Weak-Form Efficiency of NASDAQ Stocks

Stanford STATS 207 Project

Johannes Fuest
Department of Statistics
Stanford University
jfuest@stanford.edu

Diane Sarkis
Department of Mathematics
Stanford University
diane.sarkis@stanford.edu

Jonathan Williams
Department of Statistics
Stanford University
jonwill@stanford.edu

Abstract

If it is the case that non-random patterns in stock prices change over time and are difficult to model and exploit ex-ante, then traditional backward-looking weak-form efficiency tests that uncover non-randomness ex-post need not be seen as credible evidence of stock-market inefficiencies. In this paper we perform traditional weak-form efficiency tests on NASDAQ-listed stocks between 1960 and 2024 and then try to fit carefully tuned ARIMA models to exploit the supposedly inefficient stocks. This approach results in poor out-of-sample performance. Our results show that non-randomness and efficiency may not be mutually exclusive in practice due to the computational and statistical difficulties that arise when attempting to exploit time-dependent non-randomness in stock price time series.

1 Introduction

In the landscape of financial markets, the question of whether stock markets are efficient has captivated the attention of investors, financial analysts, and policymakers for decades. The efficient market hypothesis (EMH) in its weak form, suggests that stock prices fully reflect all available historical price information, rendering it impossible to achieve excess returns through technical analysis of past prices. This assertion is based on the belief that stock markets immediately adjust to available information, meaning that stock prices should thus evolve in a manner that is inherently unpredictable, driven solely by new information as it becomes available and not by historical price patterns.

This project seeks to rigorously test the weak-form efficiency of the NASDAQ stock market, one of the major stock markets in the U.S. known for its high daily trading volume and significant market capitalization. By using a comprehensive dataset of daily stock prices from Kaggle, this study applies an array of statistical tests to analyze the randomness of stock prices and then attempts to build predictive models for stocks that exhibit evidence of non-random historical prices.

As financial markets continue to evolve in complexity and scope, the findings of this study aim to contribute valuable insights into the ongoing debate about market efficiency, offering implications for theoretical finance and practical investment strategies. This paper documents our comprehensive analysis, methods, and findings.

2 Related Work

This work is related to previous literature which has investigated the efficiency of capital markets since the 1950s [1][2] [3] [4]. Within this literature, this work focuses on the question of the presence of efficiency of stocks, rather than trying to ascertain the driving factors behind it, which is done by work on factor models[5] and related attempts to forecast prices.

As such, this work is directly related to statistical financial literature that tests whether stock returns follow random walks. It directly applies the variance ratio test for weak-form efficiency developed by Lo and MacKinley [6], as well as unit root tests [7], autocorrelation tests[8], and runs tests[9]. There are many works applying one or multiple of the aforementioned tests for randomness of stock returns to different markets all around the world, such as India [10], Bangladesh [11], Nigeria [12], Japan [13], and many others [14] [15]. Nasdaq-specific work relating to market efficiency also exists, examining not only the presence of market efficiency [16], but also its relation to other factors[17] [18], as well as its changes over time[19].

This work contributes to the existing literature in two ways. First, it provides an updated account of the weak-form efficiency of the Nasdaq stock exchange, doing so based on all four commonly used tests, as well as on a comprehensive, stock-by-stock basis, rather than only looking at the exchange as a whole. Secondly, this work attempts to put forth an additional dimension for testing weak-form efficiency, by trying to generate predictive models for stocks which traditional tests deem to be inefficient.

3 Data

Original Dataset The empirical analysis for this project is based on an extensive dataset containing daily stock price data of 3404 U.S. stocks traded on the NASDAQ stock exchange from 1960 to 2024. This dataset, sourced from Kaggle, provides comprehensive coverage with over 11 million observations of open, close price, high, and low prices for each stock. We chose this dataset in large part for its accessibility, comprehensive coverage and large number of observations.

Data Preprocessing The dataset contained very few missing values, and the data cleaning revealed almost no outliers and anomalies that could skew the analysis. While the dataset encompasses data from 1960 to 2024, not all stocks have complete records for the entire duration. For the purpose of uniformity and to ensure robust statistical inference, the analysis was confined to stocks with at least 60 days of trading history. This filtering ensured that the models were trained and tested on sufficient data points.

