Modeling Stonk[1] Prices Using Stochastic Differential Equations

Alan Bruner, Evan Cochrane, Theodore Guetig, Ben Wilfong

1 Introduction

The stonk market is a system consisting of buyers and sellers that share ownership of a business. Sellers are the individuals who manage the business, while buyers purchase portions of the business, known as stonks, in hopes of making a profit. Futures are contracts that oblige buyers to buy a fixed amount of a commodity or currency and sell at a later date, and function similarly to stonks for our purposes. The return for buyers depends on a variety of factors that are internal and external to the market. To increase the probability of profit, we use a stochastic differential equation model to predict future stonk prices based on the historical trends of 60 businesses and commodities.

2 Model

We model the price of stonks as a stochastic differential equation whose price is dependent on a random variable $z(t)$ as a function of time. This differential equation is [1]

$$\mathrm{d}P = \mu P \mathrm{d}t + \sigma P \mathrm{d}z. \tag{1}$$

Where μ represents the drift, or general trend of a stonk, and σ represents the volatility of a stonk. If a stonk has no volatility, $\sigma = 0$, then the differential equation simplifies to $\mathrm{d}P = \mu P \mathrm{d}t$, which has the solution

$$P(t) = P(0)e^{\mu t}.$$

Thus, the model for a stonk's price is a continually compounding interest model with an added volatility parameter. The random variable z is assumed to be a Weiner process, satisfying

- $z(0) = 0$,
- $z(t_1)$ and $z(t_2)$ are independent random variables for any $t_1 \neq t_2$,
- $z(t)$ is continuous, and
- $z(t_1) - z(t_2) \overset{\$}{\leftarrow} \mathcal{N}(0, t_1 - t_2)$.[2]

[1] We define a *stonk* as a stock, commodity, or a cryptocurrency.

[2] The notation $\overset{\$}{\leftarrow}$ denotes selecting a value from a probablilty distribution.

The last condition implies $z(t+\Delta t) - z(t) = \sqrt{\Delta t}\phi$, where ϕ is a standard normal variable. In differential form,

$$\mathrm{d}z = \sqrt{\mathrm{d}t}\phi.$$

Subsituting into (1) gives

$$\mathrm{d}P = \mu P \mathrm{d}t + \sigma P \sqrt{\mathrm{d}t}\phi. \tag{2}$$

We use Itō 's Lemma to help evaluate the stochastic differentials in (1).

2.1 Itō 's Lemma

Let $F(P,t)$ be twice-differentiable and P be a solution to (1). Then

$$df = \left(\frac{\partial f}{\partial t} + \mu P \frac{\partial f}{\partial P} + \frac{\sigma^2 P^2}{2}\frac{\partial^2 f}{\partial P^2}\right) dt + \sigma P \frac{\partial f}{\partial P} dz. \tag{3}$$

This lemma is used to compute values of df given a particular solution P and function f, provided P and f satisfy (3). For our model, we select

$$f(P,t) = \ln(P).$$

Using Itō 's Lemma,

$$\begin{aligned} d\ln(P) &= \left(\frac{\partial \ln(P)}{\partial t} + \mu P \frac{\partial \ln(P)}{\partial P} + \frac{\sigma^2 P^2}{2}\frac{\partial^2 \ln(P)}{\partial P^2}\right) dt + \sigma P \frac{\partial \ln(P)}{\partial P} dz \\ &= \left(0 + \mu P \left(\frac{1}{P}\right) + \frac{\sigma^2 P^2}{2}\left(\frac{-1}{P^2}\right)\right) dt + \sigma P \left(\frac{1}{P}\right) dz \\ &= \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma dz. \end{aligned}$$

Assuming that z is a Weiner process, the logarithmic ratio of the price at any future time τ is calculated as

$$\begin{aligned} \ln\left(\frac{P(\tau)}{P(0)}\right) = \ln(P(\tau)) - \ln(P(0)) &= \int_0^\tau d\ln(P) \\ &= \int_0^\tau \left(\mu - \frac{\sigma^2}{2}\right) dt + \int_0^\tau \sigma dz \\ &= \left(\mu - \frac{\sigma^2}{2}\right)\tau + \sigma(z(\tau) - z(0)) \\ &= \left(\mu - \frac{\sigma^2}{2}\right)\tau + \sigma\sqrt{\tau}\phi. \end{aligned}$$

