Economic Behavior in two different Political Regimes

Investigating the economic behavior of the top large cap and top mid cap companies during the political tenures of UPA and BJP

Himadri Mandal

Bachelor's in Statistics

Statistical Methods - III

Kolkata, December 2024

Economic Behavior in two different Political Regimes

Investigating the economic behavior of the top large cap and top mid cap companies during the political tenures of UPA and BJP

Himadri Mandal

Supervisor: Prof. Ayan Basu

Professor, ISRU, Indian Statistical Institute, Kolkata

Bachelor's in Statistics

Statistical Methods - III

Project

Kolkata, December 2024

This page intentionally left blank.

Declaration of Authorship

Has undersigned, hereby it his declared that this work entitled "Economic Behavior in two different Political Regimes" is the original work and that it has not previously in its entirety or in part been submitted at any university or higher education institution for the award of any degree, diploma, or other qualifications. It is also hereby declared that to the best of the knowledge, this work contains no material previously published or written by another person, except where due reference, acknowledgement, and citation is made.

Kolkata, December 2024

Himadri Mandal

This page intentionally left blank.

Abstract

Investigating the behavior of the stock prices of top Large Cap and Small Cap companies in the **UPA-II** tenure and comparing it with **BJP-I**.

This page intentionally left blank.

Contents

This page intentionally left blank.

List of Figures

This page intentionally left blank.

1

The Goal, and the Data Sources

1.1 Goal

The Goal of the project is to understand the stock market behavior of the top large cap and mid cap companies, in India, during different political tenures. We try to address two crucial questions:

1. does the political atmosphere in the BJP tenure **encourage** large-cap companies and **discourage** mid-cap companies as opposed to the UPA tenure, as one would assume from all the hear-say in the media?
2. is there something you can say about the volatility of the economic behavior of top large cap and small cap companies during these two different political environments?

1.2 Data Source

We use data from **Yahoo! Finance**, which hosts stock market data from around the Globe. We used data from **YF** because it is well-supported by Python libraries written by the community. Also, certain analysis was done after correcting the prices using the **Consumer Price Index**, the **CPI** data was obtained from **https://data.worldbank.org** The following Python libraries have been used extensively in this project:

1. yfinance: **https://github.com/ranaroussi/yfinance**
2. numpy: **https://numpy.org/**
3. pandas: **https://pandas.pydata.org/**

4. matplotlib: **https://matplotlib.org/**

We use the **^NSEI, ^NSEMDCP50** indices for our analysis, which represent the **NIFTY50** and **NIFTY MID-CAP 50** indices. We collect data of both the indices from

- **2009-05-25 - 2014-05-17**: **UPAII tenure**
- **2014-05-26 - 2019-05-24**: **BJPI tenure**
- **2019-06-03 - 2024-05-15**: **BJPII tenure**

We wished to use data from the **UPAI tenure** as well, but couldn't find well-organized data of the indices used, from back then.

2

Exploration of the data

2.1 The Stock Data

Price	Adj Close	Close	High	Low	Open	Volume
Ticker	^NSEI	^NSEI	^NSEI	^NSEI	^NSEI	^NSEI
Date						
2009-05-25 00:00:00+00:00	4237.549805	4237.549805	4270.049805	4205.100098	4238.100098	0
2009-05-26 00:00:00+00:00	4116.700195	4116.700195	4256.049805	4092.250000	4239.549805	0
2009-05-27 00:00:00+00:00	4276.049805	4276.049805	4286.450195	4115.250000	4117.299805	0
2009-05-28 00:00:00+00:00	4337.100098	4337.100098	4354.850098	4254.850098	4276.149902	0
2009-05-29 00:00:00+00:00	4448.950195	4448.950195	4488.049805	4340.750000	4340.750000	0
...	...	...	...	...	...	...
2014-05-12 00:00:00+00:00	7014.250000	7014.250000	7020.049805	6862.899902	6863.399902	157700
2014-05-13 00:00:00+00:00	7108.750000	7108.750000	7172.350098	7067.149902	7080.000000	232200
2014-05-14 00:00:00+00:00	7108.750000	7108.750000	7142.250000	7080.899902	7112.000000	177200
2014-05-15 00:00:00+00:00	7123.149902	7123.149902	7152.549805	7082.549805	7111.299805	186800
2014-05-16 00:00:00+00:00	7203.000000	7203.000000	7563.500000	7130.649902	7270.200195	393200

