Forecasting Hewlett Packard Enterprise's Earnings Yield with Random Forests

Nicolas Newberry | nnicolas@umich.edu | +1 734 634 1562

Abstract

This study employs Random Forest to forecast the earnings-to-price (E/P) ratio for Hewlett Packard Enterprise (HPE). Using financial and market indicators from February 2006 to January 2023, the model was trained (2006–2015), validated (2016–2022), and tested (January 2023). These splits were chosen due to their similar variance and characteristics balancing out recessionary events like the 2008 Financial Crisis and the 2020 Covid-19 Pandemic. Random Forests were ideal for this analysis due to their ability to capture complex, non-linear relationships in financial data. Their ensemble approach averaged out errors, ensuring reliability across volatile markets. The Random Forest analysis achieved a validation set R^2 of 0.1259, outperforming a vanilla decision tree's R^2 of 0.1072 by 18%. Random Forests produced an entire dataset R^2 of 0.1470, demonstrating strong performance over the best single decision tree's R^2 value of 0.1290. Beating it by 13.7%, highlighting Random Forests' enhanced generalization abilities, and ability to counteract overfitting.

The Random Forests analysis identified the top five key metrics influencing the earnings-to-price (E/P) ratio for Hewlett Packard Enterprise (HPE): book-to-market ratio (finlag1bm, 44.4% importance), short interest ratio (fing09shoadj, 18.7%), R&D-to-sales ratio (fing05rdsadj, 12.6%), earnings revisions (fing14erevadj, 6.2%), and lagged market capitalization (lag1mcreal, 5.7%). The analysis revealed that companies with higher book-to-market ratios exhibited elevated E/P ratios, indicating that value stocks, perceived as undervalued, typically command higher earnings yields potentially due to slower future growth prospects. Conversely, higher short interest ratios correlated with higher E/P ratios, suggesting that stocks with increased bearish sentiment, often viewed as riskier, offer higher yields to compensate investors. A lower R&D-to-sales ratio was associated with higher E/P ratios, as increased R&D spending may reduce earnings, and firms with low R&D expenditures may be slow-growing businesses. Slightly positive earnings revisions signaled modest analyst optimism, exerting a subtle downward pressure on E/P. Lastly, larger lagged market capitalization, indicative of stable firms, correlated with lower E/P ratios, reflecting higher valuations typical of established companies. For HPE, the model's predicted E/P of 0.0679 (6.79%) in January 2023 was driven by its high book-to-market ratio (1.104), low short interest (0.030), modest R&D-to-sales ratio (0.020), slightly positive earnings revisions (0.001), and large market capitalization (21,664,747). In conclusion, the analysis estimates a fair E/P ratio for HPE at 6.79%, suggesting that investors consider purchasing shares in January 2023 if the earnings yield meets or exceeds this threshold, potentially indicating an undervalued stock.

Contents

Tables

Figures

Introduction

This project leverages Random Forest techniques to predict the earnings-to-price ratio for Hewlett Packard Enterprise (HPE) utilizing a range of financial and market-based indicators. I employed a dataset spanning February 2006 to December 2023, comprising 447,188 monthly stock-level observations across 14 financial variables and 12 “miss” variables acting as a placeholder for missing data. The Random Forest model was trained on data from 2006 to 2015, validated on data from 2016 to 2022, and tested on January 2023, ensuring a time-based split that captures market dynamics across economic cycles. The goal is to accurately predict HPE's earnings to price ratio to recommend a ratio that represents the fair value of HPE stock.

Data

Introduction

This dataset contains a set of financial and market variables, derived from a wide range of corporate and stock market data, with observations from 2006 to 2023. The dataset includes metrics from 12 teams such as dividends, ESG scores, insider buying, free cash flow yield, research and development expenditures, sales growth, short interest, and various financial ratios. Features like "lag1mcreal," the natural log of market capitalization, and "adjlag1bm," the book-to-market ratio, are widely researched financial indicators associated with future stock returns.