Substituting t for τ gives a normally distributed random variable with mean $(\mu - \sigma^2/2t)$ and variance $\sigma^2 t$. When the mean is greater than zero the prices usually increase with time, but a high volatility may still result in a net loss. Solving for P as a function of t yields

$$P(t) = P(0)e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma\sqrt{t}\phi}.$$

2.2 Portfolio Selection

Now that we know how to predict the price of a stonk, the next step in our model is to determine what stonks to invest in. Our goal is to maximize profits, of course, but our model is stochastic, so there is also a chance of loss. Greater profit potentials usually mean greater loss potentials. We use a parameter $\alpha \in [0,1]$ to quantify this trade-off. If $\alpha = 0$, the model focuses entirely on minimizing risk or variance in profit, and if $\alpha = 1$, the model focuses entirely on maximizing mean profit.

If there are n stonks, i, each stonk has price $P_i(t)$, and we refer to the collection of stonks as a *portfolio*. Let $x_i(t)$ denote the number of shares invested in each stonk at time t. Then, the total value of the investment is

$$V(t) = \sum_{i=1}^{n} P_i(t) x_i(t).$$

Let $w_i(t)$ be the percentage or weight invested in stonk i at time t,

$$w_i(t) = \frac{P_i(t) x_i(t)}{V(t)}.$$

If $R_i(t)$ is the return on each stonk per unit dollar invested, the total return per dollar invested is computed as

$$\begin{aligned} R(t) &= \frac{V(t+\Delta t) - V(t)}{V(t)} \\ &= \sum_{i=1}^{n} \frac{(P_i(t+\Delta t) - P_i(t)) x_i(t)}{V(t)} \\ &= \sum_{i=1}^{n} \left(\frac{P_i(t) x_i(t)}{V(t)} \right) \left(\frac{P_i(t+\Delta t) - P_i(t)}{P_i(t)} \right) \\ &= \sum_{i=1}^{n} w_i(t) R_i(t) \end{aligned}$$

From (2), the forward difference approximation of $R_i(t)$ is

$$R_i(t) = \frac{P_i(t+\Delta t) - P_i(t)}{P_i(t)} \approx \mu_i(t)\Delta t + \sigma_i \sqrt{\Delta t}\phi.$$

If we wish to maximize expected profit ($\alpha = 1$), then we ignore the stochastic term and maximize

$$\sum_{i=1}^{n} w_i(t) \mu_i(t) \Delta t = \sum_{i=1}^{n} w_i r_i = \mathbf{r}^T \mathbf{w}.$$

If we instead wish to minimize risk ($\alpha = 0$), we can not use the sum of the variances of each stock i,

$$(w_i\sigma_i\sqrt{\Delta t})^2 = w_i^2\sigma_i^2\Delta t,$$

since the variables P_i are not independent - stonk prices are typically correlated. Instead, we use the sum of the covariances of all pairs of stonks,

$$\sum_{i=1}^{n}\sum_{j=1}^{n}(w_i(t)\sigma_i(t)\sqrt{\Delta t})(w_j(t)\sigma_j(t)\sqrt{\Delta t}) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j C_{ij} = \mathbf{w}^T C\mathbf{w}.$$

For an abitrary α strictly between zero and one, we minimize a convex combination of the variances and negative means

$$\begin{aligned} \min \quad & (1-\alpha)(\mathbf{w}^T C\mathbf{w}) + (-\mathbf{r}^T\mathbf{w}) \\ \text{s.t.} \quad & \sum_{i=1}^{n} w_i = 1 \\ & w_i \geq 0 \qquad\qquad \forall i. \end{aligned}$$

neglecting any stonks that receive a weight less than 0.00001. This is a quadratic convex optimization problem.

2.3 Making Predictions

Once the parameters μ and σ are determined for each stonk, the future prices are predicted discretely via

$$P(t+\Delta t) = P(t) + \mu P(t)\Delta t + \sigma P(t)\sqrt{\Delta t}\phi$$

where ϕ is a standard normal variable. We also consider predictions where correlations in stonk prices are considered. We replace ϕ with w, a a multivariate sample with mean 0 and covariance C by defining

$$w = A^T C$$

where A is some matrix such that $AA^T = C$. We consider two different matrices A for comparision:

- A is the lower triangular matrix of the Cholesky decomposition
- $A = U\sqrt{D}$, where $A = UDU^T$ is the orthogonal eigenvector-eigenvalue decomposition of C.