1221 rows × 6 columns

Figure 2.1: *UPAII NSEI Dataset*

The dataset of both the indices has the same form. We ensure that the number of rows is same throughout the three tenures as that becomes important in the subsequent analysis.

The stock price data has 6 columns:

- **Close**: The price at the end of the trading session.
- **Adj Close**: The adjusted closing price (includes dividends, stock splits, rights offerings, mergings, etc.)
- **High**: The highest price throughout the trading session.
- **Low**: The lowest price throughout the trading session.

- **Open**: The price at the start of the trading session.
- **Volume**: Number of shares traded. (0 indicates that the data isn't recorded.)

2.1.1 Accurate Partitioning of the Stock Data

```
UPAI = {
    "start" : "2004-05-22",
    "end" : "2009-05-21",
    "PrimeMinister" : "Manmohan Singh",
    "Party" : "Indian National Congress",
    }

UPAII = {
    "start" : "2009-05-25", #Offsetting for Monday
    "end" : "2014-05-17", #Offsetting for period
    "PrimeMinister" : "Manmohan Singh",
    "Party" : "Indian National Congress",
    }

BJPI = {
    "start" : "2014-05-26", #Offsetting for Number of days in UPAII
    "end" : "2019-05-24",
    "PrimeMinister" : "Narendra Modi",
    "Party" : "Bharatiya Janata Party",
    }

BJPII = {
    "start" : "2019-06-03", #Offsetting for Monday
    "end" : "2024-05-15", #Offsetting for Period
    "PrimeMinister" : "Narendra Modi",
    "Party" : "Bharatiya Janata Party",
    }
✓ 0.0s
```

Figure 2.2: *Data Partition*

As mentioned already, we partition the stock data into the time periods of different political regimes. We ensure that the partition keeps the days of the week aligned throughout regimes (for the most part). Although initially we wished to include the **UPAI** regime, there wasn't enough data available.

We proceed to download the partitioned stock data using the **yfinance** module.

```
#download stock data

stock_data = {}
print(
    {UPAII['start'], UPAII['end']}, {BJPI['start'], BJPI['end']}, {BJPII['start'], BJPII['end']}
)

for index in Index:
    stock_data[("UPAII", index)] = yf.download(index, start=UPAII["start"], end=UPAII["end"])
    stock_data[("BJPI", index)] = yf.download(index, start=BJPI["start"], end=BJPI["end"])
    stock_data[("BJPII", index)] = yf.download(index, start=BJPII["start"], end=BJPII["end"])
```

2.2 CPI Data

For further comparitive analysis across regimes it is important to use the **Consumer Price Index** to offset the effect of **inflation**. What is **CPI**? The **CPI** measures the average change in prices of a basket of goods and services over time, reflecting the cost of living.

$$\text{CPI} = \left(\frac{\text{Cost of Basket in Current Year}}{\text{Cost of Basket in Base Year}}\right) \times 100$$

```

import matplotlib.pyplot as plt

years = list(CPI.keys())
cpi_values = list(CPI.values())

plt.figure(figsize=(10, 6))
plt.plot(years, cpi_values, marker='o', linestyle='-', color='b')
plt.xlabel('Year')
plt.ylabel('CPI Value')
plt.title('Consumer Price Index (CPI) Over Years')
plt.grid(True)
plt.show()
```

