SHARE REPURCHASE STRATEGY OPTIMIZATION WITH GENETIC ALGORITHMS ON GEOMETRIC BROWNIAN MOTION SIMULATED PRICE DATA

Ze Kai Lin, Zelin Xu, Charles Wang
Columbia University
New York
{zl3359, zx2409, clw2198}@columbia.edu

1 Introduction

The concept of a buyback purchase involves a corporation repurchasing its own shares, a strategic action that carries profound implications for both the company's future and the broader stock market landscape. In groundbreaking research undertaken by Osterrieder and Seigne[1], this endeavor is characterized as a "free lunch" phenomenon, positing that the skillful execution of such buyback operations can surpass existing benchmarks and yield tangible benefits for both the company and its shareholders. Building upon the foundation laid by Osterrieder and Seigne, our research delves further into the realm of optimal buyback scheduling, leveraging the efficacy of Genetic Algorithm as a computational tool. Our investigation substantiates the effectiveness of Genetic Algorithm in generating optimal solutions, notably surpassing conventional benchmarks such as the simple average price method. Moreover, we undertake a meticulous examination of these optimal solutions in various market conditions characterized by distinct levels of volatility and drift. Importantly, our findings underscore the enhanced efficacy conferred by temporal optionality in optimizing buyback strategies.

2 Stock Price Simulation using Geometric Brownian Motion

Embarking on this research project, we have explored the intricacies of stock market dynamics through the lens of Geometric Brownian Motion (GBM). To harness the market's complexity, we have incorporated a trio of sophisticated models, each shedding light on distinct facets of volatility, market regimes, and abrupt jumps in price dynamics.

GBM assumes that the stock price S_t follows a log-normal distribution, which can be represented by the stochastic differential equation:

$$S_t = e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)}$$

where:
S_t is the stock price at time t
μ is the expected return
σ is the volatility of the stock's return
W_t is a Brownian motion

The first model, simple GBM with varying volatility [1], incorporates each price path with different volatility that captures the potential of all possible market conditions. This model allows for a more general view of how might a stock perform under different fluctuating scenarios.

The Regime-Switching Model extends the simple GBM framework by allowing for different 'regimes' or states of the market. Each regime represents a different set of parameters (such as drift and volatility) that govern the behavior of stock prices. The model simulates stock prices by switching between these regimes based on a user-defined or Markov process. This method is particularly useful in capturing the bull and bear phases in the stock markets, as well as other phenomena like market crashes or hyperinflation periods.

Lastly, the Merton Jump Diffusion Model adds sudden [2], discrete events to the continuous GBM process governed by Poisson Distribution. These events, or 'jumps,' can represent unforeseen information arriving in the market, such as earnings announcements, economic news, or other market shocks.

3 Genetic Algorithm

Our genetic algorithm (GA) is designed to identify superior trading strategies through an evolutionary process. This process commences with the amalgamation of optimal parent solutions to yield a subsequent generation of potentially superior solutions. A stochastic mutation is incorporated to circumvent local optima, ensuring the diversification of our solution space.

The foundation of our GA's implementation is rooted in the methodologies espoused within seminal research papers. Our primary optimization target, which doubles as the fitness function for each trading strategy, is the difference between the final day's Volume Weighted Average Price (VWAP) and a predetermined benchmark.

To illustrate, we may consider a hypothetical benchmark such as the arithmetic mean of daily VWAPs. Here, the optimization objective is to maximize the discrepancy between this benchmark and the VWAP.

Alternatively, we can express the fitness value as the differential between the total benchmark prices and the total prices actually paid by traders, a method paralleled in our research.

Our empirical data suggests that these two objective formulations are congruent in terms of their resultant outputs within our study.

3.1 Implementation Specifics

We have integrated a standard GA model, incorporating an impact cost into our objective function[3]. Prior to the inclusion of this cost, its coefficient can be ascertained through empirical analysis. Additionally, our model incorporates a variable trading duration parameter, which imbues traders with temporal flexibility[4].

The GA software extends the capability for users to select from various crossover and mutation techniques. These include one-point, two-point, and uniform crossover methods, as well as polynomial and uniform mutation[3]. A Gaussian mutation has also been introduced as an alternative option. The program autonomously selects a tournament size 'k' of strategies for each evolutionary cycle, subsequently applying the chosen crossover and mutation processes.