3 Results

The stonks in our portfolio are listed in Appendix B. We begin optimizing our portfolio by examining the risk versus reward frontiers for various constraints on the optimization model to determine the points of diminishing return. We explored risk versus reward frontiers for three cases:

1. No Restrictions
2. Maximum investment of 20% in any one stonk and no minimum investment
3. Minimum investment of 1% and maximum investment of 20% for each stonk in the portfolio

The risk versus reward frontiers for each of these three cases is shown in Figure 1. The associated risk of diminishing returns for each case are provided in Table 1.

Table 1: Risks of Diminishing Return

Case	Risk (α)
1	0.4974
2	0.4976
3	0.6675

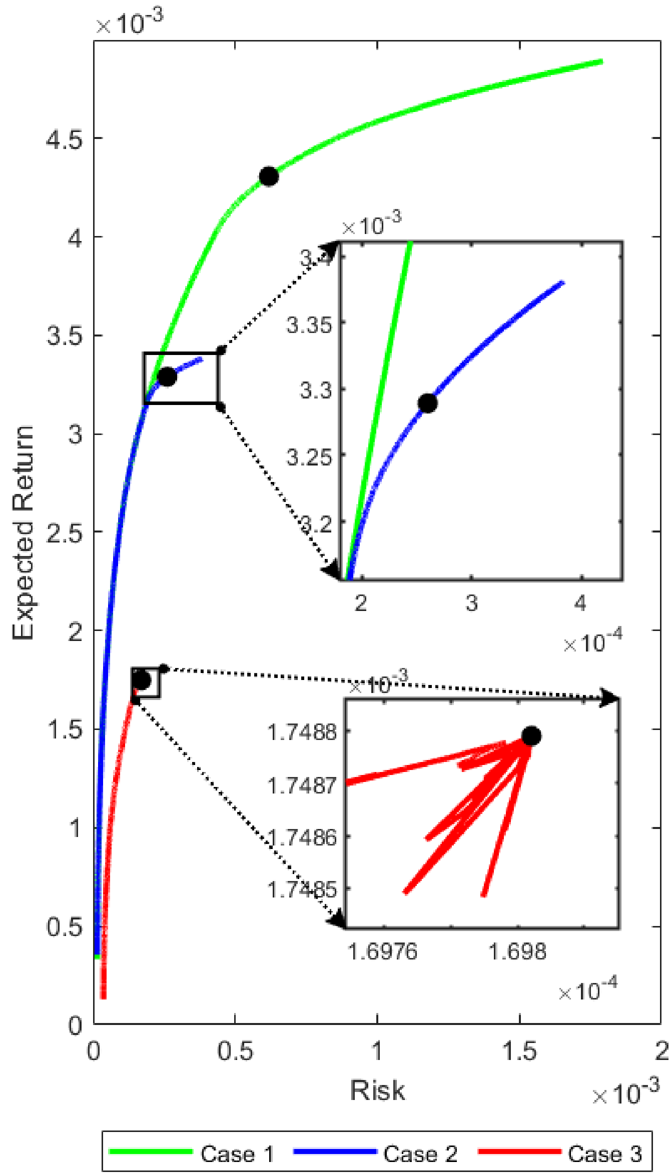

Figure 1: Risk versus Reward Frontiers

It becomes quite clear from Figure 1 that applying constraints limits both expected risk and return. By applying a maximum investment of 20% in any one stonk, the risk frontier is nearly identical to that of the unconstrained frontier until a risk of approximately 2×10^-4. After this, the constrained frontier breaks away from the unconstrained frontier and terminates at a lower maximum estimated risk and return. As expected, applying a lower bound of 1% investment in each stonk decreases expected return, but in the lower bound it surprisingly increases risk slightly over cases one and two. This is most likely due to the forced inclusion of a small subset of high risk stonks due to the 1% minimum investment. Table 1 suggests a sharp increase in risk of diminishing return between cases two and three, but this increase is suspect given the noise apparent in the lower zoomed plot in Figure 1. Despite this, we elect to continue with this risk of diminishing return because this noise appears to be independent of the sampling frequency of risk.

To explore the previously selected model, we utilize the alphas of diminishing return in Table 1 to build portfolios and make six month predictions, both with and without stonk price correlation. In order to make useful comparisons, a random number seed of 05282022 was used in Matlab for all comparative explorations.

To set a baseline return, we run the model with no portfolio restrictions and no price correlation (Case A). This led to a portfolio containing three stonks that made only a few small transactions in the six months predicted. No new stonks were purchased or old stonks sold in their entirety. The monthly returns were -3.27%, 0.66%, -7.98%, -6.38%, -6.53%, and 3.07%. Including a 5% transaction fee on sales yields monthly returns of -3.27%, 0.62%, -8.03%, -6.43%, -6.59%, and 2.75%. The daily trajectories for are shown in Figure 2 and appear to move independently of each other through time.