After this data preprocessing, we were left with 3,273 stocks and over 11 million observations.

4 Methods

To test the veracity of the EHM, we first ran statistical test to identify non-random stock time series. We conducted a set of four different tests on the time series of every single one of the over 3000 stocks in our dataset. We then built ARIMA models to forecast the prices of a subset of stocks that did not pass the tests for weak-form efficiency. Violation of the EHM would be implied if the ARIMA models elicited low RMSEs and strong out-of-sample predictive performance. Due to computational constraints, our analysis was restricted to 100 stocks and ARIMA class models on a relatively small grid search with cross validation that is described in the following sections.

4.1 Stationary Tests

Recall that under the EMH, a stock time series is a random walk. To identify stocks which violate the EMH, we ran four statistical tests which check for significant divergence from theoretical random walk properties.

4.1.1 Autocorrelation Test

We know from lecture that random walks do not exhibit significant autocorrelation. The Ljung-Box-Pierece Test fashions a test statistic which is a function of the sum of time series autocorrelations ($\hat{\rho_k}$) and follows a known distribution. The Null Hypothesis of the test is that the time series is random. The Alternate Hypothesis is that the time series is not random. The Test Stat:

$$Q = n(n+2)\sum_{k=1}^{h} \frac{\hat{\rho}_k^2}{n-k} \tag{1}$$

asymptotically follows the chi-squared distribution [20]. This we reject the null if:

$$Q > \chi^2_{1-\alpha,h} \tag{2}$$

Rejection of null implies the stock time series exhibits significant autocorrelation, which implies stock time series is not random, which in turn implies the EMH is violated.

4.1.2 Runs Test

It is known that in a random data sample, the probability of the $(k+1)$th value being greater than the kth value follows a binomial distribution [21]. This fact determines the distribution of the number of runs in a random data series and directly underpins the runs test. The Null Hypothesis of the test is that the time series is random. The Alternate Hypothesis is that the time series is not random. The Test Stat:

$$Z = \frac{R - \mu_R}{\sigma_R} \tag{3}$$

is asymptotically normal [21]. Thus we reject the null if:

$$|Z| \geq Z_{1-\alpha/2} \tag{4}$$

Rejection of null implies the stock time series exhibits a very low or high run #, which implies the stock time series is not random, which in turn implies the EMH is violated

4.1.3 Augmented Dickey Fuller Test (ADF)

We know from lecture that random walks can be written as an AR(1) process:

$$y_t = \phi y_{t-1} + Z_t \tag{5}$$

where $\phi = 1$. Notice that if a given time-series is a random walk, then it must follow:

$$\begin{aligned} \Delta y_t &= y_t - y_{t-1} && (6)\\ &= \phi y_{t-1} + Z_t - y_t && (7)\\ &= (1-\phi) y_{t-1} + Z_t && (8) \end{aligned}$$

The ADF Test functions by running t-tests for the slope parameter from the basic auto regression:

$$\Delta y_t = \beta_0 + \beta_1 y_{t-1} \tag{9}$$

The ADF Test also runs slope t-tests for slight modifications of (9) which account for trend, higher-order autoregrssive dynamics, or random walks with drift. The Null Hypothesis is that the time series is not stationary (auto regression slope $B_1 = 0$). The Alternate Hypothesis is that the time series is stationary (auto regression slope $B_1 \neq 1$). The Test Stat:

$$\text{DF} = \frac{\hat{B}_1}{\text{SE}(B_1)} \tag{10}$$

asymptotically follows the t-distribution [22]. Thus we reject the null if:

$$\text{DF} \geq t_{1-\alpha/2} \tag{11}$$

Rejection of null implies the stock time series is stationary, which implies the stock time series is not random, which in turn implies the EMH is violated.

