IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Modelling Stock Price Predictability: Insights from Random Walk and Autoregressive Simulations

PARAMITA KARMAKAR

(Independent Financial Analyst)[1]

Abstract

Predicting stock prices is a central challenge in finance. This paper examines stock price behavior through two fundamental models: the Random Walk and the Autoregressive model of order one (AR (1)). Using Monte Carlo simulations, we generate 50 independent price paths over 100 time periods, all starting at the same initial price. The Random Walk paths diverge rapidly, demonstrating unpredictability and randomness, while the AR (1) paths exhibit smoother, persistent trends, indicating some short-term predictability. Through these visual and statistical comparisons, the study highlights how memory in price dynamics influences predictability and provides a hands-on illustration of these opposing views on financial market behavior.

Keywords: Random Walk, Autoregressive Model, Simulation, Efficient Market Hypothesis, Predictability

I. Introduction

Predicting stock prices has been one of the most sought-after tasks over the past decades. According to the *Efficient Market Hypothesis* (EMH), a flagship theory in finance, all available information in the market is already incorporated into current stock prices, provided there is perfect dissemination of information (Fama, 1970). Therefore, predicting the future movement of stock prices is not possible. This phenomenon of stock price unpredictability is mathematically modeled by what is popularly known as Random Walk (RW) model. In a random walk model, the future value depends only on the present value—not on how the present value was reached. The future value one step ahead is determined by a new shock, often in the form of unexpected news or information that randomly affects prices. Since the effect of past values (prices) is non-existent, the process is considered *memory-less*—only the present matters.

In contrast to the nature of price movements that follow a *random walk*, empirical evidence suggests that there are situations where yesterday's price movements may influence today's—at least temporarily. The existence of such *temporal correlation* indicates the presence of *memory* in the price formation process. This behavior can be modeled using what is known in the literature as an *autoregressive* (AR) model.

This paper seeks to address the question: Are stock prices completely random, or do they exhibit short-term predictability? The answer is sought with the help of *simulation* techniques. Stock prices are simulated using both the RW and AR models, starting from identical initial conditions and generating multiple independent paths. These simulations reveal how the same starting point can lead to different price trajectories and end points when a RW model is simulated. In contrast, simulations under an AR setup display some degree of *convergence* and *predictability*.

[1] Formerly affiliated to Surendranth College for Women, Kolkata and ICRA Ltd. (A Moody's Investors Service Company). Email: paramitaa86@gmail.com

The structure of the paper is as follows: Section 2 provides a brief overview of the science of simulation. Section 3 outlines the methodology. Section 4 presents the results, followed by a discussion in Section 5. Finally, Section 6 concludes the paper.

II. The Science of Simulation

Simulation is a scientific technique used to replicate complex real-world systems through computer-generated artificial scenarios. Its main objective is to model situations that cannot be easily experimented on in reality. Financial markets, characterized by complex and dynamic interactions, can be effectively modeled using simulation techniques. By specifying simple but realistic rules that define the *Data Generating Process* (DGP), simulation allows us to generate many possible future outcomes. Now as per the model specifications, the simulated outcomes as generated replicate the real life scenarios as well as the range of probable future paths revealing a deeper understanding of the DGP. The efficacy of theoretical models and intuitive understanding is tested and enriched through visual and statistical examination.

The present study seeks to conduct an empirical experiment on the evolution of stock prices over time using simulation just like a *virtual laboratory*. It is observed how stock prices, starting from the same initial position, may evolve differently under two distinct scenarios: the random walk, which represents a memory-less process, and the autoregressive (AR) model, which incorporates memory. In this context, the term *autoregressive* (AR) refers to a process where the present value of a variable is regressed on its previous values, known as "lags." While the number of lags can be determined using appropriate statistical criteria, the present study focuses on an AR (1) process, which considers just one lag. This approach can be generalized to any number of lags as needed[2].

III. Methodology

1. The Random Walk Model

The Random Walk assumes that the stock price at time t, denoted P_t, is the previous price plus a random shock: $P_t = P_{t-1} + \in_t ; \in_t \sim N(0, \sigma^2)$

Here, $\in_t$ $'s$ are independent, normally distributed shocks with zero mean and variance σ^2. This model embodies the core EMH idea that price changes are completely unpredictable and have no memory.