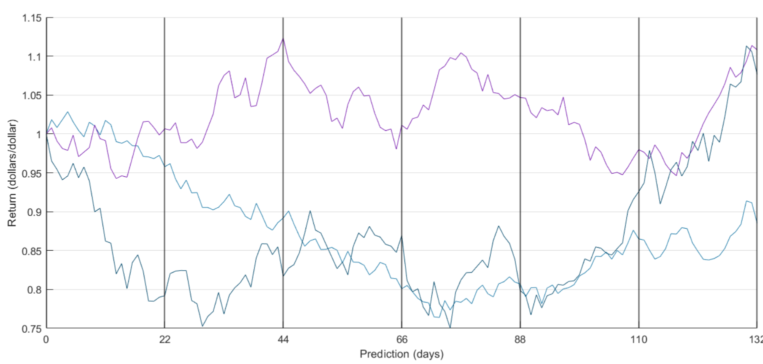

Figure 2: Case A Daily Trends

If we continue with no portfolio restrictions or transaction fees but add stonk price correlation via the lower triangluar matrix of the cholesky decomposition of the correlation matrix, we get the daily response shown in Figure 3 (Case B).

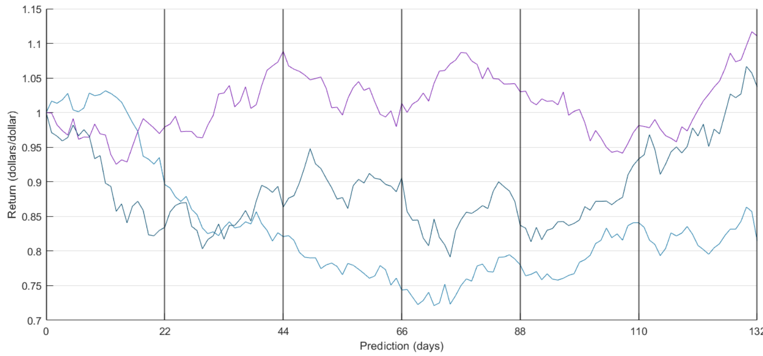

Figure 3: Case B Daily Trends

The model still selects the same three stonks, but their trajectories are slightly different. The monthly returns for this case are slightly lower at -6.65%, -3.34%, -9.56%, -7.80%, -7.16%, and 0.52%. Including a 5% transaction fee slightly reduces these returns to -6.65%, -3.38%, -9.61%, -7.86%, -7.21%, and 0.19%. In both cases, this correlation predicts returns slightly lower than in Case A.

Alternatively, if we define the correlation as $A = U\sqrt{D}$ where $C = UDU^T$ is the orthogonal eigenvector eigenvalue decomposition of the correlation matrix, we get the prediction shown in Figure 4.

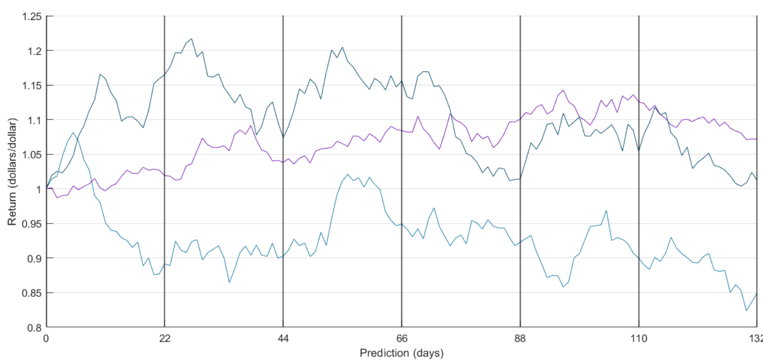

Figure 4: Case C Daily Trends

The returns in this case are -1.59%, -0.86%, 4.31%, 3.04%, 4.31%, and -0.59% without transaction fees and -1.59%, -0.93%, 4.25%, 2.98%, 4.25%, and -0.98% with transaction fees.

Next, we consider the predictions when any one stonk can make up at most 20% of the portfolio. Again, as a baseline, we first examine the case where there is no correlation in stonk prices (Case D). The six month predictions for this case are shown in Figure 5 and appear to move independently of each other.