4.1.4 Variance Ratio Test

It is known from [23] that if a time series is a random walk, then the variance of a k period: $\text{Var}(X_t, X_{t-1}, \cdots X_{t-k+1})$ equals k times the variance of a one period: $\text{Var}(X_t, X_{t-1})$. This fact determines the distribution of the k to 1 period variance ratio and underpins the Variance Ratio Test. The Null Hypothesis of the test is that the time series is random. The Alternate Hypothesis is that the time series is not random. The Test Stat:

$$Z(k) = \frac{\text{VR}(k) - 1}{\sqrt{\phi(k)}} \tag{12}$$

where $\phi(k) = \frac{2(2k-1)(k-1)}{3kT}$. Because the test statistic is asymptotically normal [24], we reject the null if:

$$|Z(k)| \geq Z_{1-\alpha/2} \tag{13}$$

Rejection of the null implies the stock time series exhibits low/high k to 1 period variance ratios, which implies the stock time series is not random, which in turn implies the EMH is violated.

4.2 Modelling

While the efficient market hypothesis is often argued to require stock prices to follow a random walk, this argument rests on the fundamental assumption that any non-random stock price movement can be exploited and morphed into a trading strategy that generates excess returns. Next to the matter of transaction costs, which could provide a significant obstacle to this, we argue that the modelling process itself presents a challenge to the connection between the efficient market hypothesis and the idea that stock prices must follow a random walk. If it is the case that the non-random patterns are extremely difficult to model and exploit ex-ante or shift significantly over time, then backward-looking tests that uncover ex-post non-randomness need not be seen as credible evidence of stock-market inefficiencies. In order to demonstrate this, we apply ARIMA models, a powerful, flexible tool for modelling time series to the stock price data and examine their performance.

4.2.1 Grid Search and Cross Validation

To both efficiently use all available data, and maximize the potential for out-of-sample performance, we utilized cross validation as our model selection schema. For each stock, we ran a shallow grid search over ARIMA parameters and selected the model which elicited the highest 3-fold Cross Validation RMSE. We considered up to four autogressive and moving average terms, as well as up to double differencing. The exact cross-validation algorithm used is described below:

Algorithm 1 ARIMA Cross-Validation

Require: Potential ARIMA Models $M_1, \ldots, M_\ell$
Require: Folds Number K
  for each model M_i **do**
    for $k = 0, \ldots, K-1$ **do**
      Fit M_i to the data from the first k folds.
      Forecast M_i over the remaining $K-k$ folds & compute RMSE.
    end for
    Let CV_i = average M_i RMSE over all folds
  end for
  return M_i with the lowest associated CV_i

4.2.2 Out of Sample Prediction Generation

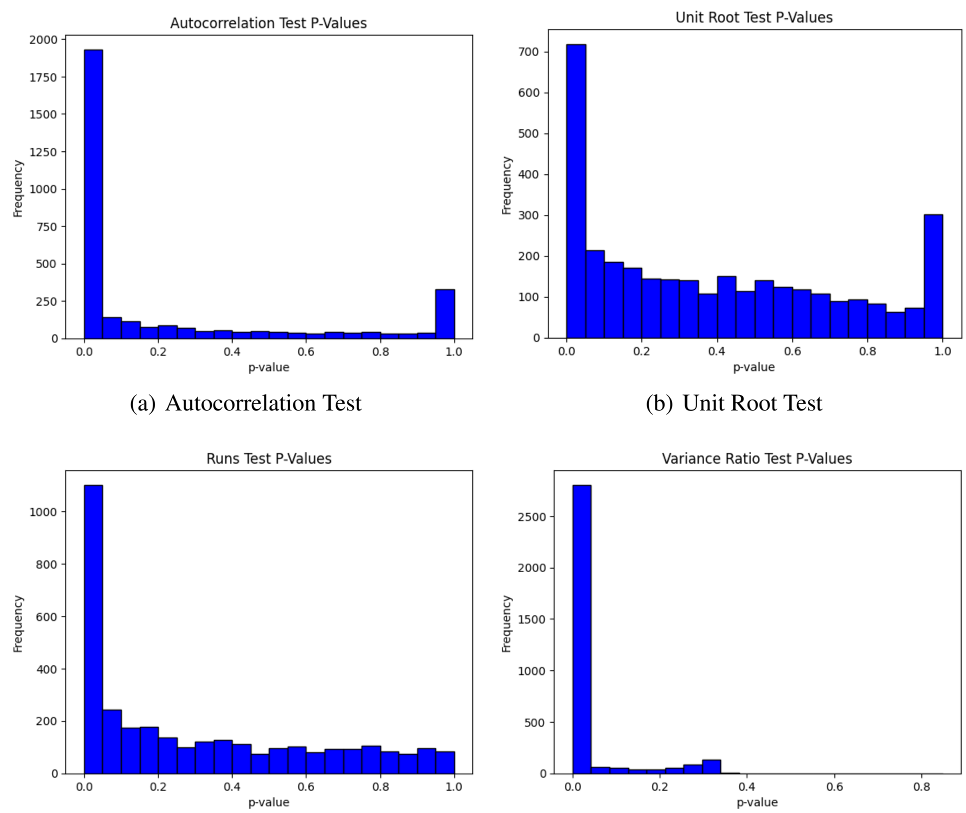
Upon finding the most suitable ARIMA configuration for every stock, we then proceeded to train the model on all but the last 30 days of historical stock prices, and predicted the final 30 days, refitting the model on a rolling basis with the most recent day after each prediction to give the best possible out-of-sample prediction at each timestep.