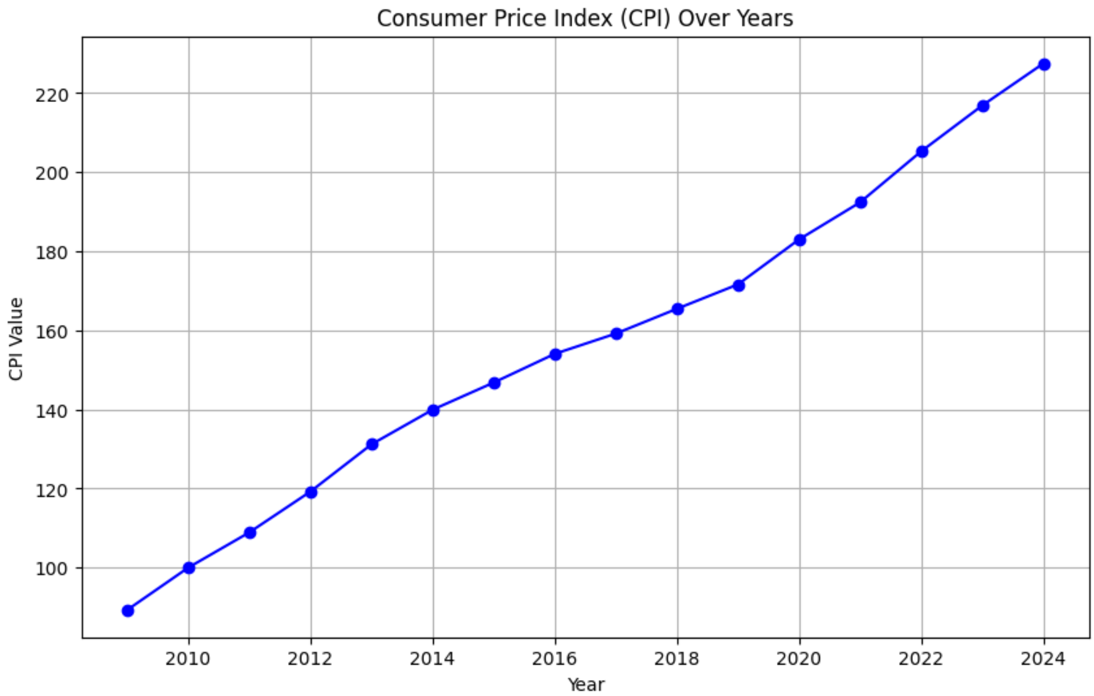

Figure 2.3: *Consumer Price Index*

3

Analysis

3.1 Question #1

Question

Is the political environment in **BJP** regimes more favourable for the top large-cap companies, and unfavourable for the top mid-cap companies as opposed to the **UPA** regimes, as one following establishment critical media sources might think?

3.1.1 Setup

We don't have data from the **UPAI** regime, so we have to test our hypothesis with **UPAII** regime only.

We use two indices **^NSEI** (NIFTY50) and **^NSEMDCP50** (NIFTY MID-CAP 50) which represent the behaviour of the stocks of the top 50 large cap and mid cap companies.

We adjust the prices throughout the time period using CPI to offset for inflation. This makes the observations comparable throughout different periods of time.

Our variables of comparison would be adjusted price deltas. That is, if $[X_1, X_2, X_3, \dots]$ are the **prices** (we are using opening prices throughout) and $[f(1), f(2), f(3), \dots]$ are the **CPI values** then we consider

$$\left[\left(\frac{X_2}{f(2)} - \frac{X_1}{f(1)}\right) \cdot 100, \left(\frac{X_3}{f(3)} - \frac{X_2}{f(2)}\right) \cdot 100, \left(\frac{X_4}{f(4)} - \frac{X_3}{f(3)}\right) \cdot 100, \dots\right]$$

$$= \Big[Y_1, Y_2, Y_3, \dots \Big]$$

as our **adjusted price deltas**.