2. The AR (1) Model

The AR (1) model introduces a dependency of the present value (price) on its previous value (price), representing short-term memory:

$$P_t = \alpha + \rho P_{t-1} + \in_t ; \in_t \sim N(0, \sigma^2)$$

In the present context the values of the model parameters are set as below:

α=10 introduces a positive drift pushing prices upward over time; ρ=0.9 reflects strong persistence in price changes. The value of σ is taken to be 2. The initial price P_0 is taken to be 100.

This model assumes that prices are strongly impacted by their immediate past values and also randomness does still play a role in formation of prices.

3. Simulation Design

For both the RW and AR models, 50 independent simulations are carried, each for 100 time periods. All paths start from the same initial price P_0=100. Random shocks for each time step are independently drawn from

[2] See Hamilton (1994) for a detailed discussion on this.

their respective normal distributions. This design allows us to observe and compare how different model assumptions influence the dispersion, trend, and overall behavior of price paths over time.

IV. Results

Looking at the simulated paths one may observe the visual contrasts. The RW paths after beginning at the same initial position diverge rapidly over time. The randomness of shocks coupled with the absence of any drift or memory makes some price paths go up on the one hand and let others fall down drastically, on the other. This spread between the max and min is the characteristic of the Random Walk.

In contrast to this pattern, a much more structured pattern is generated by AR (1) simulated paths. Looking at the first few simulated paths, one may observe that prices tend to move upward gradually staying within a tighter band. Although the random error term is there, it is due to the presence of the autoregressive term which introduces memory in the process resulting in price changes which are influenced by previous values.

Whether the underlying model is that of *persistence* or of *randomness* is what determines the price trajectory. This is what the simulated paths show. In essence we observe that the RW paths (Fig. 1) are characterized by divergence in contrast to the AR (1) paths that have a tendency to cluster and follow a shared upward pattern. This happens due to the effect of the positive intercept term (α =10) and the high persistence factor (ρ=0.9). The influence of random noise is dampened by the presence of the autoregressive structure, resulting in the observed trajectory. The two price trajectories are now summarized in terms of their first two moments.

For the RW model, while the mean that is the expected price, remains unchanged at the fixed initial price ($P_0 = 100$), the variance however increases proportionally with the passage of time. Specifically variance at time t, $V(P_t) = \text{t}.\sigma^2$. With a value of σ being equal to 2, variance thus becomes 100, 200 & 400 respectively at t=25, 50 and 100.

The AR (1) model (Fig. 2) however behaves completely differently. With α=10, ρ=0.9, the process has a long-run mean of: $\mu = \frac{\alpha}{1-\rho} = \frac{10}{1-0.9} = 100$. It may be noted that the initial price is also 100. That is, the long-run equilibrium price is equal to the initial price of the stock. The process starting from the initial equilibrium level thus hovers around the same. This is due to the fact that the fluctuations that occur constantly are characterized by zero mean implying that positive shocks cancel negative ones. As a consequence of this, the expected mean price of the process remains stable over time unlike in a trending process. The main point of departure between a RW and AR(1) process is the magnitude of variance which increases rather slowly and non-linearly due to the partial absorption and smoothening of random shocks because of the presence of memory in an AR(1) process.

This behavior is corroborated by the Simulation exercise implemented in this paper. It is observed that even after 100 periods, the dispersion across the paths is much narrower in an AR (1) process as against the explosive divergence observed in the RW trajectories. The phenomenon of *persistence* embedded in the AR process gives rise to clustering around the long-run average revealing the effect of memory in the Data Generating Process (DGP) in achieving greater short-term predictability and reduced volatility.