The dataset, spanning 2006 to 2023, underwent preprocessing to handle missing data and ensure consistency for time-series analysis. Absent values in financial variables were replaced with zeros, while corresponding “miss” dummy variables were set to 1 to indicate missing data. This approach ensures a complete dataset with only valid values, which is essential for the Random Forest analysis.

Training, Validation, and Testing Splits

To ensure a balanced and meaningful training-validation split for the analysis, I selected December 31, 2015, as the cutoff point. This choice provides a ~59/41 split of the data while maintaining comparable levels of market volatility on both sides. The dataset spans 2006 to 2023, covering significant economic events, including the 2008 financial crisis, the European debt crisis (2011–2012), the COVID-19 market shock (2020), and recent post-pandemic fluctuations. By using 2006 to 2015 for training, the analysis captures a diverse range of volatility regimes, including the extreme spike of 2008 (VIX October 2008 highest close: 59.89). Meanwhile, the 2016 to 2023 validation period reflects more recent market conditions, incorporating both pre-pandemic stability (VIX in 2017: 11.04) and heightened uncertainty post 2018 (VIX March 2020 highest close: 53.54). Overall, the average VIX during the training period was 20.4 and the average during the validation period was 19.1 whereas the long run VIX average (1990 to 2022) is approximately 19.7.

Cutting off the date at the end of 2015 balances volatility, ensuring that both training and validation periods include periods of high and low volatility, crucial for testing robustness in the models. Furthermore, the choice aligns with financial data best practices, where models trained on multiple economic cycles tend to generalize better. The volatility data is synced from CBOE's historical VIX records, a widely used measure of market sentiment and uncertainty (CBOE VIX Index).

To ensure this split was appropriate, I ran a test on the variance between the dependent variable in training and validation sets. As shown in Table 1, earnings-to-price ratio variance is similar between both sets of data at 0.00328 in the training set and 0.00320 in the validation set, a difference of approximately 2.4%. This similarity supports the split by ensuring the statistical properties of the E/P ratio are consistent across both sets.

Table 1. E/P Variance During Training and Validation

Set	E/P Ratio Variance
Training (2006-2015)	0.00328
Validation (2016-2022)	0.00320

Descriptive Statistics on the Whole Sample

Descriptive statistics are presented in Table 2. Values are separated into different categories including Min, 10th percentile (10%), 25th percentile (25%), mean, median, Standard Deviation (Std), 75th percentile (75%), 90th percentile (90%), Max, Count, and % Missing. (the percentage of the values that are missing).

Table 2. Descriptive Statistics on the Whole Sample

	Min	10%	25%	Mean	Median	Std	75%	90%	Max	Count	% Missing
lag1mcreal	1887	100451	338642	10822027	1370842	51479707	5124846	19646205	3270662741	477188	0%
finlag1bm	0.00	0.00	0.20	0.54	0.46	0.49	0.77	1.09	4.74	412203	14%
fing01dyadj	0.00	0.00	0.00	0.02	0.01	0.02	0.02	0.04	0.11	477188	0%
fing02esg	0.00	0.00	0.00	17.90	0.00	23.86	33.96	55.93	95.68	210534	56%
fing03nibadj	-0.20	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.05	173489	64%
fing04fcfyadj	-0.19	0.00	0.00	0.00	0.00	0.03	0.00	0.02	0.95	143871	70%
fing05rdsadj	0.00	0.00	0.00	0.02	0.00	0.07	0.01	0.09	1.39	420917	12%
fing08sadadj	0.00	0.00	0.00	0.51	0.00	0.67	1.01	1.21	4.81	220000	54%
fing09shoadj	0.00	0.00	0.01	0.05	0.03	0.07	0.06	0.11	0.53	476894	0%
fing10shiadj	0.00	0.00	0.00	0.03	0.00	0.07	0.03	0.09	0.65	176881	63%
fing11ret5adj	-0.98	-0.27	0.00	0.88	0.38	1.70	1.19	2.39	13.49	398589	16%
fing12empadj	0.00	0.66	0.89	0.89	0.96	0.30	1.02	1.09	2.12	436110	9%
fing13sueadj	-9.19	-1.41	0.00	0.93	0.00	2.95	1.87	4.38	10.81	339214	29%
fing14erevadj	-0.19	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.22	361614	24%