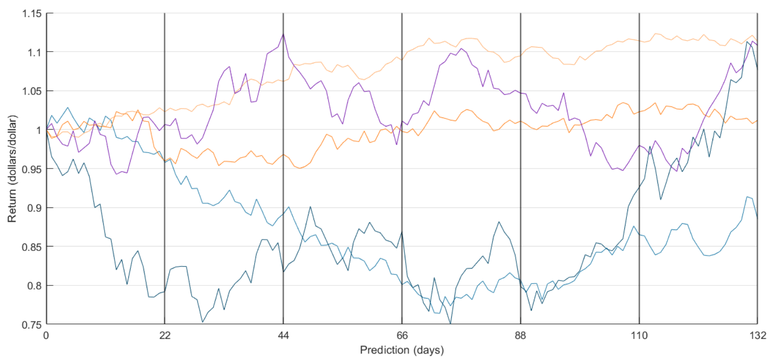

Figure 5: Case D Daily Trends

The first change to note is that five stonks are purchased at 20% of the portfolio each and there is no change in the portfolio throughout the six months. The monthly returns both with and without transaction fees are -5.23%, -2.76%, -4.64%, -4.84%, -1.85%, and 3.89%.

When the stonk prices are correlated using the lower triangular Cholesky factorization of the correlation matrix (Case E), the returns increase but the same five stonks are purchased at the beginning without any portfolio change throughout the prediction. The daily returns are shown in Figure 6. The monthly returns both with and without transaction fees are -6.35%, -4.63%, -6.16%, -5.60%, -3.43%, and 0.48%.

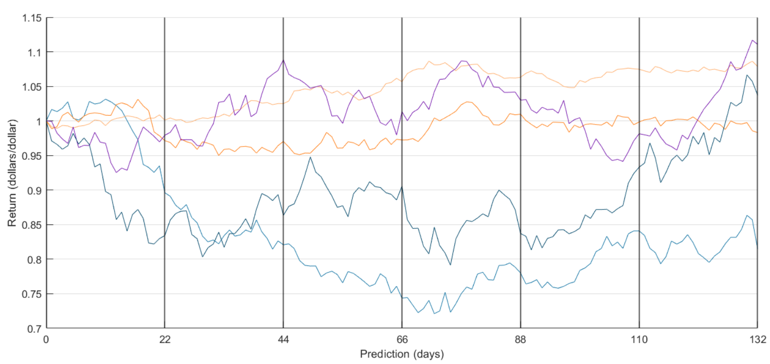

Figure 6: Case E Daily Returns

Alternatively if we define the correlation using the orthogonal eigenvector eigenvalue decomposition, we get the daily returns for Case F shown in Figure 7.

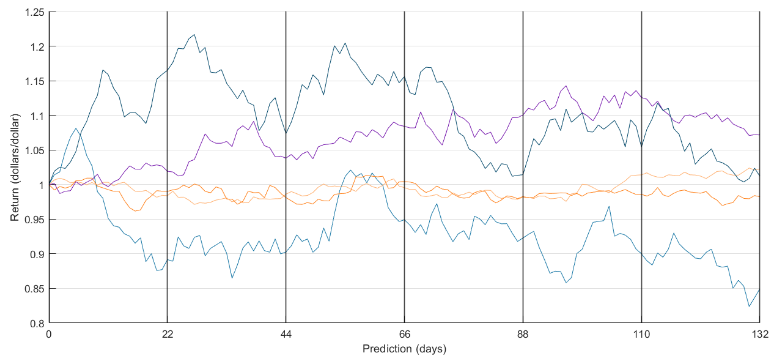

Figure 7: Case F Daily Returns

There is still a lack of transactions in this case so the monthly return both with and without transaction fees are 1.00%, -0.37%, 3.78%, 0.03%, 1.60%, and -1.23%.

Lastly, we consider a portfolio with a maximum single stonk investment of 20% and a minimum stonk investment of 1%. This greatly diversifies the stonks purchased. We begin with Case G in Figure 8 showing the daily responses without transaction fees.