3.1.2 Averaged Adjusted Price Deltas

Listing 3.1: *Adjusting Price Deltas using CPI*

```
df_cpi_adj = {}
for regime in df:
    df_cpi_adj[regime] = df[regime]
print(df)
print('-'*100)
for regime in df:
    for i, df_stock in enumerate(df[regime]):
        df_stock['AdjCloseCPI'] = df_stock['Adj Close'].values
        df_stock['OpenCPI'] = df_stock['Open'].values
        index = df_stock.index

        tmpOpen = df_stock['Open'].values
        tmpAdjClose = df_stock['Adj Close'].values
        for j, date in enumerate(index):
            year = date.year
            if year in CPI:
                CPI_value = CPI[year]
                tmpOpen[j] /= CPI_value / 100
                tmpAdjClose[j] /= CPI_value / 100
        df_stock['OpenCPI'] = tmpOpen
        df_stock['AdjCloseCPI'] = tmpAdjClose
        df_cpi_adj[regime][i] = df_stock

df_cpi_adj
```

Now, we plot the charts of **unadjusted price deltas** and **adjusted price deltas**.

Listing 3.2: *Plotting Adjusted Price Deltas*

```
import matplotlib.pyplot as plt

# Define colors for the indices
colors = {'^NSEI': 'blue', '^NSEMDCP50': 'red'}

# Plotting the means and standard deviations in a single graph
plt.figure(figsize=(6, 10))

for regime in deltas_cpi:
    for i, index in enumerate(Index):
        means = deltas_cpi[regime][i].mean()
        std_devs = deltas_cpi[regime][i].std()
        mid_point = len(deltas_cpi[regime][i]) // 2

        plt.plot(deltas_cpi[regime][i].index, [means] * len(deltas_cpi[regime][i]),
        ↪  color=colors[index])
        plt.errorbar(deltas_cpi[regime][i].index[mid_point], means, yerr=std_devs,
        ↪  fmt='o', alpha=0.5, capsize=10, color=colors[index])
        plt.fill_between(deltas_cpi[regime][i].index, means - std_devs, means +
        ↪  std_devs, color=colors[index], alpha=0.1)

# Custom legend
handles = [plt.Line2D([0], [0], color=color, lw=2) for color in colors.values()]
labels = [index for index in colors.keys()]
plt.legend(handles, labels, title="Indices", fontsize="small", markerscale=0.5,
↪  borderpad=0.5, loc="upper right")

plt.title("Means and Standard Deviations of Adjusted Price Deltas")
plt.tight_layout()
plt.show()
```