The descriptive statistics table summarizes the 14 core independent variables used to predict the adjusted earnings-to-price (E/P) ratio, encompassing 477,188 observations from February 2006 to January 2023. These variables capture a variety of financial characteristics, such as firm size, value orientation, innovation focus, and market sentiment, across a period marked by varying economic conditions. Data completeness varies significantly across the variables. Market capitalization (lag1mcreal) and dividend yield (fing01dyadj) are fully populated. In contrast, free cash flow yield (fing04fcfyadj) exhibits the highest missingness at 70%, followed by net income before extraordinary items (fing03nibadj, 63.6%) and short interest adjusted by institutional ownership (fing10shiadj, 62.9%), reflecting challenges in data availability for certain financial metrics, particularly for smaller firms or those with limited reporting.

The variables have large differences in their distributional patterns. Market capitalization (lag1mcreal) has a mean of $10.82 million but a median of $1.37 million, indicating a heavily right-skewed distribution (standard deviation: $51.48 million), with values ranging from $1,886 to $3.27 billion. This skewness highlights the presence of a few large firms, consistent with financial datasets where giants dominate market cap. Similarly, the book-to-market ratio (finlag1bm, mean: 0.542, median: 0.459) shows moderate right-skewness, with a maximum of 4.74, reflecting a range of value-oriented firms. Variables like R&D-to-sales ratio (fing05rdsadj, mean: 0.024, median: 0) and ESG scores (fing02esg, mean: 17.9, median: 0) are also right-skewed, with many firms reporting zero values, indicating that innovation and ESG focus are concentrated among a subset of firms.

Variability is notable across other predictors. Short interest (fing09shoadj, mean: 0.048, median: 0.026) and earnings revisions (fing14erevadj, mean: -0.00001, median: 0) exhibit relatively low means but wide ranges, suggesting a variety of market sentiments and analyst expectations across the sample. The employment growth proxy (fing12empadj, mean: 0.893, median: 0.963) and standardized unexpected earnings (fing13sueadj, mean: 0.925, median: 0) show more symmetric distributions around their medians, with tighter interquartile ranges, indicating less extreme variation.

Listed below is an explanation of each of the variables used in the dataset:

- Group 1 (g01dyadj) focuses on Dividend Yield, calculated as the trailing 12-month dividend divided by the stock price at the end of the month.
- Group 2 (g02esg) uses the ESG (Environmental, Social, Governance) score, which is calculated annually and reflects a company's overall commitment to sustainable and ethical practices.
- Group 3 (g03nib) tracks Net Insider Buying, the percentage of shares purchased by company insiders minus those sold.
- Group 4 (g04_fcfyadj) looks at Free Cash Flow Yield, which is the free cash flow as a percentage of market capitalization.
- Group 5 (g05_rdsadj) examines the ratio of Research & Development expenses to sales. A higher ratio indicates that a company is investing more in innovation.
- Group 8 (g08_sadadj) compares the growth rates of sales and advertising expenses.
- Group 9 (g09shoadj) tracks Short Interest, or the percentage of shares sold short.
- Group 10 (g10shiadj) also measures Short Interest, similarly, reflecting investor pessimism.
- Group 11 (g11ret5adj) measures the trailing 5-year return, including delisting returns. Companies with higher 5-year returns tend to experience lower subsequent returns, based on the idea that past performance may not always predict future success.
- Group 12 (g12empadj) looks at Employee Turnover relative to Revenue Change, comparing the percentage change in the number of employees to the percentage change in revenue. While the relationship with returns isn't clear, high turnover relative to revenue changes may indicate operational inefficiencies.
- Group 13 (g13sueadj) focuses on Standardized Unanticipated Earnings (SUE), which reflects the difference between actual earnings and analysts' consensus forecasts.
- Group 14 (g14erevadj) tracks Analyst EPS Revisions, scaling changes in earnings forecasts by the stock price.