V. Discussion

The *mean-reverting* behaviour of asset prices as observed under the AR (1) DGP —where they tend to return to their average over time—can be understood through two key investors' biases: *representative bias* and *conservative bias*. The former causes investors to overreact to small or recent pieces of information, while the latter makes them slow in updating their beliefs as new information comes in. Barberis et al. (1998) explain that these biases result in market *over-reactions* and *under-reactions*. Additionally, Daniel et al. (1998) offer further insight, suggesting that investor behaviour is also shaped by how they process information. They argue that overconfident investors, who believe their successes are due to their own abilities and their failures are due to external factors, tend to rely heavily on their private information and ignore broader public information. This can lead to *mispriced stocks* and *distorted market dynamics* (Karmakar, 2024).

The idea of *market efficiency* holds major implications as far as predictability of stock prices is concerned. This becomes particularly so when random walk process as a specific realization is considered. In an *efficient market*, stock prices quickly incorporate all available information, meaning that any new data or news is reflected in the price almost instantaneously. This creates a framework where prices exhibit random walk characteristics: just like a random walk, with each step being independent of the last, stock prices move in a way that follows no definite pattern, making it impossible to predict the future value of a stock based on its past movements. The randomness in movements makes a Random Walk series inherently unpredictable. In this context, the expected price at time t, conditioned on all past information, is given by the equation: $E[P_t \mid P_{t-1}, P_{t-2}, \dots \dots] = P_{t-1}$ As the movements of prices do follow a random pattern, past prices fail to have any impact on the present implying that the best predictor of tomorrow's price is simply today's price. Under the assumption that the markets are efficient, prices already reflect all available information meaning that future price movements are determined by new, unforeseen information alone, reiterating the notion of unpredictability that is unique to stock prices. The root of this unpredictability lies in the essential randomness in the effect of any new information which may come into the system, be it economic or otherwise, implying that past will have little or no effect on the present. As prices reflect new information almost instantaneously in an efficient market, there is no room for investors to exploit historical data for future gains. This invalidates strategies based on technical analysis or historical price patterns highlighting the challenges investors face when attempting to forecast future price movements (Karmakar, 2025).

Dealing with AR (1) model is suitable for Traders and Investors for forecasting near-future prices on the basis of short-term momentum. At the same time, risk managers should account for the effects of this persistence when estimating volatility and assessing risk over different time horizons[3].

VI. Conclusion

This study tries to look into the behavior of stock prices from an empirical point of view. Depending on the underlying structural parameters, i.e., the Data Generating Process (DGP), a stock series may be considered as *stationary* that is one whose statistical characteristics remain unchanged over time or as *non-stationary* which is characterized by time-varying characteristics. While theory and intuition suggests that the *stationary time series* is mean-reverting and hence predictable, the *non-stationary* one is not mean-reverting, rather having a tendency to divert away from mean. It is not possible to do any predictive analysis from a non-stationary DGP. The specific non-stationary DGP considered in this study is random walk (RW) model and the specific stationary DGP considered in this study is the AR (1) model. The novelty of the present study is that it implements a *simulation* study to observe the theoretical notions of predictability or un-predictability from an empirical point of view. By artificially generating sample data set characterized by a stationary or non-stationary DGP, the study demonstrates how the price dynamics behave completely differently under the two scenarios. Starting from the same initial point one observes that the end points are quite different under the different model assumptions. The end points of the non-stationary random walk process have a much wider variability in contrast to the stationary AR(1) DGP resulting in some sort of *convergence* among the price trajectories. In fine, the Random Walk model emphasizes unpredictability and divergence, embodying the Efficient Market Hypothesis, while the AR (1) model reveals persistence and clustering, suggesting some short-term predictability. This paper offers a foundational step that bridges theory with intuitive, visual exploration[4]. The simulation results provide a clear visual and statistical narrative: Starting from the same initial position does not imply that the end points will always be the same. Depending on the underlying structure, final outcomes may either converge to some extent or diverge significantly. This divergence or convergence fundamentally reflects the nature of predictability in financial markets and underscores the crucial role that model structure plays in shaping price behavior.