Initially, our program was challenged by its incapacity to produce strategies with zero trading volume, an issue attributable to the normalization process post-mutation. To resolve this, we have implemented a feature that allows users to assign a probability to the mutation of trading volumes to zero on specific days. This feature is designed to be robust against renormalization, as it is programmed to normalize only the non-selected days.

3.2 Parameter Optimization

Through observation, we found that a GA utilizing one-point crossover in tandem with polynomial mutation converges most expeditiously to the optimal solution. We have set the tournament size at 3 and the number of generations at 60. It was noted that beyond this threshold, additional generations resulted in an inconsequential improvement in fitness value and incurred excessive computational resources.

4 Results

4.1 GA Optimal Strategies

Consistent with the findings of Osterrieder and Seigne [5], our results corroborate the effectiveness of the Genetic Algorithm (GA) in devising optimal trading strategies. The GA's remarkable capability is showcased under varying market conditions, such as zero drift environments with annual volatilities of 10%, 20% and 30%. Our study aims to not only confirm the GA's efficiency but also to provide a granular understanding of its operational mechanics in diverse market scenarios.

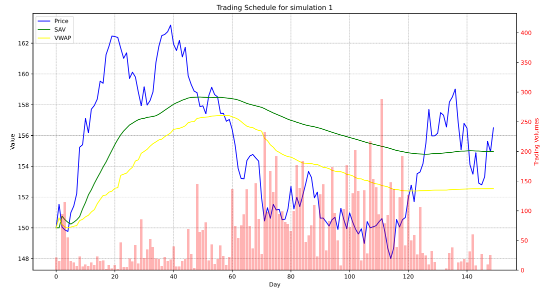

Figure 1: optimal strategy on simulated path of annual volatility 10% and 0 drift

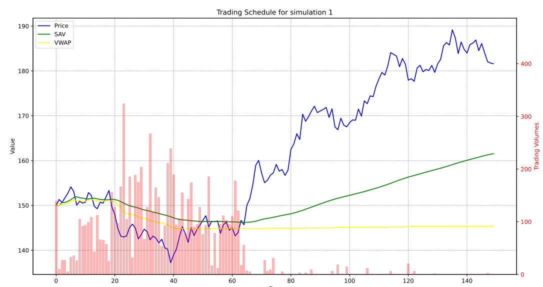

Figure 2: optimal strategy on simulated path of annual volatility 20% and 0 drift

In the first scenario, with an annual volatility of 10%, the GA demonstrated a conservative strategy. The trading schedules indicated a tendency for the algorithm to engage in buying activities when prices dipped below the daily Volume Weighted Average Price (VWAP) benchmark. This cautious approach aimed at minimizing risk while capitalizing on lower prices to maximize returns. The graph in Figure 1 illustrates this behavior, where repurchase activities clustered around these lower price points.

As the volatility increased to 20%, the GA's strategy evolved. It began to exhibit more frequent trading activities, capitalizing on short-term price fluctuations. The algorithm's ability to quickly adapt to these conditions was evident in the higher density of trade executions, as seen in Figure 2. This strategy appeared to balance the need for risk mitigation with the potential for higher returns in a more volatile market.

At the highest volatility level of 30%, the GA displayed its most dynamic and aggressive trading strategy. The increase in volatility led to more significant price swings, which the GA leveraged to execute trades at optimal points. The trading schedule for this scenario, depicted in Figure 3, showed a marked increase in trading frequency and a strategic approach to exploit large price movements for higher gains.

In scenarios characterized by zero drift, the schedules derived from the Genetic Algorithm (GA) exhibit consistent alignment with theoretical expectations. In environments marked by increased volatility, stock prices exhibit greater fluctuations. Consequently, the GA demonstrates an enhanced capacity to identify and exploit periods of lower pricing. This phenomenon is evidenced by the observation that schedules generated under conditions of higher volatility tend to cluster more densely around periods when prices are at their lowest.

In our research, we further conducted a thorough evaluation of the Genetic Algorithm's (GA) efficacy across simulated trajectories characterized by varying degrees of drift. Anticipating that the drift would significantly influence the temporal clustering of trading activities, we formulated specific hypotheses for each drift scenario. For paths exhibiting positive drift, we hypothesized an intensified concentration of trading activities in the initial stages, under the assumption that prices would escalate subsequently. Conversely, in the context of negative drift, we predicted that trading would predominantly occur in the later stages. In scenarios with zero drift, we expected a more uniform distribution of trading throughout the entire execution period. However, we acknowledged that even in these circumstances, clustering might still occur, contingent upon the stochastic nature of price movements.

