STAT 27410 Final Project Proposal - A Bayesian Approach to Portfolio Management

Andrew Su, Scarlett He

1. Introduction

Portfolio management is a complex endeavor that involves not only selecting appropriate investments but also navigating the uncertainties posed by market dynamics and economic conditions. While diversification across sectors and asset classes is widely regarded as a cornerstone of effective portfolio construction, managing risk and optimizing returns remains a significant challenge. This study seeks to explore innovative approaches to enhance investment decision-making by comparing traditional frequentist methods with Bayesian techniques in the context of portfolio optimization.

Over a three-year period from January 1st, 2022, to January 1st, 2025, we analyze historical stock prices for five individual stocks:

- Apple Inc. Common Stock (AAPL) (Nasdaq, n.d.b)
- Coca-Cola Company (The) Common Stock (KO) (Nasdaq, n.d.c)
- Costco Wholesale Corporation Common Stock (COST) (Nasdaq, n.d.d)
- Advanced Micro Devices Inc. Common Stock (AMD) (Nasdaq, n.d.a)
- Salesforce Inc. Common Stock (CRM) (Nasdaq, n.d.f)

Two sector specific ETFS:

- Invesco QQQ Trust (QQQ) (Nasdaq, n.d.e)
- the Consumer Discretionary SPDR Select Sector Fund (XLY) (Nasdaq, n.d.h)

As well as S&P 500 index (SPX) (Nasdaq, n.d.g) as a baseline for comparison.

All data used is from the 3-year timespan of January 1st, 2022 to January 1st, 2025 (753 trading days) and obtained directly from the Nasdaq stock exchange. Data was sourced directly from the Nasdaq stock exchange, ensuring a robust empirical basis for our analysis.

The selection of these securities reflects a strategic focus on diverse industries and market segments. We chose individual companies and ETFs to represent both the Technology and Consumer sectors, ensuring a mix of defensive and cyclical stocks while avoiding undue

similarity among the selected investments. This approach aims to capture variations in market performance and provide a comprehensive view of portfolio dynamics.

For each of the securities, separate datasets were used. For each security, the data set include the following variables:

- **Date**: The date of the trading day
- **Close/Last**: The price of the security at the end of the trading day
- **Volume**: The total number of shares traded during the trading day
- **Open**: The price of the security at the start of the trading day
- **High**: The highest price of the security during the entire trading day
- **Low**: The lowest price of the security during the entire trading day

For the purpose of our analysis, we will only be looking at the Date, Close/Last, Open and Volume categories across each of the securities to build a fundamental AR model to capture essential market dynamics and trends.

To determine weights of allocation, we will run regression models to predict the future prices of each stock, calculating the percentage changes of each and assigning portfolio weights through convex optimization. For instance, for a certain stock A, our model will include the previous prices of stock A into our calculation for future price of A, while also incorporating the sector-ETF and SPX as measures of how well the economy is doing.

We define the stock price of stock i at time t as $P_{i,t}$. The predictive model for stock prices is given as follows:

For each technology sector stock, the model is expressed as:

$$\begin{aligned} P_{i,t+1} = \beta_1 P_{i,t} + \beta_2 P_{i,t-1} + \\ \beta_3 \text{QQQ}_t + \beta_4 \text{SPY}_t + \sum_{j \neq i} \text{COV}(P_{i,t}, P_{j,t}) \end{aligned}$$

For each consumer sector stock, the model is expressed as:

$$\begin{aligned} P_{i,t+1} = \beta_1 P_{i,t} + \beta_2 P_{i,t-1} + \\ \beta_3 \text{XLY}_t + \beta_4 \text{SPY}_t + \sum_{j \neq i} \text{COV}(P_{i,t}, P_{j,t}) \end{aligned}$$

The percentage change in stock price from time t to $t+1$, denoted as $R_{i,t}$, is given by:

$$R_{i,t} = \frac{P_{i,t+1} - P_{i,t}}{P_{i,t}}$$

As opposed to a pure Auto-Regressive (AR) model, we will consider exogenous variables that aim to capture aspects of the economy that cannot be fully captured with our limited selection of stocks. This would allow for specific analysis while also factoring in more complex market trends. *Note: Our models are subject to change.*