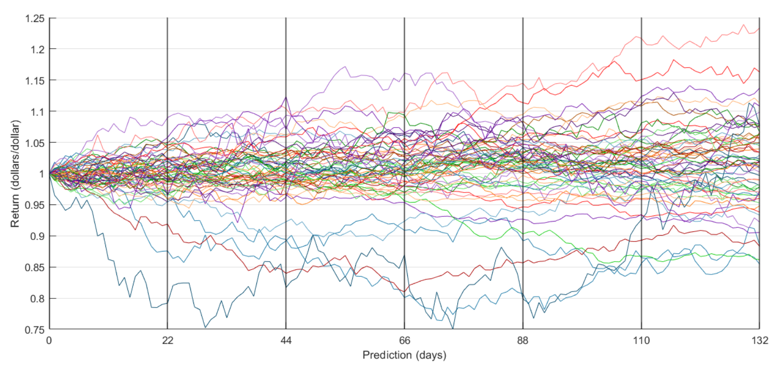

Figure 8: Case G Daily Returns

The most obvious difference is that the figure is far busier than those shown previously. This portfolio requires all sixty stonks to be in the portfolio every month. As a result, there are more transactions made than in the previous cases yielding transaction free returns of -3.00%, -0.79%, -2.74%, -2.75%, -1.50%, and 3.73%, and returns of -3.00%, -0.83%, -2.79%, -2.79%, -1.55%, and 3.45% with transaction fees.

Next, we consider the inclusion of correlation based on the lower triangular cholesky factorization of the correlation matrix. This yields the daily predictions for Case H in Figure 9.

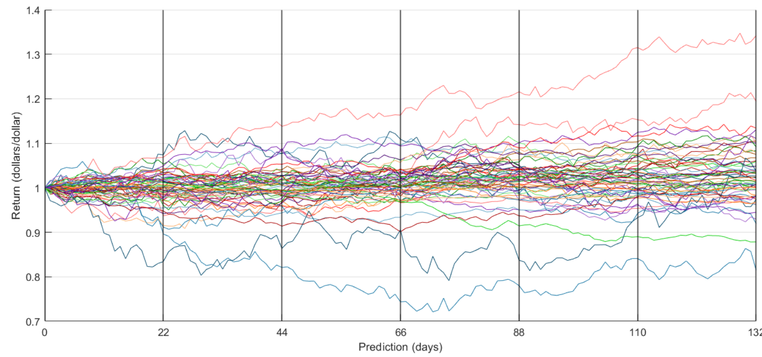

Figure 9: Case H Daily Responses

The portfolio in this case is far more diversified, so the change in trajectories is less clear visually. The monthly returns were -3.64%, -1.47%, -2.45%, -2.05%, -0.73%, and 3.55%. With transaction fees the monthly returns were slightly different at -3.64%,- 1.47%, -2.50%, -2.10%, -0.78%, and 3.27%.

Lastly, we consider the effect of using correlation based on the orthogonal eigenvector eigenvalue decomposition of the correlation matrix. This yields the responses for Case I in Figure 10.

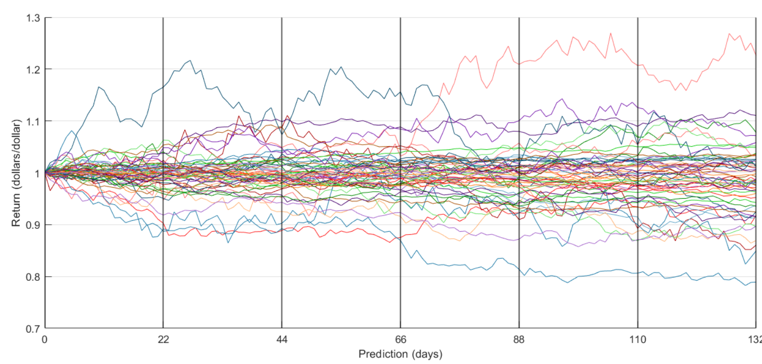

Figure 10: Case I Daily Returns

This resulting returns are 1.23%, 0.83%, 3.30%, 1.20%, 2.23%, and 0.12% without transaction fees and 1.23%, 0.77%, 3.25%, 1.15%, 2.18%, and -0.18% with transaction fees.

The results for Cases A-J are summarized in Table 2 for comparison. The most obvious difference between cases is the low and sometimes even negative returns when correlation is considered. The correlation matrix our portfolio yields has a mean value of 0.1808. This positive mean correlation leads to the stonks in our portfolio tending to trend in the same direction on average. This makes it difficult for our portfolio to return from losses. The returns in this table also show a range in the affect of transaction fees on the six-month return ranging from 0% to 173% based on relative difference. By excluding cases with six month returns less than 1%, this range decrease to 0% to 10.79%. Lastly, sim-month returns in this table don't quite agree with the expectation that increasing portfolio diversity decreases return as suggested by Figure 1.