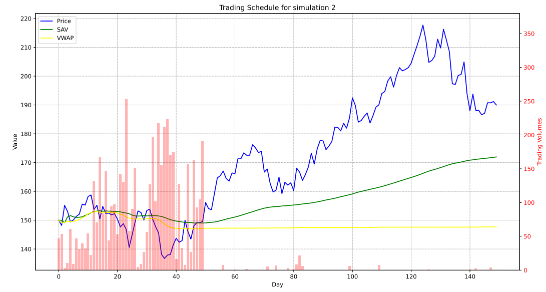

Figure 3: optimal strategy on simulated path of annual volatility 30% and 0 drift

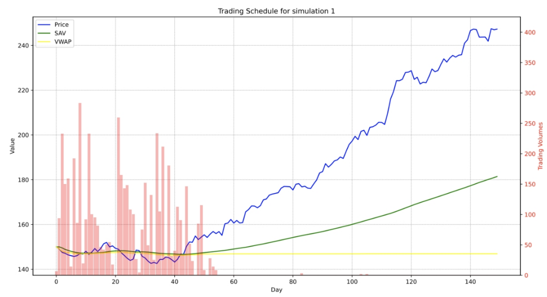

Figure 4: optimal strategy on simulated path of annual volatility 15% and 0.005 drift

Our results largely affirm the hypotheses previously formulated. As evidenced in Figure 4 and Figure 5, the majority of simulations featuring positive drift paths demonstrate a propensity for trading to occur predominantly in the initial stages, in line with our expectations. However, the scenario with negative drift paths, as shown in Figure 6, presents an interesting divergence. While we did not observe a distinct clustering of trades towards the end of the period, the volume of trading in these later stages was notably higher. This outcome appears to challenge the expected symmetry between the two scenarios. We hypothesize that this anomaly may stem from the specific formulation of our optimization problem or the operational characteristics of our Genetic Algorithm (GA) implementation. We intend to delve deeper into these aspects in subsequent sections, aiming to unravel the complexities underlying this phenomenon.

We also observed that the trading behaviors identified in scenarios with varying levels of annual volatility were similarly manifested in scenarios featuring non-zero drift. This consistency suggests that the influence of volatility on trading patterns is not exclusively dependent on the presence or absence of drift, indicating a more complex interplay between volatility and drift in shaping trading behaviors. This finding underscores the need to consider both volatility and drift as influential factors in the analysis of trading strategies and market dynamics.

It's important to highlight that, based on our empirical research, we determined a drift value of 0.005 as an effective parameter for testing. When the drift value was significantly higher than this point, it resulted in an overly rapid increase in prices, subsequently leading to the underperformance or failure of our Genetic Algorithm. Conversely, when the drift was substantially lower than 0.005, the clustering effect became too subtle to be effectively analyzed. These findings reinforce the significance of empirically identifying an optimal combination of volatility and drift. Such a balanced approach is crucial to ensure the effectiveness of the GA and to reliably capture and interpret the nuances of market dynamics.

Across all scenarios, the GA consistently demonstrate its proficiency in timing purchases at lower prices, thereby maximizing the differential between the simple average on the final day and the daily VWAP. The algorithm's ability to identify and act on these price differentials was a testament to its design and optimization capabilities. The emphasis on timing and price sensitivity in the GA's strategy is a crucial aspect of its success.

Our study focused on evaluating the performance of our Genetic Algorithm (GA) optimized strategies against a rudimentary benchmark – the simple average price, referred to as the "bogus benchmark." named by Osterrieder and Seigne [5]. We observed that the distribution of outperformance, quantified in basis points, across different drift and volatility scenarios appears to align closely with a log-normal distribution, characterized by distinct sets of parameters.