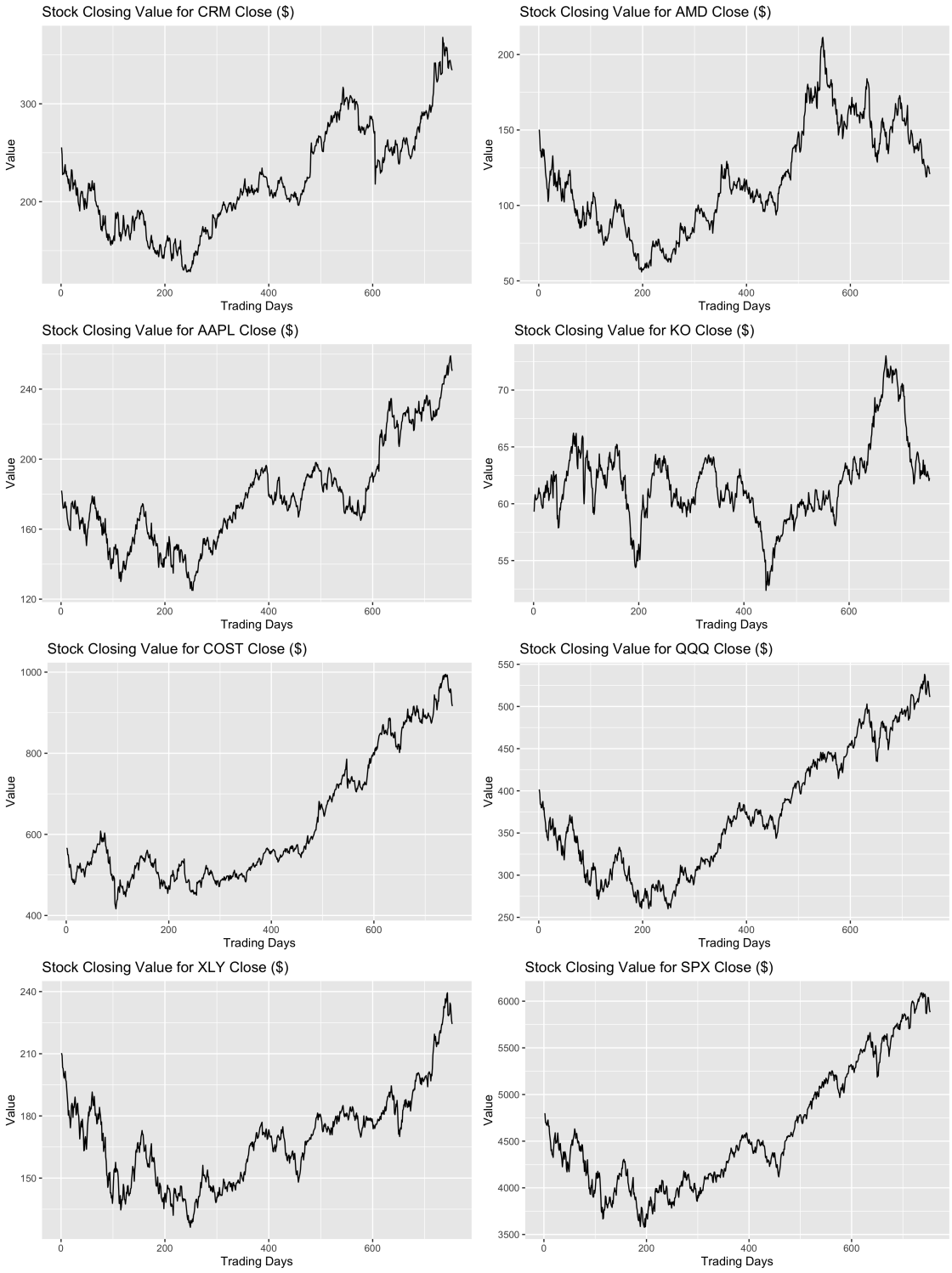
With computed stock price percentage changes, we will use convex optimization to determine the portfolio weights, where volume traded will play a role in managing the risk assigned to each stock.