Table 2: Cases A-J Summary

		Month						
Case	Fees	1	2	3	4	5	6	% Diff
A	n	-3.27%	0.66%	-7.98%	-6.38%	-6.53%	3.07%	10.4%
	y	-3.27%	0.62%	-8.03%	-6.43%	-6.59%	2.75%	
B	n	-6.65%	-3.34%	-9.56%	-7.80%	-7.16%	0.52%	173%
	y	-6.65%	-3.38%	-9.61%	-7.86%	-7.21%	0.19%	
C	n	-1.59%	-0.86%	4.31%	3.04%	4.31%	-0.59%	66%
	y	-1.59%	-0.93%	4.25%	2.98%	4.25%	-0.98%	
D	n	-5.23%	-2.76%	-4.64%	-4.84%	-1.85%	3.89%	0%
	y	-5.23%	-2.76%	-4.64%	-4.84%	-1.85%	3.89%	
E	n	-6.35%	-4.63%	-6.16%	-5.60%	-3.43%	0.48%	0%
	y	-6.35%	-4.63%	-6.16%	-5.60 %	-3.43%	0.48%	
F	n	1.00%	-0.37%	3.78%	0.03%	1.60%	-1.23%	0%
	y	1.00%	-0.37%	3.78%	0.03%	1.60%	-1.23%	
G	n	-3.00%	-0.79%	-2.74%	-2.75%	-1.50%	3.73%	7.51%
	y	-3.00%	-0.83%	-2.79%	-2.79%	-1.55%	3.45%	
H	n	-3.27%	-1.70%	-3.59%	-2.48%	-1.46%	1.68%	7.89%
	y	-3.27%	-1.72%	-3.61%	-2.50%	-1.49%	1.54%	
I	y	-0.42%	-0.31%	1.83%	0.49%	0.98%	-1.40%	10.79%
	n	-0.42%	-0.29%	1.86%	0.51%	1.00%	-1.25%	

To further evaluate the model, we look at the results of several random samples. Table 3 shows the mean and standard deviation of the six-month return of cases A through J for 2,0000 random samples. The returns in this table do follow the return trends suggested in Figure 1 with each additional constraint decreasing the average return, unlike the previous results. Comparing the standard deviations of Case B with Case C, Case E with Case F, and Case H with Case I shows that both approaches to defining correlation yield similar return ranges. The six month returns also show that both methods yield similar mean returns, so it is difficult to make an argument for either approach being superior to the other. Interestingly, the addition of correlation doesn't always significantly reduce the standard deviation of the return as one might expect given the additional constraints it puts on the predictions. In some cases, the mean return with transaction fees is higher than the mean return without transaction fees. This would likely be resolved with a larger set of random samples, but time prohibited.

Table 3: Cases A Through J Average Six Month Return

Case	Fees	Mean	Stdev
A	n	10.13%	14.20%
	y	8.99%	13.58%
B	n	9.68%	13.54%
	y	9.92%	13.92%
C	n	9.18%	13.48%
	y	9.69%	13.54%
D	n	7.22%	9.37%
	y	6.99%	9.11%
E	n	6.94%	9.41%
	y	7.13%	9.13%
F	n	7.30%	6.49%
	y	7.06%	6.55%
G	n	4.34%	7.72%
	y	4.33%	7.18%
H	n	4.51%	7.30%
	y	4.07%	6.92%
I	n	4.49%	4.39%
	y	4.30%	4.36%

To conclude, we show an example of Case E with a risk tolerance of 0.1. The resulting daily returns for this case are shown in Figure 11. This decrease in risk leads to an increase both in portfolio diversity as well as transactions. There are cases where one stonk is sold in its entirety. The monthly returns are -0.12%, 1.02%, -0.47%, -1.97%, -1.35%, and -0.90%. These returns are lower in magnitude than those found previously due to the decreased risk associated with this prediction.

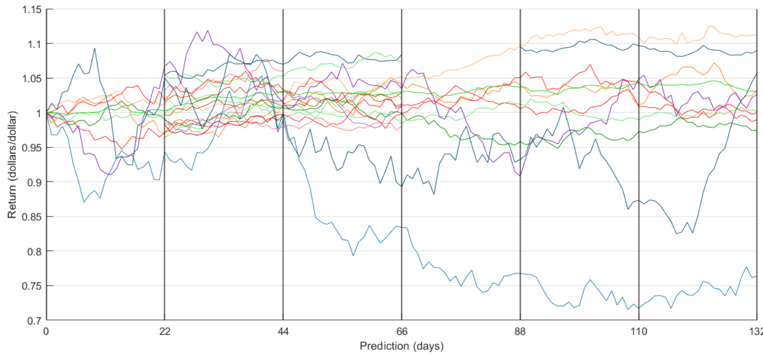

Figure 11: Case E with $\alpha = 0.1$.

