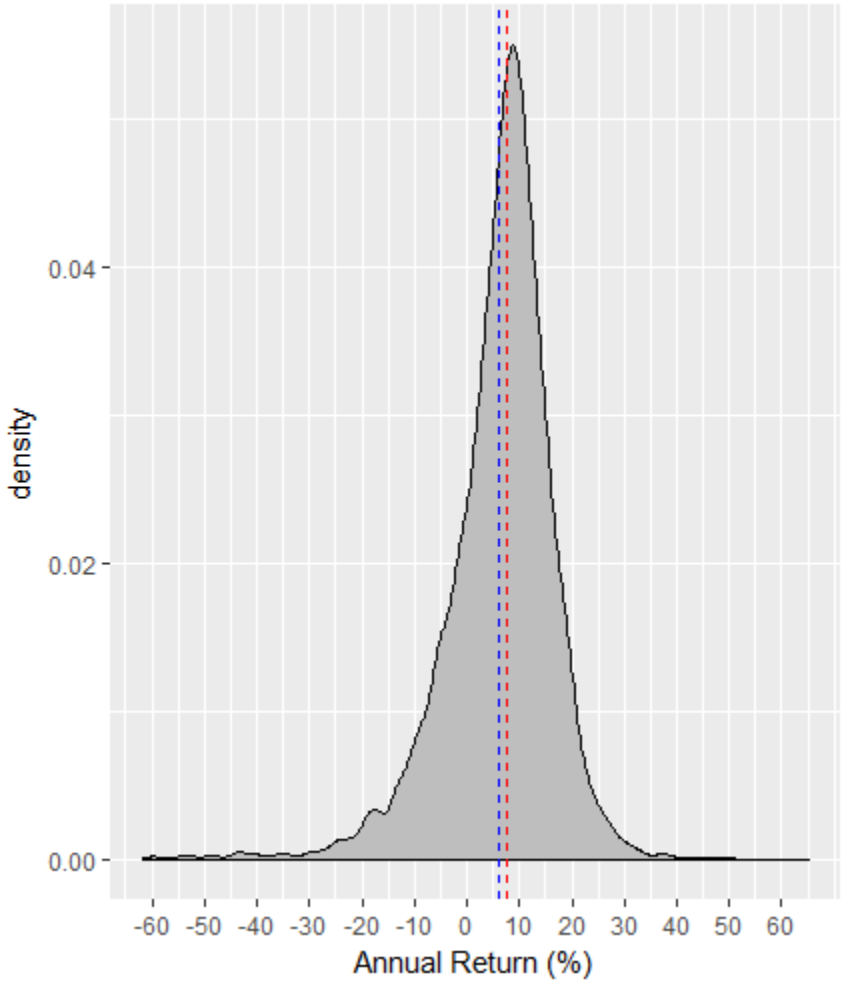
[2] In Datastream these are represented as P for Adjusted Price, UP for Unajusted Price, RI for Return Index, PTBV for Price/Book, WC09104 for Price/Earnings, WC05601 for Stock Ticker and RI for Return Index

[3] For further discussion on why a negative book value is note useful in analysis see Brown, S., Lajbcygier, P. and Li, B. (2008). Going Negative: What to do with Negative Book Equity Stocks. *SSRN Electronic Journal*.

```
PB        1 6368 3.88  6.48   2.61    2.96 1.54   0.22 243.48 243.26 16.07   413.06 0.08
Return    2 6368 6.21 10.58   7.63    6.98 7.86 -61.62  64.98 126.61 -1.24     5.10 0.13
```

The sample period, 1980-2018, had two major crises: the Dot Com Bubble (2001) and the Financial Crisis (2007-8). This is key in understanding the huge standard deviation (10.58) in annualized returns in this period. The median absolute deviation (7.63) is lower than the standard deviation and therefore we would expect a skewed distribution. For testing correlation through regression this suggest that an OLS regression results may be compromised by outlier stocks.