We will then compare these frequentist approaches with Bayesian methods, where we will assume a prior distribution (multivariate normal or student-t) on an asset, and then continuously update the distribution using the market data as the likelihood function to generate a posterior distribution that gives the most likely expected return for assets and measuring uncertainty through credible intervals. These returns can then be used in combination with the uncertainty measures to allocate portfolio weights.

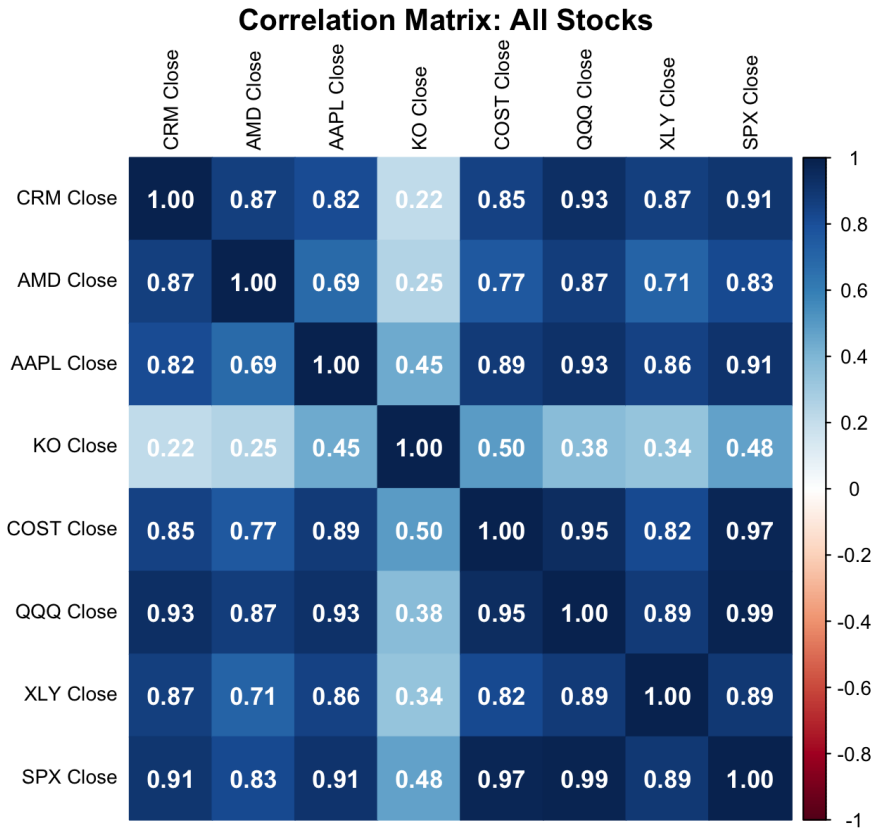
This study not only sets the stage for a detailed comparison between frequentist and Bayesian methods in portfolio optimization but also provides insights into the important features to capture comprehensive market dynamics.