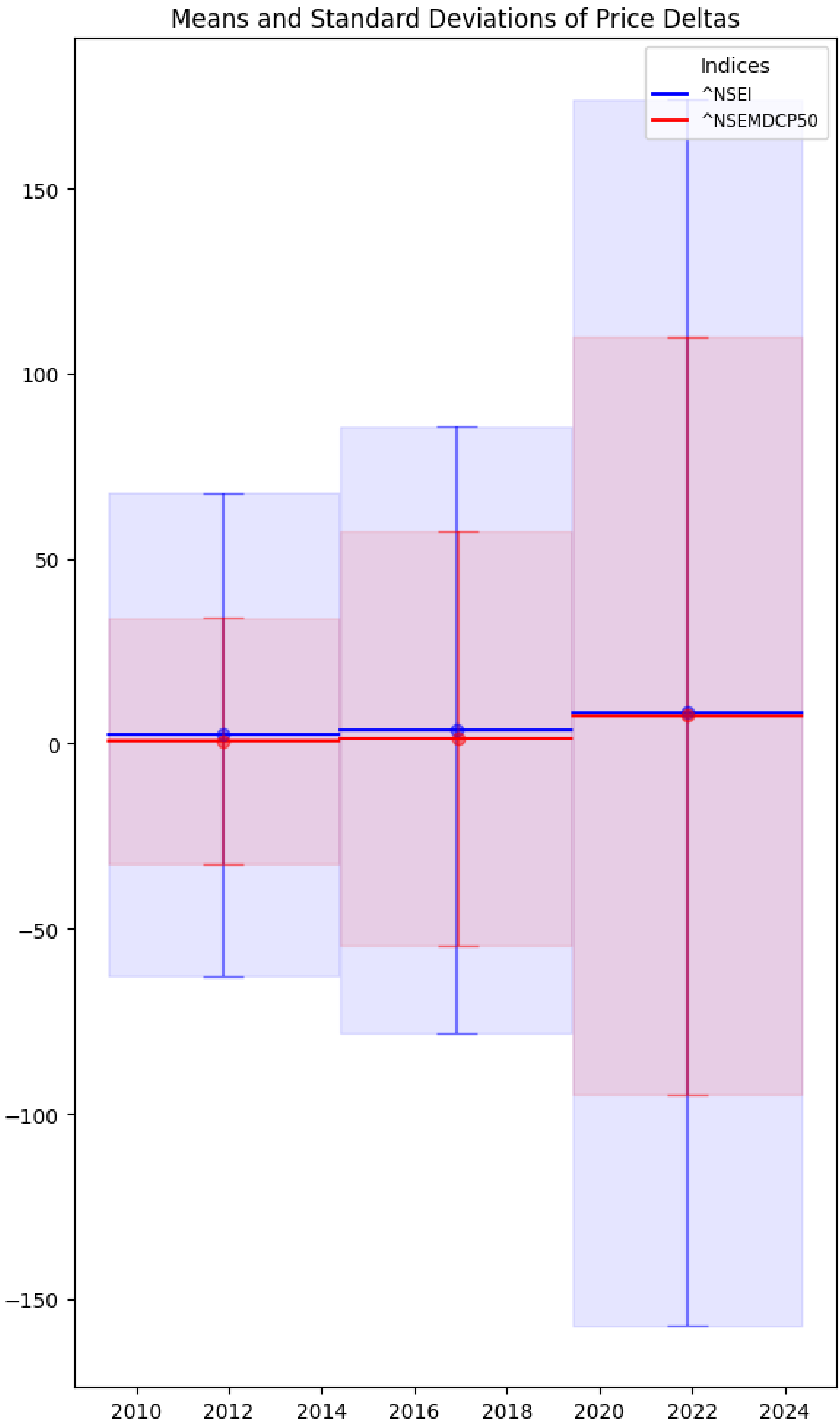

Figure 3.1: *Unadjusted Price Deltas*

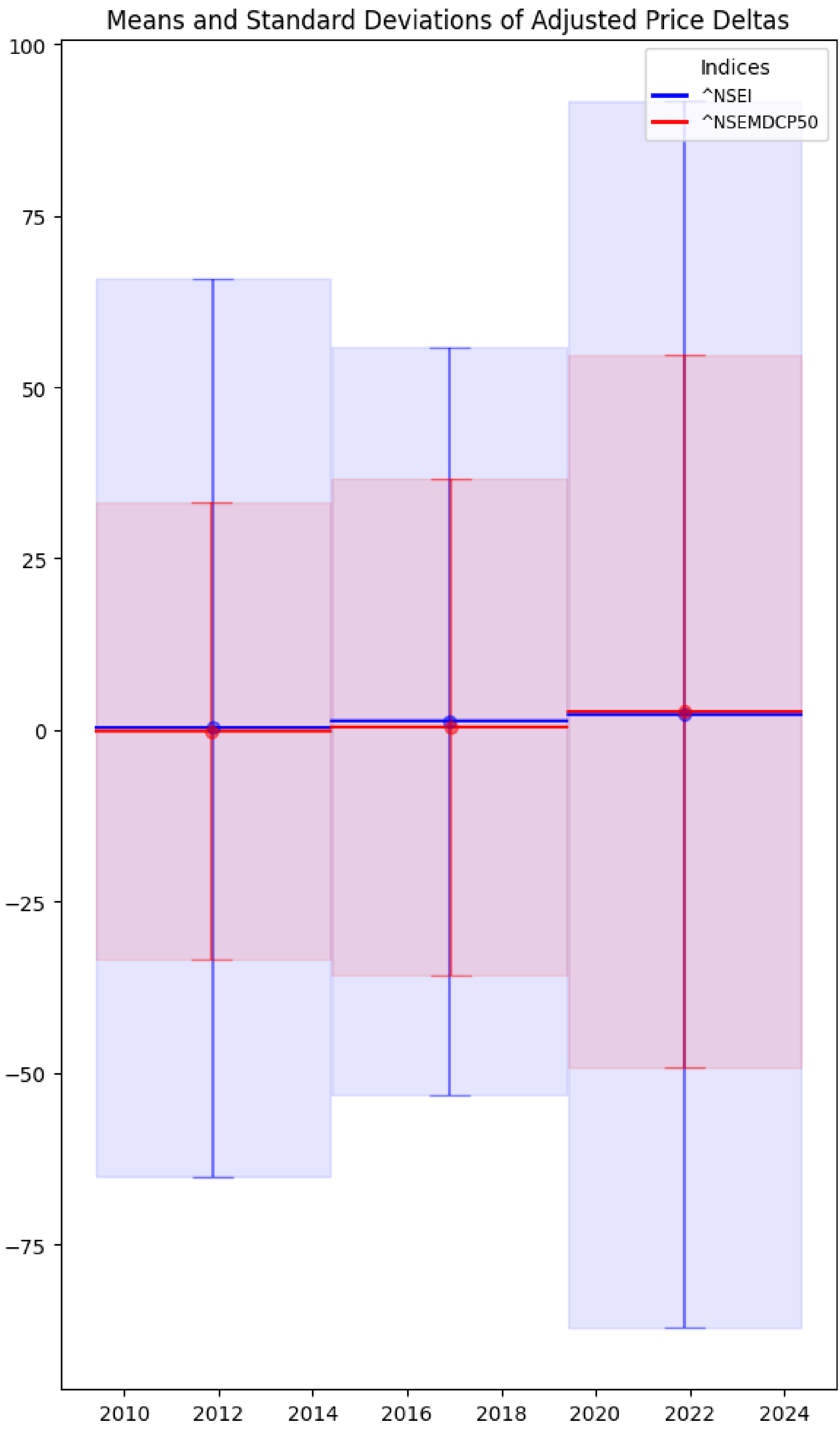

Figure 3.2: *Adjusted Price Deltas*

There is clearly a very big difference in unadjusted and adjusted price deltas. But there is a problem, it appears that there is high volatility in adjusted price deltas.

To counter for this, I averaged out the price deltas week-wise (i.e. average the first 5 observations, then the second 5 observations, and so on...) which leads me to the **Averaged Adjusted Price Deltas**

$$\left[\frac{Y_1 + Y_2 + \cdots + Y_5}{5}, \frac{Y_6 + Y_7 + \cdots + Y_{10}}{5}, \ldots\right]$$

$$= [Z_1, Z_2, Z_3, \ldots]$$

3.1.3 Split Adjusted Price Deltas

To better understand the behavior throughout the different regimes, I also plot out the Adjusted Price Deltas with the periods split into years.