Results

Random Forest

This study employs Random Forests, an ensemble machine learning method that enhances predictive accuracy by combining multiple decision trees, to model the E/P ratio for Hewlett Packard Enterprise (HPE). Random Forests work by building many trees, each trained on a random subset of the data and considering a random subset of features at each split, then averaging their predictions. This approach reduces overfitting to a training set when compared to a single Decision Tree. Random Forests are particularly useful for financial data like the one in this project, as they capture non-linear, context-dependent patterns without requiring distributional assumptions, handle high-dimensional data with 26 features, and generalize across economic regimes, such as the volatile 2006–2015 training period and the more stable 2016–2022 validation period.

To ensure the Random Forests model performed optimally, I tuned its hyperparameters using the validation set (2016–2022). Hyperparameters are critical settings that govern the structure and behavior of the Random Forest and its individual decision trees, influencing the model's ability to balance complexity and generalization. I employed Bayesian optimization to test hyperparameters. This is a technique that efficiently searches the hyperparameter space to maximize predictive performance, as measured by the R^2 on the validation set. Bayesian optimization works by constructing a probabilistic model of the objective function—in this case, the validation R^2—and iteratively selects hyperparameter combinations to evaluate based on past results. Bayesian optimization intelligently balances exploration (trying new combinations) and exploitation (focusing on promising regions), making it efficient for complex models like Random Forests.

I tuned the following hyperparameters by testing a range of values to identify the combination that maximized the validation R^2 value:

- **max_depth:** This parameter restricts the maximum depth of each tree, preventing excessive growth that could lead to overfitting. A deeper tree can capture more complex patterns but risks fitting noise in the training data. Values between 3 and 15 were tested.
- **min_weight_fraction_leaf:** This specifies the minimum weighted fraction of the total sum of weights (of the input samples) required at a leaf node, acting as a regularization parameter to control the robustness of the leaves. A higher value forces leaves to contain a larger fraction of samples, reducing sensitivity to small sample variations and preventing overfitting. Values between 0.01 and 0.15 were tested to ensure the trees were neither too shallow (underfitting) nor too sensitive to noise (overfitting).
- **n_estimators:** This determines the number of trees in the Random Forest. A larger number of trees generally improves predictive accuracy by averaging out individual tree errors, but it increases computational cost. Values between 100 and 500 were tested.
- **max_features:** This sets the maximum number of features considered for splitting at each node in a tree, introducing randomness to ensure diverse trees within the forest. A larger number of trees generally improves predictive accuracy by averaging out individual tree errors, but it increases computational cost. Values between 1 and 26 were tested.
- **max_samples:** This controls the fraction of the original dataset sampled with replacement (via bootstrap sampling) for training each tree, a key component of the bagging process in Random Forests. A smaller fraction increases diversity among trees by training them on more varied subsets, while a larger fraction makes trees more similar to each other, potentially improving individual tree accuracy but reducing the ensemble's variance reduction. Values between 0.5 and 0.9 were tested.

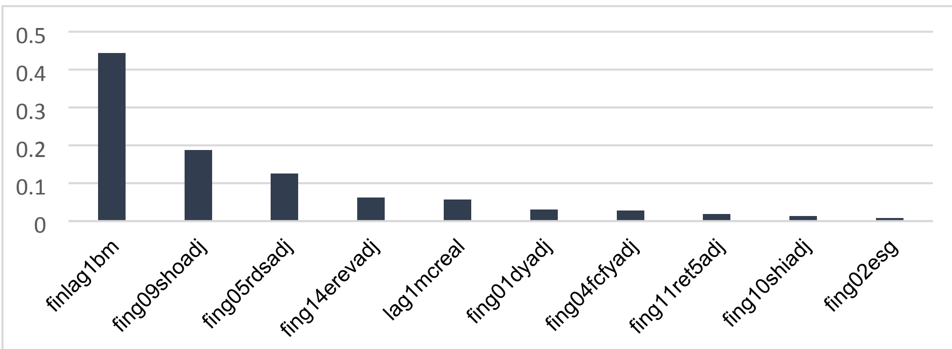
After running 50 trials, the optimal result had a combination of 15 on the max depth, 0.01 min_weight_fraction_leaf, 500 n_estimators, 13 max_features, and 0.9 max_samples to be the best performing hyperparameters leading to a 0.1320 R^2 in the training set (r2_train), and an 0.1259 R^2 in the validation set (r2_valid). In Table 3, the top ten trials are ranked from highest to lowest R^2 in the validation set.