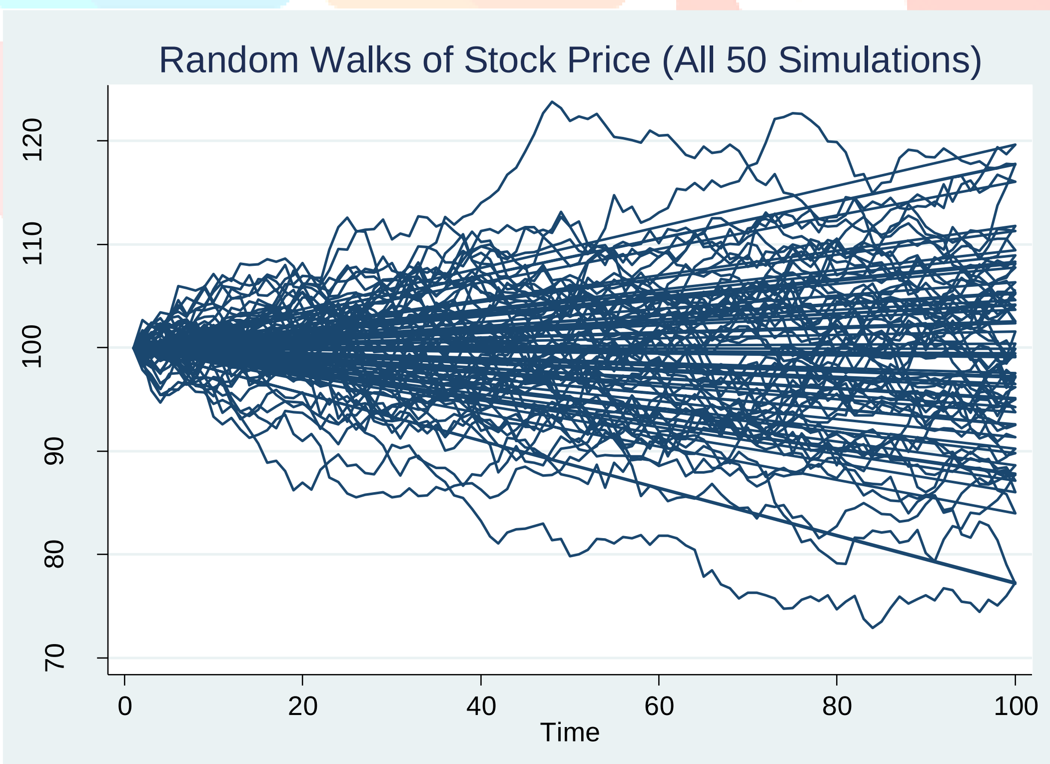
[3] The limitation of the present analysis is that it has considered a simple first order autoregressive structure to investigate into the nature of predictability. Real markets exhibit more complex phenomena such as volatility clustering, jumps, and regime shifts. Future research can extend this work by integrating: GARCH models to capture changing volatility, ARIMA models to account for trends and seasonality, State-space models for latent influences.

[4] See Tsay (2010), Box, Jenkins, Reinsel & Ljung (2015) & Hull (2015).

References

1. Barberis, N., Shleifer, A., &Vishny, R. W. (1998). A model of investor sentiment. Journal of Financial Economics, 49(3), 307-343.
2. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: Forecasting and control (5th ed.). Wiley.
3. Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. Journal of Finance, 53(6), 1839-1885.
4. Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383–417. https://doi.org/10.2307/2325486
5. Hamilton, J. D. (1994). Time series analysis. Princeton University Press.
6. Hull, J. C. (2015). Options, futures, and other derivatives (9th ed.). Pearson.
7. Karmakar, P. (2024). Predictability of stock prices in India: Reconciling the efficient market hypothesis with technical analysis. International Journal of All Research Education and Scientific Methods, 12(9), 147–158.
8. Karmakar, P. (2025). Stock price fluctuations: Understanding the forces at play. International Journal of Creative Research Thoughts, 13(1), 2320–2882.
9. Tsay, R. S. (2010). Analysis of financial time series (3rd ed.). Wiley.

- **Figure 1:** Fifty Random Walk simulated price paths over 100 periods, starting at 100, showing wide divergence.



- **Figure 2:** First five AR (1) simulated price paths over 100 periods, starting at 100, showing smooth, persistent trends.