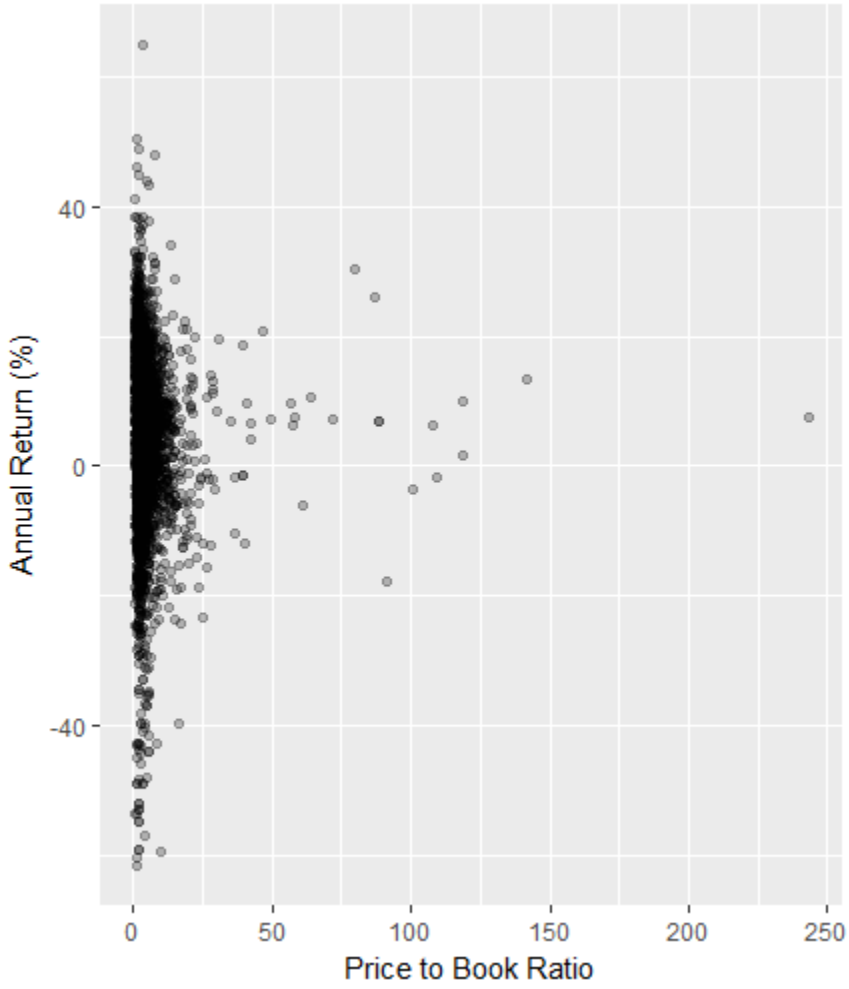
It is important to note that this sample differs from the S&P 500's annualized return for the same period. The sample equally weighs all stocks and has a mean and median return of 6.21% and 7.63% respectively. The S&P 500 is not an equally weighted index (it is weighted on market capitalization) and therefore the results will differ.



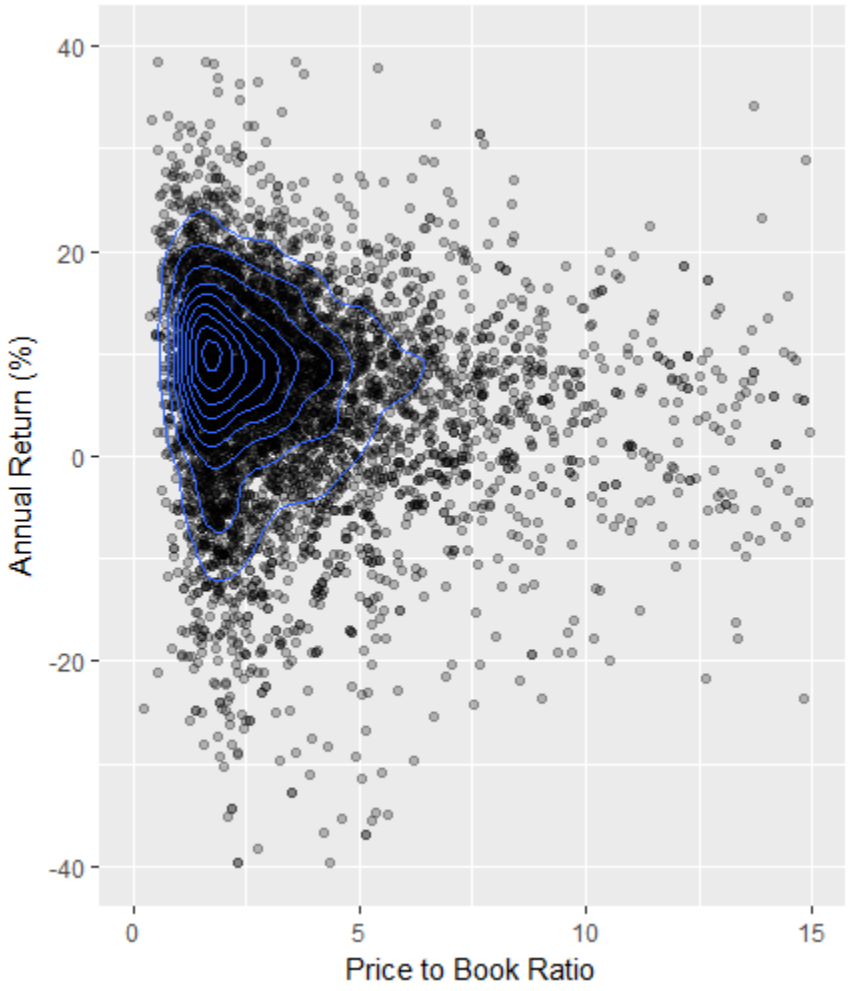
The density plot visually demonstrates the strong left skew (-1.24) of annualized returns. This is likely the effect of the two crises and tech stocks that were overrepresented in the S&P 500 that have never recovered since 2001. The red and blue dashed lines are added as a reference for