Table 3. Hyperparameter Optimization Results

Rank	r2_valid	r2_train	n_estimators	max_depth	min_samples_leaf	max_features	max_samples
1	0.1259	0.1320	500	15	0.01	13	0.9
2	0.1259	0.1320	500	14	0.01	13	0.9
3	0.1259	0.1318	500	13	0.01	12	0.9
4	0.1259	0.1320	450	15	0.01	13	0.9
5	0.1258	0.1320	372	15	0.01	13	0.9
6	0.1257	0.1320	147	13	0.01	12	0.9
7	0.1255	0.1310	500	11	0.01	13	0.9
8	0.1254	0.1322	500	13	0.01	14	0.9
9	0.1251	0.1306	463	15	0.01	11	0.9
10	0.1251	0.1299	402	12	0.01	12	0.9

Feature importance scores, derived from the final trained model, measure the relative contribution of each input variable to reducing mean squared error across the forest. These scores provide insight into which features most influenced the model's predictions during the validation period. Figure 1Figure 1 presents the feature importance scores for the features used in the model that had greater than 1% importance. The variable with the highest importance factor by a large margin was finlag1bm (book to market ratio), with 44.4% importance. This was followed by fing09shoadj (short interest) with a 18.7% importance followed later by fing05rdsadj (previous five years returns) at 12.6% and fing14erevadj (analyst EPS revisions) at 6.2%.

Figure 1. Random Forest Importance Values



Comparison to Vanilla Decision Trees

To evaluate the predictive performance of the models for estimating the E/P ratio, I trained both a Random Forest and a Vanilla Decision Tree on a dataset spanning 2006 to 2023, with training (2006–2015), validation (2016–2022), and test (2023) sets, and computed R^2 values for the training set, validation set, and the entire dataset (Table 4). The Vanilla Decision Tree was trained under the same process of the Random Forest analysis, training the tree on the training set and optimizing the hyperparameters on the validation set. Hyperparameters were optimized using Bayesian Optimization, testing minimum weight fraction leaf values between 0.01 and 0.15, and maximum depth values between 3 and 15.

The Random Forest achieved R^2 values of 0.1319 (training), 0.1259 (validation), and 0.1470 (entire dataset), while the Decision Tree recorded 0.0845 (training), 0.1072 (validation), and 0.0937 (entire dataset). The Random Forest outperforms the Decision Tree across all sets, with its validation R^2 of 0.1259 notably higher than the Decision Tree's 0.1072, reflecting better generalization to unseen data. This result aligns with expectations, as Random Forest's ensemble approach reduces overfitting compared to a single Decision Tree. Given these findings, the Random Forest is better suited for estimating a fair E/P ratio for Hewlett Packard Enterprise, as its reduced overfitting ensures more reliable predictions across diverse market conditions.

Table 4. Vanilla Decision Tree and Random Forest Comparison

	Validation R^2	Training R^2	All R^2
Random Forest	0.1259	0.1319	0.147
Vanilla Decision Tree	0.1072	0.0845	0.0937

To further ensure the Random Forest model was working properly, I computed the R^2 values of all 500 trees within the Random Forest analysis (n_estimators) across the training and validation dataset and ranked them. In Table 5, the top ten trees are shown with a maximum R^2 value of 0.1290, followed by .1278, .1277, and so on.

Table 5. Top Ten Individual Decision Trees R^2 Values

Rank	Score	Rank	Score
1	0.12900	6	0.12738
2	0.12780	7	0.12732
3	0.12770	8	0.12731
4	0.12766	9	0.12724
5	0.12741	10	0.12702

In Table 6, the following comparison is made with the first two columns being R^2 values for vanilla decision trees, and the last column being R^2 values for the Random Forest analysis across both the training and validation datasets combined. The average R^2 of 0.1206 reflects the typical performance of a Decision Tree across multiple runs or configurations, while the maximum R^2 of 0.1290 represents the best-case scenario for a single Decision Tree. However, even the best Decision Tree falls short of the Random Forest's R^2 of 0.1467, a difference of 0.0177 or about 13.7% relative improvement over the maximum Decision Tree R^2. These results are in line with expectations as statistically, the mean prediction of the ensemble has a lower error than the average error of individual trees, as averaging smooths out individual tree errors, leading to better generalization across the dataset.