4 Conclusion

In this report we demonstrate a stochastic differential equation model for predicting the future price of stonks and an optimization approach to selecting the best portfolio given a tolerated risk. The model yielded the expected results of decreased returns with increased portfolio constraint and reduced risk.

References

[1] Allen Holder and Joseph Eichholz. *An Introduction To Computational Science*. International Series in Operations Research & Management Science. Cham, Switzerland: Springer, 2019.

A Summary

A stochastic differential model is used to increase our chances of a profit in the stonk market. To develop the model, we used Itō 's lemma which simplified the expression as a function of time. To understand which stonks had the highest chances to make return, future stonk prices were determined discretely as shown below,

$$P(t + \Delta t) = P(t) + \mu P(t)\Delta t + \sigma P(t)\sqrt{\Delta t}\phi.$$

The parameters μ and σ are dependent on the previous history of the stonk. Where μ would represent the general trend of a stonk while σ embodies the volatility of a stonk. The parameter ϕ is a standard normal variable and predicts the immeasurable behaviour of the stonk market. While there exist a chance to make profits, it is always possible to loss profits. The portfolio selection is dependent on parameter $\alpha \in [0, 1]$ which measures the relationship between risk and rewards. Low values of α are associated with low risk or low reward while large values include high risk or high reward. The weights for each stonk are determined by solving the optimization problem

$$\begin{aligned} \min \quad & (1-\alpha)(\mathbf{w}^T C\mathbf{w}) + (-\mathbf{r}^T\mathbf{w}) \\ \text{s.t.} \quad & \sum_{i=1}^{n} w_i = 1 \\ & w_i \geq 0 \qquad \forall i. \end{aligned}$$

where $\mathbf{r}$ is an array of mean returns for each stonk and C is a covariance matrix. With this model, we were able to predict future prices for stonks both with and without correlation. Our results show that our correlation models have little affect on average six-month return, but risk tolerance and portfolio constraints do as demonstrated in Table 3.

B Portfolio

The following table lists of all stonks that were selected from when creating the portfolio

Name	Symbol	Name	Symbol
Alphabet Inc	GOOG	Apple Inc	AAPL
Tesla	TSLA	General Electric Company	GE
Coca-Cola Co	KO	General Motors Company	GM
Kroger Co	KR	Peloton Interactive Inc	MRO
Twitter Inc	TWTR	McDonald's Corp	MCD
Wendys Co	UNH	Merck & Co., Inc.	LLY
AeroVironment, Inc.	AVAV	BAE Systems PLC	BA.L
Lockheed Martin Corporation	LMT	Boeing Co	BA
Northrop Grumman Corporation	NOC	Kubota Corp	KUBTY
Deere & Company	DE	Smith & Wesson Brands Inc	SWBI
Sturm Ruger & Company Inc	SPWH	American Outdoor Brands Inc	AOUT
Oshkosh Corp	OSK	Rolls-Royce Holding PLC ADR	RYCEY
Leonardo SPA Unsponsored ADR	FINMY	Rheinmetall ADR	RNMBY
Gold	GC=F	Platinum	PL=F
Intel Corporation	INTC	Walt Disney Co	DIS
Crude Oil	CL=F	Unity Software Inc	U
American Airlines Group Inc	AAL	Nokia Oyj	NOK
Silver	SI=F	Visa Inc	V
Exxon Mobil Corp	XOM	Pfizer Inc.	PFE
Natural Gas	NG=F	Artificial intelligence Tech Solutns Inc	AITX
Wipro Limited	WIT	Nintendo 8 ADR	NTDOY
Ubi Soft Entertainment ADR	UBSFY	NVIDIA Corporation	NVDA
Amazon.com, Inc.	AMZN	3M Co	MMM
Chipotle Mexican Grill, Inc.	CMG	United Parcel Service, Inc.	UPS
Sherwin-Williams Co	SHW	Aterian Inc	ATER
Waste Management, Inc.	WM	PepsiCo, Inc.	PEP
Microsoft Corporation	MSFT	O'Reilly Automotive Inc	ORLY
Uranium Energy Corp	UEC	Fidelity Advisor Semiconductors	FIKGX

C Memes

Figure 1: Stonks Meme 1

Figure 2: Stonks Meme 2

Figure 3: Stonks Meme 3

Figure 4: Stonks Meme 4

Figure 5: Stonks Meme 5