mean and median returns respectively. The kurtosis of 5.10 suggests implies a leptokurtic distribution. Beyond the crises, this is also explained by the equity risk premium.

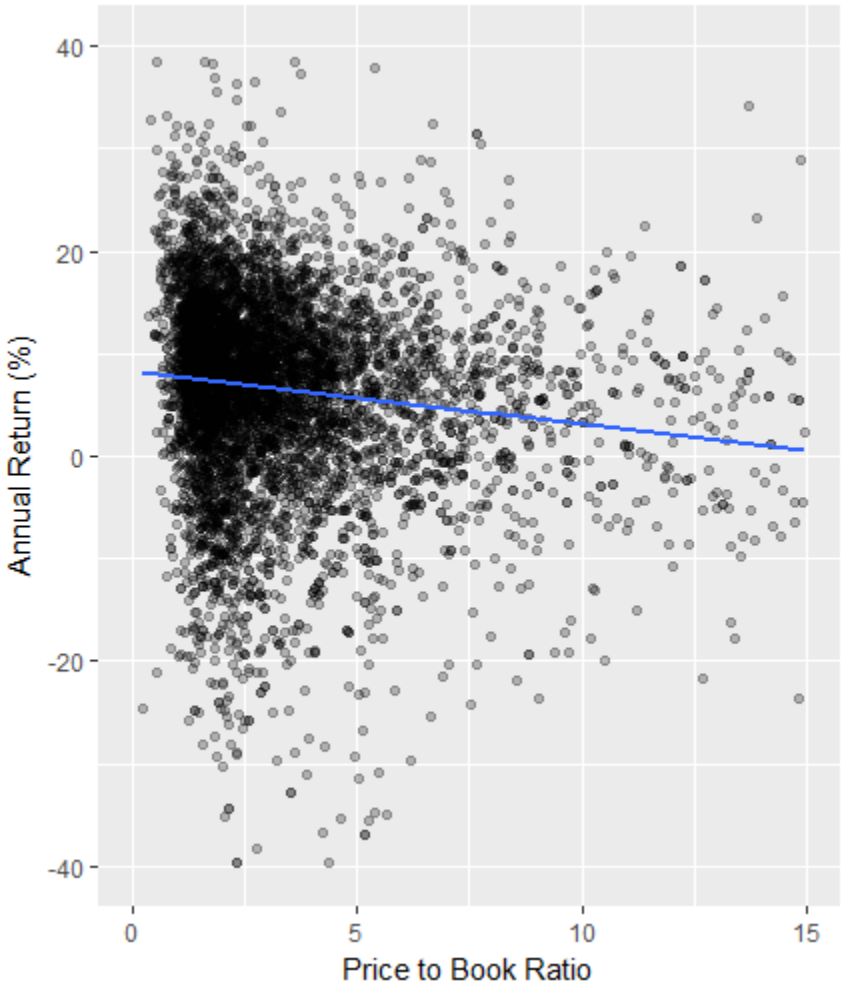
Estimating the relationship



A scatter plot of Price/Book to Annualized Returns of the sample shows no clear visual relationship and clearly has a few outliers. One explanation for such outliers is that a high Price/Book implies investor confidence in a security – the price is much higher than its book value because of future earnings expectation. Thaler and De Bondt's theory that "winner" stocks are pushed far beyond their reasonable valuation is visible here (Bondt and Thaler, 1985).