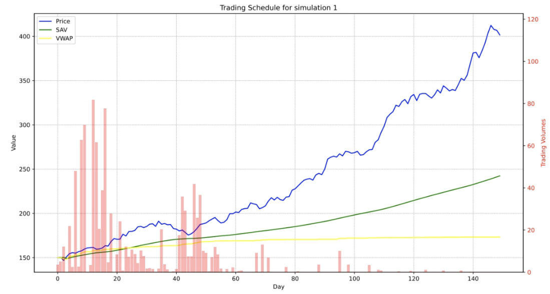

Figure 5: optimal strategy on simulated path of annual volatility 25% and 0.005 drift

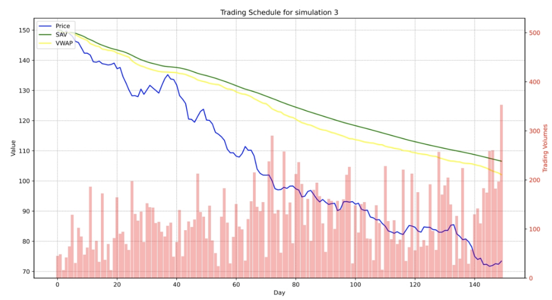

Figure 6: optimal strategy on simulated path of annual volatility 20% and -0.005 drift

We present the outperformance data for scenarios with an annual volatility of 10%, but varying drifts. Notably, these distributions exhibit variations in skewness and mean. In the zero drift scenario in Figure 9, the outperformance distribution is denser in higher value regions, suggesting a greater likelihood of achieving low levels of outperformance. In contrast, scenarios with non-zero drift in Figure 7, Figure 8 and show a more symmetrical distribution around the mean. Significantly, in the positive drift scenario, the magnitude of overall outperformance is markedly higher compared to the other scenarios. This aligns with our earlier observations that the GA is particularly adept at identifying lower-price opportunities in positive drift environments, thus leveraging this advantage to surpass the benchmark.

Furthermore, similar patterns in the interplay between drift and outperformance distribution are observable across different levels of volatility. However, higher volatility scenarios, as illustrated in Figure 10, correspond to a more dispersed distribution, indicating a generally higher magnitude of outperformance. Figures 11, 13, and 12 elucidate the relationship between volatility and mean outperformance when the drift is held constant. These findings suggest a positive linear relationship, wherein higher volatility is associated with increased mean outperformances. This observation is congruent with our previous findings that greater volatility, entailing more pronounced price fluctuations, enables the GA to capitalize on trading at lower prices, thereby enhancing performance.

Nonetheless, the scenario with positive drift presents a more erratic relationship, diverging from our expectations. This anomaly might be attributable to the limitations of our GA program, or it might represent an intriguing area for further investigation in future research. This discrepancy underscores the complexity and nuanced nature of market dynamics.

4.2 Temporal Optionality

To emulate the discretionary aspect of trade duration exercised by traders, we incorporated temporal optionality into the GA. This was achieved by allowing the algorithm to select the trading end day within a range of percentage of the total trading period (T) for each individual strategy as suggested by Osterrieder and Seigne[4]. Our anticipation was that the evolutionary process would identify the most advantageous duration.

This expectation is corroborated by our empirical findings in scenarios characterized by varying drift rates under a constant volatility regime. Illustrated in Figure 14, the case of zero drift yields a more dispersed duration distribution. This dispersion arises due to the equal probability of lower asset prices occurring at any given point in time. Conversely, Figure 15 elucidates the scenario with positive drift, wherein the optimal execution time predominantly congregates

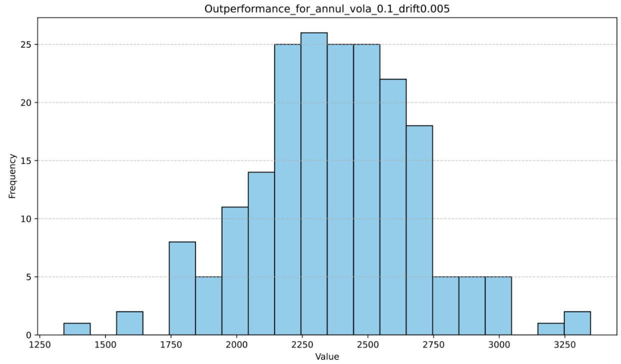

Figure 7: Distribution of outperformance using GA over simple average benchmark with annual volatility 10% and 0.005 drift