Table 6. Individual Vanilla Decision Tree and Random Forest R^2 Values

Average (Vanilla Decision Tree)	Maximum Value (Vanilla Decision Tree)	Random Forest
0.1206	0.1290	0.1467

Quintile Analysis

To investigate the relationships between the E/P ratio and key predictor variables, I conducted a quintile analysis on the entire dataset, grouping observations into five quintiles based on actual E/P ratios, from the lowest (Quintile 1) to the highest (Quintile 5). For each quintile, I calculated the mean and median of the actual E/P ratio, the predicted E/P ratio (pred_ep), and the top five predictor variables by Random Forest feature importance: book-to-market ratio (finlag1bm), short interest ratio (fing09shoadj), R&D-to-sales ratio (fing05rdsadj), earnings revisions (fing14erevadj), and lagged market capitalization (lag1mcreal). The results (Table 7) reveal patterns. The actual E/P ratio increased from a mean of 0.038 in Quintile 1 to 0.094 in Quintile 5, with the predicted E/P ratio following a similar trend (0.047 to 0.090), indicating the Random Forest model's ability to rank observations effectively, though it slightly overestimates at the lower end (0.047 vs. 0.038) and underestimates at the higher end (0.090 vs. 0.094).

The book-to-market ratio (finlag1bm) showed a strong positive relationship, rising from a mean of 0.188 in Quintile 1 to 1.028 in Quintile 5, suggesting that value stocks with higher book-to-market ratios are associated with higher E/P ratios. The short interest ratio (fing09shoadj) also exhibited a positive relationship, increasing from 0.034 to 0.074, indicating that higher E/P ratios are linked to increased bearish sentiment. In contrast, the R&D-to-sales ratio (fing05rdsadj) showed a strong negative relationship, decreasing from 0.093 to 0.004, suggesting that firms with lower R&D spending relative to sales tend to have higher E/P ratios, possibly prioritizing profitability over innovation. Earnings revisions (fing14erevadj) displayed a weak negative relationship, decreasing from 0.001 to -0.003, indicating slightly more negative analyst revisions for stocks with higher E/P ratios, though the small magnitude suggests a limited impact. Finally, lagged market capitalization (lag1mcreal) showed a negative relationship, decreasing from 36,083,314 to 6,207,836, suggesting that smaller firms are associated with higher E/P ratios, potentially due to higher risk or undervaluation.

Table 7. Random Forest Quintile Analysis

	Actual E/P		Predicted E/P		finlag1bm		fing09shoadj		fing05rdsadj		fing14erevadj		lag1mcreal	
Quintile	mean	median	mean	median	mean	median	mean	median	mean	median	mean	median	mean	median
1	0.039	0.033	0.043	0.044	0.202	0.197	0.036	0.024	0.078	0.047	0.001	0.000	21865586	3486168
2	0.054	0.047	0.056	0.056	0.279	0.293	0.035	0.024	0.025	0.000	0.001	0.000	13160279	1864212
3	0.064	0.055	0.065	0.065	0.445	0.489	0.042	0.028	0.009	0.000	0.001	0.000	7189138	1165329
4	0.074	0.064	0.074	0.074	0.689	0.727	0.042	0.025	0.004	0.000	0.000	0.000	6262266	805156
5	0.099	0.079	0.093	0.088	1.094	1.016	0.084	0.032	0.006	0.000	-0.003	0.000	5632844	772517