Limiting the axes gives a slightly clearer visualization of the relationship. In the above scatter plot, density of the different values is shown through contour lines as well as colour, darker dots imply a cluster of values around the area. The correlation between Price/Book and Annual Return is -0.05872881. This is clearly very weak. While undervalued stocks have low Price/Book values, companies that are perceived to have a fundamental weakness also have low values. Price/Book alone can not distinguish between these two groups. A linear regression also implies this presents this weak negative correlation:



The data not in view (stocks with a PB>15) have not been omitted, this scatter plot is for illustrative purposes. The negative slope of the β coefficient implies that the higher the Price/Book value of a stock, the less one would expect it to return over the succeeding 10 years.

```
                    Residuals:
     Min      1Q  Median      3Q     Max
-68.052  -4.701   1.416   6.050  58.706

                    Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.57676    0.15421  42.649  < 2e-16 ***
PB          -0.09586    0.02042  -4.694 2.74e-06 ***
                    ---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 10.56 on 6366 degrees of freedom
Multiple R-squared:  0.003449, Adjusted R-squared:  0.003293
F-statistic: 22.03 on 1 and 6366 DF,  p-value: 2.736e-06
```

The regression takes the form:

$$\text{Annualized 10 Year Returns} = 6.57676 - 0.09586\text{PB} + \varepsilon$$

with a standard error of 10.56. This says that for an increase in one unit of Price/Book, returns will decrease by 9 basis points. However R^2 value of 0.003449 is clearly very low and suggests that the fitted line is a very weak predictor of stock returns when a function of Price/Book.

The t value of -4.694 is very instructive: this suggests that we can reject the null hypothesis of there being no relationship between these two variables with a *very* high level of confidence.

```
Breusch Pagan Test for Heteroskedasticity
-----------------------------------------
      Ho: the variance is constant
    Ha: the variance is not constant

                 Data
-----------------------------------
         Response : Return
  Variables: fitted values of Return

           Test Summary
 -----------------------------
      DF            =    1
   Chi2          =    10.23817
  Prob > Chi2   =    0.001375637
```

The Breusch Pagan Test also shows us that we cannot reject H_0 implying that we cannot rule out heteroskedasticity. This is unsurprising and consistent with expectations due to the presence of outliers in our original sample, as seen in the first scatter plot. 71 of the observation have Price/Book values above 20 with a maximum Price/Book value being 243.48. This maximum value has a Cook's Distance of 0.9249 suggesting it has disproportionate leverage and influence on the regression. This distorts the regression, but the t value remains promising for estimating the relationship. Adjusting for heteroskedasticity, a robust regression may be more useful in estimating the relationship:

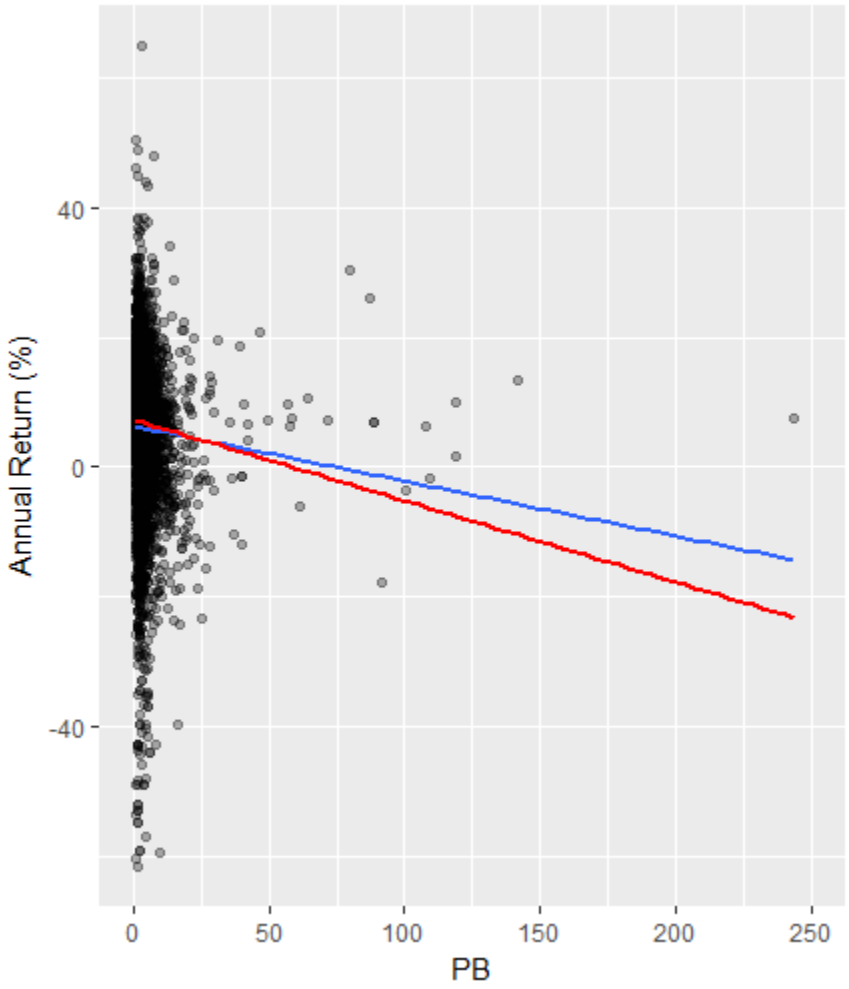
```
                    Residuals:
     Min       1Q   Median       3Q      Max
-69.3539  -5.7075   0.3525   5.0245  57.5968

                    Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  8.06712    0.29272  27.559  < 2e-16 ***
PB          -0.21857    0.07805  -2.801  0.00512 **
                        ---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Robust residual standard error: 7.871
Multiple R-squared:  0.01545,  Adjusted R-squared:  0.0153
```

This takes the form:

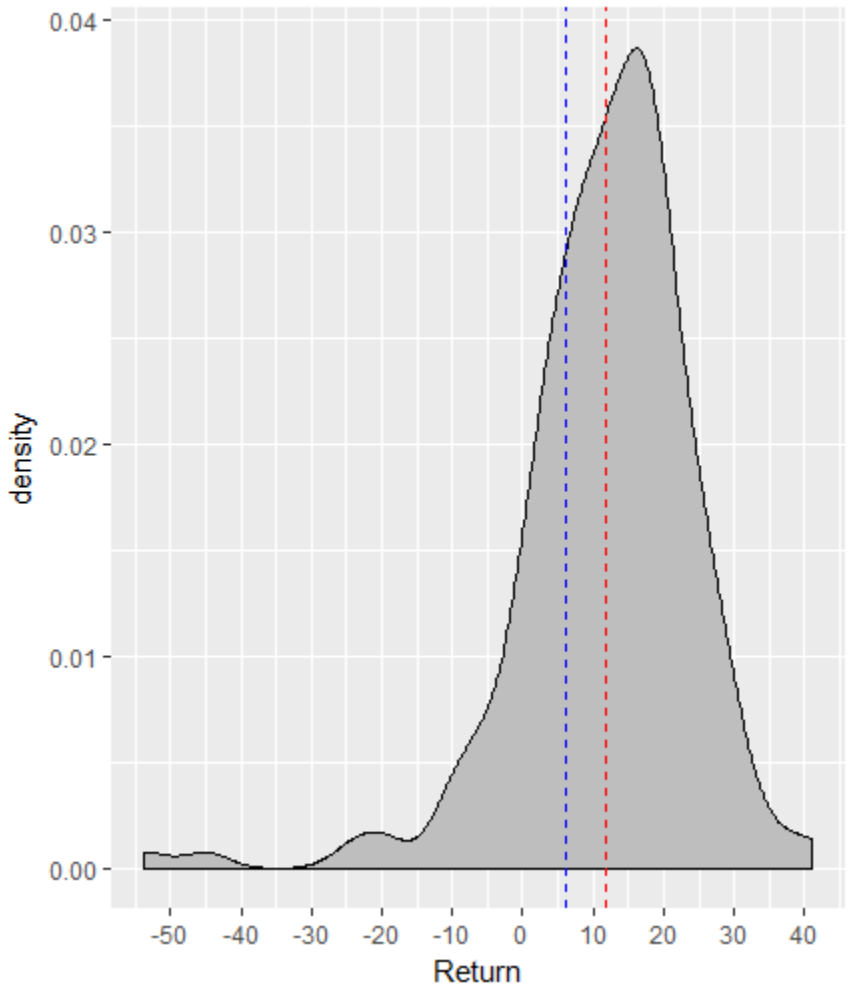
$$\text{Annualized 10 Year Returns} = 8.06712 - 0.21857\text{PB} + \varepsilon$$