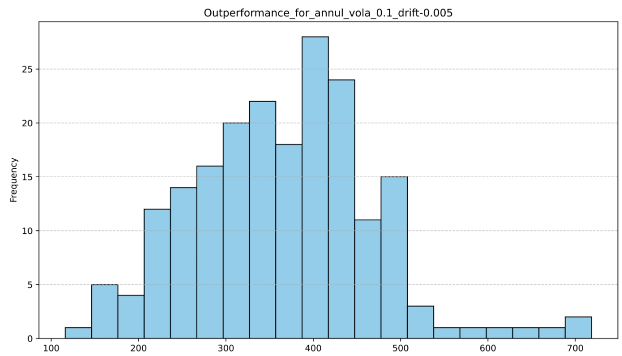

Figure 8: Distribution of outperformance using GA over simple average benchmark with annual volatility 10% and -0.005 drift

around earlier intervals. This observation aligns with intuitive expectations, given that asset prices are initially lower, thereby incentivizing maximal procurement in the nascent stages and consequently precipitating an earlier termination of the purchasing phase. In stark contrast, as depicted in Figure 16, a negative drift scenario forecasts a decline in asset prices. In this context, the strategically optimal approach would be to defer acquisition activities until the latter part of the available time frame, thereby capitalizing on the anticipated lower price points towards the end of the period.

In the advancement of our research on temporal optionality, we refined the concept of 'buyback duration' to denote the most constrained time interval within which a predetermined threshold of asset acquisition, such as 95% of shares, is achieved. In this framework, the Genetic Algorithm is afforded the flexibility to select both the commencement and termination dates of this period. In a parallel but distinct experimental setup, we alter the parameters by granting the GA the liberty solely to determine the end date, while the start date is held constant. This modification allows for the exploration of the impact of a fixed initiation point on the strategy and efficacy of the GA in achieving the defined threshold within the constraints of the given temporal window, similar to what we have discussed in the previous paragraph.

Our investigation further delved into the influence of volatility on the optimal duration of share buybacks. Figures 17, 18, and 19 illustrate the interplay between volatility and the mean duration of buybacks, with the drift rate maintained at constant values. The findings reveal a nuanced relationship between volatility and mean duration, contingent upon the drift rate, aligning partially with our initial hypotheses.

In scenarios where the drift is zero, we observe an absence of correlation between volatility and duration. This phenomenon can be attributed to the challenges faced by the Genetic Algorithm in identifying consistently favorable trading periods within a volatile environment.

Conversely, under positive drift conditions, an intriguing trend emerges: as volatility escalates, the mean duration of the buyback decreases. This can be rationalized by considering that a higher positive drift suggests an upward trend in prices. Consequently, increased volatility in this scenario provides a cushion, allowing for a safer and earlier termination of buyback operations.

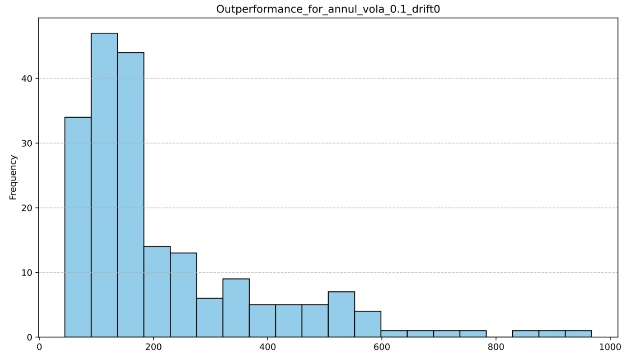

Figure 9: Distribution of outperformance using GA over simple average benchmark with annual volatility 10% and 0 drift

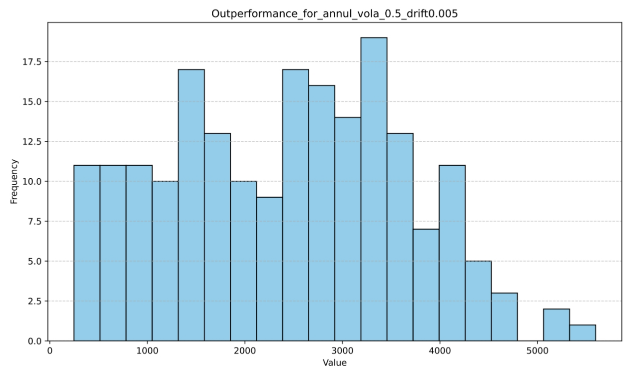

Figure 10: Distribution of outperformance using GA over simple average benchmark with annual volatility 50% and 0.005 drift

In stark contrast, when the drift is negative, the duration of buybacks tends to increase with rising volatility. This is in line with our expectations, as negative drift implies a downward price trend. In such circumstances, greater volatility necessitates a more cautious approach, extending the duration to capitalize on potentially lower prices towards the end of the period.