Conclusion

On January 31, 2023, the Random Forest model predicted an E/P ratio of 0.0679 for Hewlett Packard Enterprise (HPE), indicating a moderate earnings yield based on a quintile analysis of the entire dataset, which includes both the training and validation sets. The entire dataset was used for this analysis to provide a comprehensive historical context for interpreting the prediction, capturing a wide range of market conditions over the 17-year period. This analysis positions the predicted E/P between Quintile 3 (mean E/P: 0.065, median: 0.065) and Quintile 4 (mean E/P: 0.074, median: 0.074), suggesting that HPE's earnings yield is relatively elevated within the historical distribution of E/P ratios across the full dataset. The prediction is driven by the model's emphasis on key financial characteristics, captured through the top five predictors: book-to-market ratio (finlag1bm, 44.4% importance), short interest ratio (fing09shoadj, 18.7%), R&D-to-sales ratio (fing05rdsadj, 12.6%), earnings revisions (fing14erevadj, 6.2%), and lagged market capitalization (lag1mcreal, 5.7%), which together account for approximately 87.6% of the model's feature importance. Table 8 details the values of these predictors for HPE during the test period, as well as the observed features importance values.

Table 8. Predictor Values and Importance Values

Metric / Predictor	Value for HPE	Feature Importance
Book-to-Market Ratio (finlag1bm)	1.104	44.4%
Short Interest Ratio (fing09shoadj)	0.030	18.7%
R&D-to-Sales Ratio (fing05rdsadj)	0.020	12.6%
Earnings Revisions (fing14erevadj)	0.001	6.2%
Lagged Market Cap (lag1mcreal)	21664747	5.7%

HPE's book-to-market ratio (1.104) exceeds the Quintile 5 mean (1.094) from the entire dataset. The book-to-market ratio is a characteristic often associated with undervaluation by the market. Such firms typically exhibit higher E/P ratios as investors demand a higher earnings yield to compensate for perceived risks or lower growth prospects.

Conversely, HPE's short interest ratio (0.030) falls below the Quintile 1 mean (0.036), Short interest is often higher for stocks with elevated E/P ratios, as investors may anticipate price declines in riskier or overvalued firms. HPE's low value suggests a lack of pessimism, a trait more common in stocks with lower E/P ratios, where investor confidence supports higher valuations.

HPE's R&D-to-sales ratio (0.012) aligns closely with Quintile 2 (mean: 0.025) and is well below Quintile 1's mean (0.078), This metric highlights HPE's limited investment in research and development relative to sales, suggesting a restrained focus on innovation compared to R&D-intensive firms, which often exhibit lower E/P ratios potentially due to higher growth expectations. R&D investment directly lowers earnings, the numerator for the E/P ratio, making firms who invest heavily in R&D potentially have naturally lower E/P ratios. HPE's modest innovation profile, more typical of firms with lower E/P ratios, further moderates the prediction.

The earnings revisions mean for HPE (0.001) is slightly positive, aligning with the Quintile 1–3 mean (0.001). This predictor reflects analyst updates to earnings forecasts, with positive revisions indicating optimism about future performance. This sentiment often correlates with lower E/P ratios, as rising earnings expectations can boost stock prices, reducing the earnings yield. HPE's mild upward revisions may contribute to a small downward pressure on the predicted E/P ratio, reinforcing the moderate influences of other factors.

Finally, HPE's lagged market capitalization (21,664,747) exceeds the Quintile 1 mean (21,865,586) confirming its status as a large firm. Larger firms, often seen as stable and less risky, typically have lower E/P ratios due to higher valuations as these firms are less volatile than smaller firms. This characteristic of HPE further balances the prediction, aligning it more closely with traits of lower E/P stocks.

The predicted E/P ratio of 0.0679 reflects the dominant influence of finlag1bm, tempered by counteracting factors such as HPE's lack of investing in R&D, larger size, and reduced bearish sentiment, resulting in a balanced prediction between Quintiles 3 and 4.

Overall, I recommend that investors seek to acquire this stock at or above an earnings-to-price ratio of 0.0679 (6.79%) during the month of January 2023. Based on the analysis, a 6.79% earnings yield implies the stock is fairly valued, and values higher than 6.79% may indicate that the stock is undervalued.

References

- CBOE VIX Index: https://finance.yahoo.com/quote/%5EVIX/