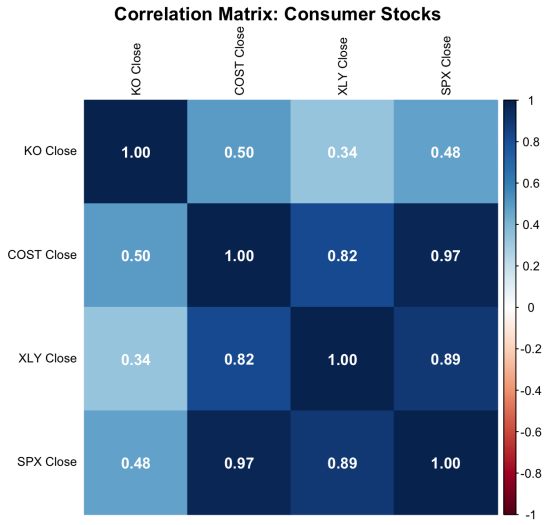
2. Exploratory Data Analysis

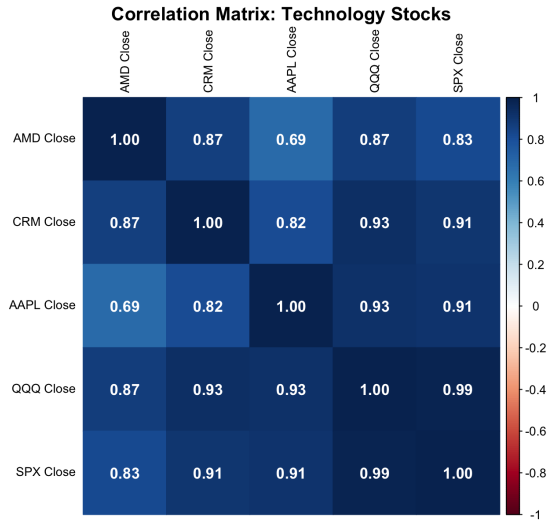
We will start by looking at the trends and changes in each of the securities over this 3 year period.



These graphs show the trends of the 7 securities over the 3-year period. It can be noted that some graphs have similar trends. Thus, we look at correlation matrix of those stock, and we found the correlation is relatively high due to the choice of our stocks are all well known and having a high market share.







3. Frequentist Analysis

Perform the frequentist analysis in this session.

3.1 Proposed Frequentist Model(s)

In this section, formulate the frequentist model(s) you are going to use to analyze your dataset. Be sure to first define the notations involved in the model(s).

3.2 Fitting the Frequentist Model(s)

In this section, * discuss how you fit the proposed frequentist model(s). * report the results. * interpret the results in the context.

4. Bayesian Analysis

Propose the Bayesian analysis you will work on during the rest of the quarter in this session.

4.1 Proposed Bayesian Model(s)

In this section,

- formulate the Bayesian model(s) you are going to use to analyze your dataset. Be sure to first define the notations involved in the model(s).
- discuss how you will elicit the prior(s).

4.2 Fitting the Bayesian model(s)

- Propose how you will fit the proposed Bayesian models.
- Propose how you will perform sensitivity analysis of the Bayesian models, i.e., how the posterior distribution is affected by the prior
- Propose how you will check the MCMC convergence.

4.3 Prediction

In this section, propose how you can make predictions using the Bayesian model.

5. Discussion

In this section, discuss how you can improve your model.

6. Contributions

In this section, discuss the percentage of your contributions to the development final project proposal. Report the number of hours you have worked on the proposal, and the sections you are involved.

Please also discuss briefly the contributions of your teammate(s), as well as the help and support you got from your teammates(s).

References

Nasdaq. n.d.a. *Advanced Micro Devices, Inc. Common Stock (AMD) Historical Quotes*. https://www.nasdaq.com/market-activity/stocks/amd/historical.

———. n.d.b. *Apple Inc. Common Stock (AAPL) Historical Quotes*. https://www.nasdaq.com/market-activity/stocks/aapl/historical.

———. n.d.c. *Coca-Cola Company (the) Common Stock (KO) Historical Quotes*. https://www.nasdaq.com/market-activity/stocks/ko/historical.

———. n.d.d. *Costco Wholesale Corporation Common Stock (COST) Historical Quotes*. https://www.nasdaq.com/market-activity/stocks/cost/historical.

———. n.d.e. *Invesco QQQ Trust, Series 1 (QQQ) Historical*. https://www.nasdaq.com/market-activity/stocks/qqq/historical.

———. n.d.f. *Salesforce, Inc. Common Stock (CRM) Historical Quotes*. https://www.nasdaq.com/market-activity/stocks/crm/historical.

———. n.d.g. *S&p 500 (SPX) Historical Data*. https://www.nasdaq.com/market-activity/index/spx/historical.

———. n.d.h. *SPDR Select Sector Fund - Consumer Discretionary (XLY) Historical*. https://www.nasdaq.com/market-activity/stocks/xly/historical.

Appendix

https://raw.githubusercontent.com/asu1-1/STAT27410-Final-Project/refs/heads/main/27410%20Final%20Project%20Code.R

https://raw.githubusercontent.com/asu1-1/STAT27410-Final-Project/refs/heads/main/Scarlett_part1_code.R