5 Results and Discussion

This section will first discuss the results of the four tests we applied to all stocks in the dataset, assessing their statistical conclusions, as well as the resulting evidence for or against weak-form efficiency of the NASDAQ. Then we will discuss the results of our ARIMA models for a subsample of 100 randomly chosen stocks via grid search and cross-validation.

Test	% rejected	Avg. p-value
Autocorrelation	59.03	0.22
Unit Root	21.94	0.39
Runs	33.64	0.30
Variance Ratio	86.22	0.03

Table 1: Weak-Form Testing Results



(a) Autocorrelation Test (b) Unit Root Test

(c) Runs Test (d) Variance Ratio Test

Figure 1: Distribution of p-values of weak-form tests across stocks

The results of the four weak-form tests are summarized in Table 1 and Figure 1. All four tests detected a mixture of weak-form efficient and inefficient stocks. The variance ratio test classified the most stocks as inefficient at 86.22%, whereas the unit root test classified the least stocks as inefficient at 21.94%. Overall, there was a wide variety in stock's efficiencies across all tests, as shown by the p-value histograms. This demonstrates that it is difficult to assess the efficiency of the NASDAQ as a whole and indicates that examining individual stocks may be more compelling.

While the differences between the results of the tests are relatively large, this can be explained by the different aspects of randomness they test for respectively. The variance ratio test for example, is concerned with with ratio of the variance of single-period returns, compared to that of multi-period returns. This test, like the autocorrelation test is sensitive to time series with bursts of short-term volatility. The unit root on the other hand only checks for a unit-root in the time series, which is a long-term observation and may not be impacted by such short-term behavior. Similarly, the runs test may not be as sensitive to short-term volatility if is expressed mainly through magnitude of price changes, rather than sign changes.