Overall, these observations underscore the complex dynamics between market volatility, price drift, and the strategic timing of share buybacks, particularly under varying market conditions.

5 Evaluation

To test the validity, consistency, and performance of the optimal trading schedule generated by the genetic algorithm, we use several key performance metrics, which are solution quality, convergence speed, diversity maintenance, computational cost, shortfall, and realized volatility.

1. **Solution Quality:** This metric measures how accurately the algorithm models the optimal solution, and is quantified by the value of the fitness function J(x) at the end of the GA run [6]. A higher solution quality value, which is the difference between the SAV and the actual VWAP paid, indicates a better performance, and is calculated with the formula:

$$\text{Solution Quality} = J(x) = \frac{1}{T}\sum_{t=1}^{T}\text{VWAP}_t - \sum_{t=1}^{T} x_t \cdot \text{VWAP}_t$$

2. **Convergence Speed:** This metric measures how fast the GA reaches the optimal solution. This is usually calculated through the number of generations needed to reach a satisfactory solution. A faster convergence is associated with a smaller number of generations [6].
3. **Diversity Maintenance:** This metric is measures the diversity of the buyback period's diversity in the population. This is important to avoid premature convergence to local optima. The way in which we calculate

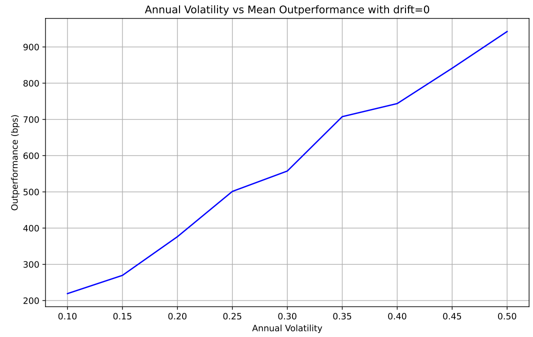

Figure 11: Annual volatility vs mean outperformance with drift of 0

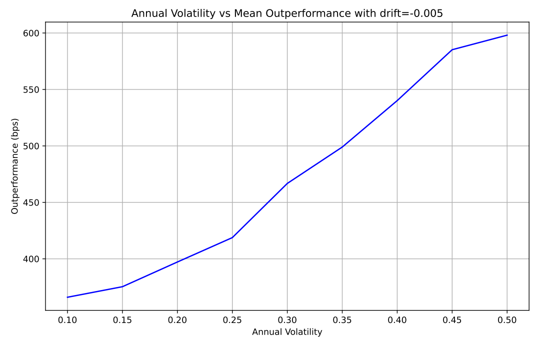

Figure 12: Annual volatility vs mean outperformance with drift of -0.005

the diversity is through the variance of fitness values in the populations each generation g, represented as

$$\text{Variance} = \frac{1}{N}\sum_{i=1}^{N}(J(x^i) - \overline{J})^2$$

4. **Computational Cost:** This measures the computational resources consumed by the GA, typically in terms of runtime. In the case of high-frequency trading schedules, the computational cost becomes critical, making it a valuable performance metric [6].
5. **Shortfall:** This metric compares the GA-based strategy and the benchmark strategy by calculating the shortfall for each trading day and the entire buyback period and taking the average of each day throughout the buyback period. The shortfall with respect to the benchmark (SF) can be calculated for both the GA-based strategy and the benchmark strategy. For the ith trading day, the shortfall is defined as:

$$SF_i = \frac{\sum_{t=1}^{T} x_t^i P_t^i - V_i \times \overline{P}_i}{V_i \times \overline{P}_i}$$

where x_i^t and P_i^t are the volume of shares traded and the trade price at time t on day i, respectively, V_i is the total volume of shares traded on day i, and $\overline{P}_i$ is the VWAP price on day i [6]. As such, the shortfall for the entire buyback period can be calculated as:

$$\text{SF} = \frac{\sum_{i=1}^{n} SF_i}{n}$$

6. **Realized volatility:** Lastly, we use realized volatility to determine the risk of the trading strategy, which we calculated using:

$$RV = \sqrt{\sum_{i=1}^{n}\left(\frac{\sum_{t=1}^{T} x_t^i (P_t^i - \overline{P}_i)^2}{V_i}\right)}$$