Red line represents the Robust Regression line of best fit

This suggests an even more negative relationship due to the β coefficient of -0.21857, in increase of 1 Price/Book implies a drop in annual return by 21 basis points.This has a slightly better fit with an R^2 of 0.01545 but is still very inconsistent with the data. The larger negative coefficient in this regression is consistent with Rosenberg's findings but we should be cautious as the relationship is still weak (Rosenberg, Reid and Lanstein, 1985).

Looking at very small Price/Book values may be more instructive. Price/Book values less than 5 returned 6.7%, slightly higher than the mean of the total sample and represents 82% of total observations. For these small values of Price/Book, the relationship is more defined. For Price/Book values less than 2, the mean return was 7.75%. These stocks still represent 34% of the total sample. A Price/Book value of less than 1 was very rare, only 2% of the sample, but had an impressive mean return of 11.73%.



Above is a density plot of returns for stocks with a Price/Book of less than 2. The Blue dashed line represents the total sample mean return and the Red is the small Price/Book mean.

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
PB	1	2166	1.49	0.33	1.51	1.51	0.36	0.22	1.99	1.77	-0.56	-0.14	0.01
Return	2	2166	7.76	10.59	8.83	8.47	7.60	-61.62	50.69	112.31	-1.37	6.55	0.23

This has similar implications to the Fama-French (1993) findings that lower Price/Book stocks have higher returns on average with higher volatility. The skewness of the distribution of returns (-1.37) confirms this, despite average return being higher. Having a 6.55 kurtosis also implies a fatter tail than the original sample distribution. Standard deviation (10.59) is also slightly higher.

A robust regression for Price/Book values lower than 5 takes the form:

$$\text{Annualized 10 Year Returns} = 9.9132 - 0.9428\text{PB} + \varepsilon$$

Residuals:

Min	1Q	Median	3Q	Max
-70.0846	-5.1196	0.4705	5.0212	58.0031

```
                Coefficients:
      Value   Std. Error t value
(Intercept)  9.9132  0.3019    32.8338
PB          -0.9428  0.1136    -8.3025

Residual standard error: 7.547 on 5207 degrees of freedom
```

With and R^2 value of 0.01481. The R^2 remains low but strengthens the consistent claim of a negative relationship between Price/Book and Returns. The very high t value -8.3025 implies we can make this claim with a very high level of confidence.

Conclusion

As expected Price/Book alone cannot successfully predict stock market returns. Stock returns are a function of far more than this variable alone. As such, all attempts at regressing the returns have produced poor R^2 values but have demonstrated a statistically significant relationship between Price/Book and annual return. Rosenberg's study on Price/Book strategies go as far as rejecting market efficiency. This study is consistent with their findings over a longer period: we can definitively say that a lower Price/Book implies a higher return although concedes that there is a higher risk premium associated with these stocks.

Basu, S. (1977). INVESTMENT PERFORMANCE OF COMMON STOCKS IN RELATION TO THEIR PRICE-EARNINGS RATIOS: A TEST OF THE EFFICIENT MARKET HYPOTHESIS. *The Journal of Finance*, 32(3), pp.663-682.

Brown, S., Lajbcygier, P. and Li, B. (2008). Going Negative: What to do with Negative Book Equity Stocks. *SSRN Electronic Journal*.

Bondt, W. and Thaler, R. (1985). Does the Stock Market Overreact?. *The Journal of Finance*, 40(3), p.793.

Fama, E. and French, K. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), p.427.

Fama, E. and French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), pp.3-56.

Lakonishok, J., Shleifer, A. and Vishny, R. (1994). Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance*, 49(5), p.1541.

Rosenberg, B., Reid, K. and Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, 11(3), pp.9-16.