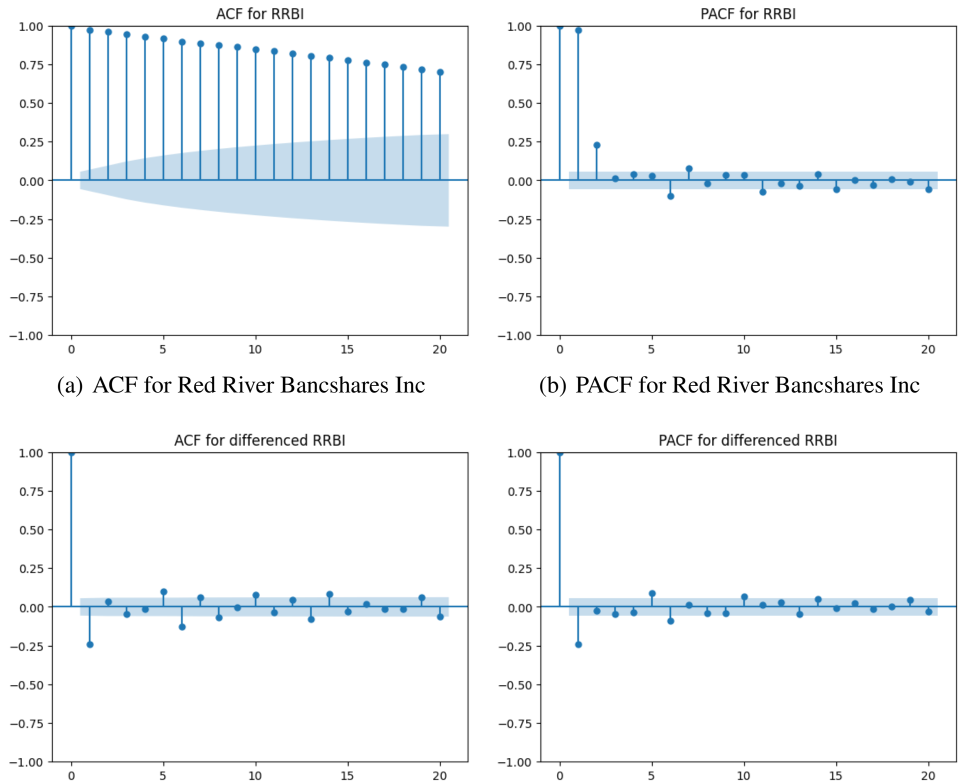
While the results of these tests are statistically significant, their implications for practitioners are thus not immediately apparent and exploiting the non-randomness detected by them requires a modelling approach that correctly captures the non-randomness observed in the pricing patterns.

Models	**% correct sign predicted**	**MAE**
All	46.22	0.34
Passed BJP Test	46.22	0.32
Failed BJP Test	46.23	0.37

Table 2: ARIMA Modelling Results

Table 2 shows the performance of the ARIMA models that resulted from applying gridsearch and cross-validation to 100 randomly chosen stocks. Due to computational constraints, this grid search only allowed up to 4 AR and MA terms, as well as differencing up to two times. All ARIMA models' residuals were tested for autocorrelation using the Ljung-Box-Pierce test. As can be seen by the sub 50% of correctly predicted sign changes in returns, the models exhibited generally poor out of sample predictive power. This poor out-of-sample performance was similarly bad for models which passed the LBP test, compared to those that did not.

There are many potential reasons for this poor performance. First, the computational constraints rendered a grid-search over a wider parameter space infeasible, which may not been in line with the structure of financial time series, which might be too complex for our ARIMA models to capture. Adding a higher number of lag terms, as well as seasonality may have lead to better models but this is not guaranteed, as exemplified by the complex shapes of the ACF and PACF of our sampled stocks as shown for the example stock RRBI in Figure 2. Even with a wider grid search and seasonality, other, more powerful methods, such as LSTMs may be required to model more complex aspects of the time series. Furthermore, our models were fit on the whole time series, which ignores time-shifting properties of the stock behind the time series. Over the decades, many firms may have switched industries, changed business models or faced other forms of disruption fundamental enough to radically change the behavior of their stock prices.



(a) ACF for Red River Bancshares Inc (b) PACF for Red River Bancshares Inc

(c) ACF for differenced Red River Bancshares Inc (d) PACF for differenced Red River Bancshares Inc

Figure 2: ACF and PACF for Red River Bancshares Inc exemplifies complexity of stock price time series

While the aforementioned reasons may be overcome, it remains unclear whether consistent out-of-sample performance is achievable through time-series modelling of these 'inefficient' stocks even with higher modelling sophistication and computing resource availability. Thus, while the weak-form tests commonly performed in the financial literature are quick to uncover many non-random properties of stock price time series, exploiting these is a distinct problem. This qualifies the criterion of weak-form efficiency as very distinct from the presence of observably non-random patterns in past stock prices.

Future Work Building on our initial findings, several avenues for future research can be explored to further investigate the efficiency of NASDAQ stocks. While our current study utilized a subset of randomly selected stocks due to computational constraints, future work could aim to include a full analysis of all NASDAQ stocks. Incorporating more complex seasonality models with an expanded grid search or exploring other classes of time series models could potentially improve predictive performance.