Figure 13: Annual volatility vs mean outperformance with drift of 0.005

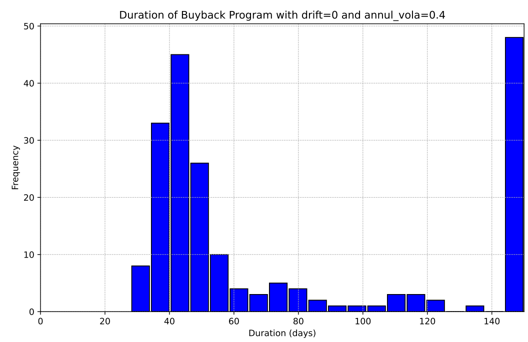

Figure 14: Purchase duration for annual 40% annual volatility and 0 drift

5.1 Leveraging Evaluation Functions

In future research, the evaluation functions devised for optimal buyback scheduling using Genetic Algorithms (GAs) offer a comprehensive framework for refining algorithmic components and parameters, integrating advanced computational techniques to mitigate Look-Ahead Bias, and dynamically adapting to changing market conditions. Researchers can conduct sensitivity analyses, explore market impact, and compare GA-based strategies with alternative methods, considering diverse economic scenarios for a more comprehensive understanding. Assessing the algorithm's long-term stability, usability, and ensuring transparent and reproducible research practices contribute to the ongoing development of robust GA strategies for buyback scheduling in real-world financial contexts.

6 Conclusion

In conclusion, this study has elucidated the substantial potential and versatility of Genetic Algorithms (GAs) in the formulation of optimal trading strategies under varying market conditions. Our findings are in harmony with those of Osterrieder and Seigne, accentuating the proficiency of GAs in capitalizing on market volatility for the timing of trades and optimization of returns. The demonstrated adaptability of GAs at diverse levels of volatility not only reinforces their efficacy in theoretical constructs but also heralds their applicability in practical trading scenarios. This research transcends the confines of academia, proffering tangible insights for traders and financial analysts within the ever-evolving landscape of finance.

Nevertheless, a salient limitation emerges in the form of Look-Ahead Bias. This bias, a manifestation of retrospective insight, can inadvertently permeate the model, culminating in over-optimistic prognostications. It arises when the algorithm gains access to data which, in a real-world context, would have been unavailable at the time of trade execution. To ameliorate this limitation and augment future research endeavors, we advocate for the integration of advanced computational techniques, such as greedy algorithms and neural networks. These methodologies possess the capability to incessantly assimilate stock trends and adapt in accordance with emerging data, thereby aligning the model's predictive capacity more closely with real-time market dynamics and diminishing the propensity for Look-Ahead Bias.

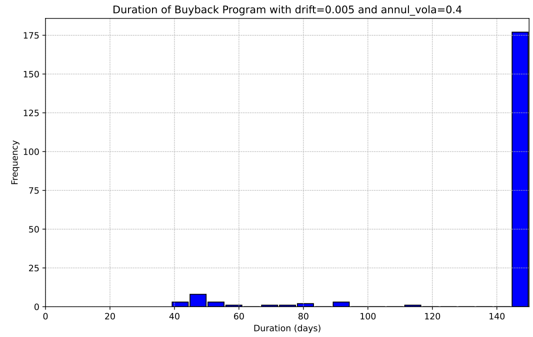

Figure 15: Purchase duration for annual 40% annual volatility and 0.005 drift

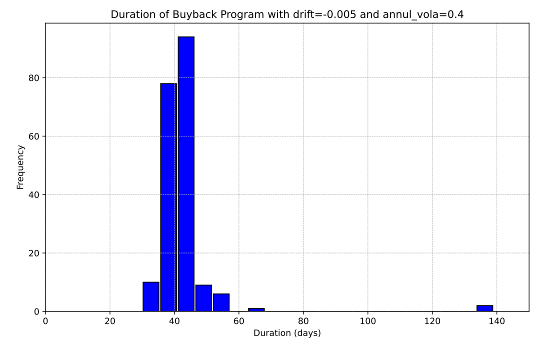

Figure 16: Purchase duration for annual 40% annual volatility and -0.005 drift

Additionally, the synthesis of deep learning methodologies with traditional GAs is proposed, aiming to cultivate a more robust hybrid model. This model would amalgamate the strategic adaptability inherent in GAs with the predictive acumen of deep learning, yielding a comprehensive analytical tool for financial market examination. Such a hybrid construct has the potential to revolutionize trading strategy development, endowing it with a more nuanced and dynamic character, well-suited for navigating the complexities of financial markets.