Furthermore, the performance of more our models could be analyzed more systematically, using optimal portfolio construction and returns baselines, such as CAPM or market portfolios to benchmark returns and evaluate the ability of the models to generate excess returns that would directly contradict weak-form efficiency.

6 Code

For a repository, including instructions on how to run the code for all of our experiments, please see github: `https://github.com/johannesfuest/time_series_for_stock_price_prediction/tree/main`

References

[1] Maurice George Kendall and A Bradford Hill. The analysis of economic time-series-part i: Prices. *Journal of the Royal Statistical Society. Series A (General)*, 116(1):11–34, 1953.

[2] Eugene F Fama. Efficient capital markets. *Journal of finance*, 25(2):383–417, 1970.

[3] Werner FM De Bondt and Richard Thaler. Does the stock market overreact? *The Journal of finance*, 40(3):793–805, 1985.

[4] Eugene F Fama and Kenneth R French. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56, 1993.

[5] Robert R McCrae and Oliver P John. An introduction to the five-factor model and its applications. *Journal of personality*, 60(2):175–215, 1992.

[6] Andrew W Lo and A Craig MacKinlay. Stock market prices do not follow random walks: Evidence from a simple specification test. *The review of financial studies*, 1(1):41–66, 1988.

[7] Graham Elliott, Thomas J Rothenberg, and James H Stock. Efficient tests for an autoregressive unit root, 1992.

[8] Greta M Ljung and George EP Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303, 1978.

[9] Marilyn A Agin and Anant P Godbole. A new exact runs test for randomness. In *Computing Science and Statistics: Statistics of Many Parameters: Curves, Images, Spatial Models*, pages 281–285. Springer, 1992.

[10] Rakesh Gupta and Parikshit K Basu. Weak form efficiency in indian stock markets. *International Business and Economics Research Journal*, 6(3):57–64, 2007.

[11] Ainul Islam and Mohammed Khaled. Tests of weak-form efficiency of the dhaka stock exchange. *Journal of Business Finance & Accounting*, 32(7-8):1613–1624, 2005.

[12] R Ayodeji Olowe. Weak form efficiency of the nigerian stock market: further evidence. *African development review*, 11(1):54–68, 1999.

[13] Venkata Rajasekhar Ryaly and RSRK Kiran Kumar. A study on weak-form of market efficiency in selected asian stock markets. 2014.

[14] Abu Towhid Muhammad Shaker. Testing the weak-form efficiency of the finnish and swedish stock markets. *European Journal of Business and Social Sciences*, 2(9):176–185, 2013.

[15] Chris Fawson, Terry F Glover, Wenshwo Fang, and Tsangyao Chang. The weak-form efficiency of the taiwan share market. *Applied Economics Letters*, 3(10):663–667, 1996.

[16] Michael J Mauboussin. Revisiting market efficiency: The stock market as a complex adaptive system. *Journal of Applied Corporate Finance*, 14(4):47–55, 2002.

[17] James Dow and Gary Gorton. Stock market efficiency and economic efficiency: Is there a connection? *The Journal of Finance*, 52(3):1087–1129, 1997.

[18] Gary C Sanger and John J McConnell. Stock exchange listings, firm value, and security market efficiency: The impact of nasdaq. *Journal of Financial and Quantitative Analysis*, 21(1):1–25, 1986.

[19] Anthony Yanxiang Gu. Increasing market efficiency: Evidence from the nasdaq. *American Business Review*, 22(2):20, 2004.

[20] PSU Department of Statistics. Testing-and-remedial-measures-autocorrelation.

[21] National Institute of Standards and Technology. Runs test for detecting non-randomness.

[22] Wikipedia Editors. Dickey–fuller test.

[23] Andrew W. Lo and A. Craig MacKinlay. Stock market prices do not follow random walks: Evidence from a simple specification test. *The Review of Financial Studies*, 1(1):41–66, 1988.

[24] Mingze Gao. Variance ratio test - lo and mackinlay (1988).