In summation, while GAs have demonstrated remarkable potential in our study, the path forward entails an integration with deep learning techniques to surmount limitations such as Look-Ahead Bias. This integrated approach harbors the promise of enhancing the efficacy, precision, and adaptability of trading strategies, paving the avenue for more refined and efficacious market predictions.

References

[1] Joerg Osterrieder and Michael Seigne. A free lunch for share buybacks. *SSRN Electronic Journal*, 2023.

[2] Robert C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1):125–144, 1976.

[3] Joerg Osterrieder and Michael Seigne. Navigating share buyback programs: A genetic algorithm approach to outperform the buyback benchmark. *SSRN Electronic Journal*, 2023.

[4] Joerg Osterrieder and Michael Seigne. Rethinking share buyback execution: Insights into temporal optionality and empirical anomalies. *SSRN Electronic Journal*, 2023.

[5] Joerg Osterrieder and Michael Seigne. The great deception: A comprehensive study of execution strategies in corporate share buy-backs. *SSRN Electronic Journal*, 2023.

[6] Joerg Osterrieder and Michael Seigne. Navigating share buyback programs: A genetic algorithm approach to outperform the buyback benchmark. *SSRN Electronic Journal*, 2023.

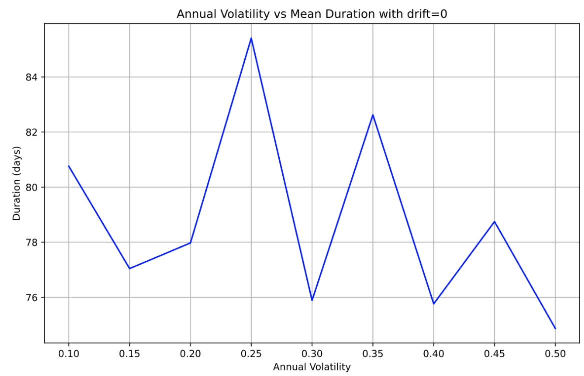

Figure 17: Annual Volatility vs Mean Duration with 0 drift

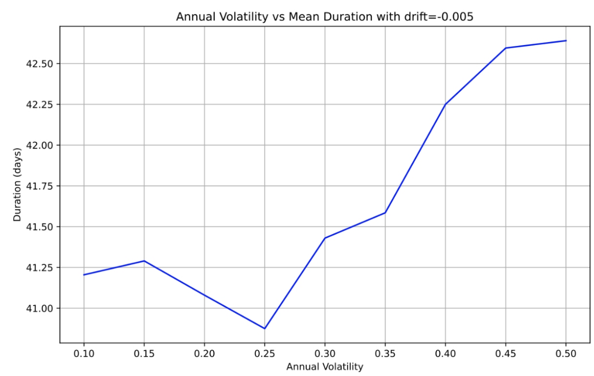

Figure 18: Annual Volatility vs Mean Duration with -0.005 drift

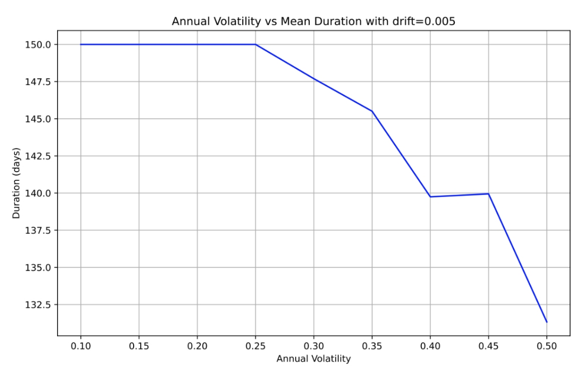

Figure 19: Annual Volatility vs Mean Duration with 0.005 drift