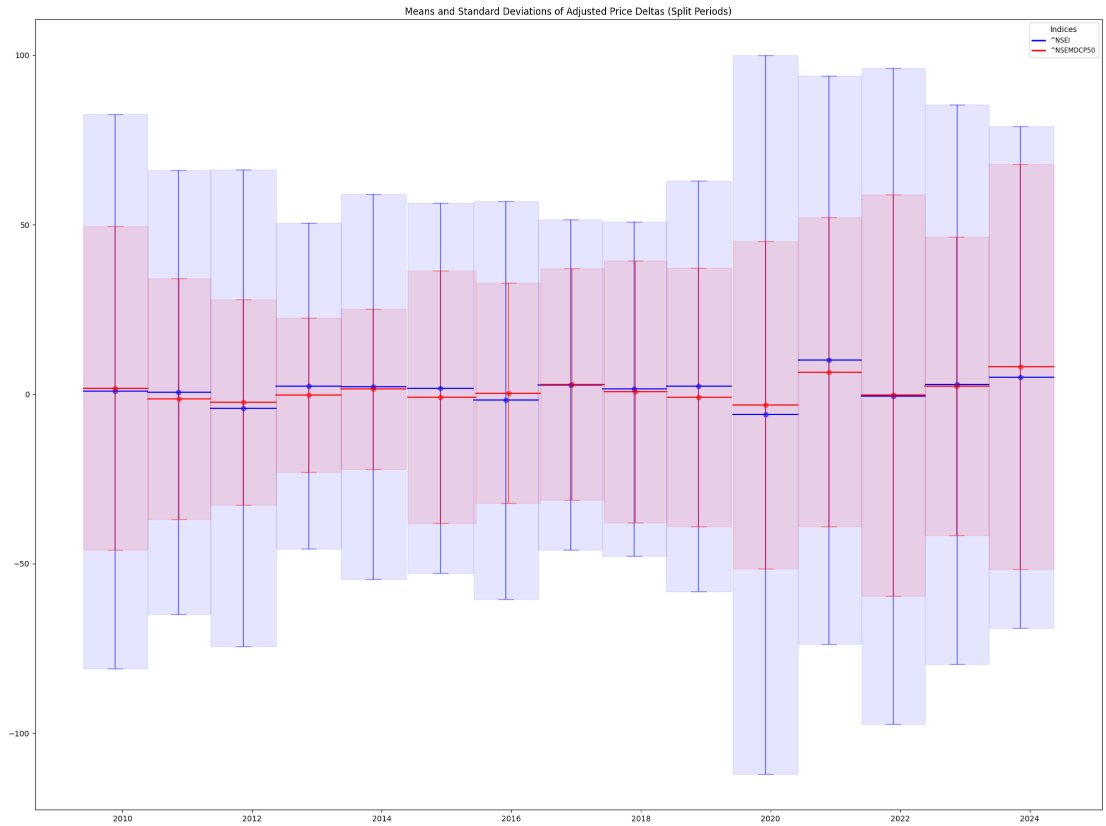

Figure 3.3: *Split Adjusted Price Deltas*

Post 2020 seems clearly anomalous possibly due to COVID-19, therefore I try to avoid drawing bold conclusions out of the **BJPII** regime. Also, it appears that the volatility of adjusted price deltas in the **UPAII** regime is higher than the **BJPI** regime for large cap companies whereas the opposite seems true for the mid cap companies. This will be our second question.

3.1.4 Statistical Analysis

Let R_1 and R_2 be the **UPAII** and **BJPI** regimes. Let $G : \{R_1, R_2\} \times N \to \mathbb{N}$ be the index function that represents the appropriate indices of the averaged adjusted price deltas through the different regimes.

Since our question wondered about the change in behaviour during R_1 and R_2, it makes sense to model the situation as follows. Let $D(n)$ represent the difference between the averaged adjusted price deltas in the n'th week of the regimes R_1 and R_2, i.e.

$$D(n) = Z_{G(R_2,n)} - Z_{G(R_1,n)}$$

with $D(i)$ following **IID** normal with mean μ. Then,

$$\frac{\sqrt{N}(\bar{D} - \mu)}{\sqrt{\sum_i (D(i) - \bar{D})^2}} \sim t_{(N-1)}$$

Large Cap

$$\begin{aligned} H_0 &: \mu = 0 \quad \text{(Null Hypothesis)} \\ H_1 &: \mu > 0 \quad \text{(Alternative Hypothesis)} \end{aligned}$$

The plot and the code is attached. Here, we find that the t-statistic is **0.42** which is well below the **5%** threshold to the right.

Clearly, there's not enough evidence to reject the Null Hypothesis.

Listing 3.3: *Large Cap t-test*

```
NSEI = {}
for regime in deltas_cpi:
    NSEI[regime] = deltas_cpi_grouped[regime]["^NSEI"].values

X = NSEI["UPAII"]
Y = NSEI["BJPI"]
n = len(X)
assert (n == len(Y))
D=np.empty(n,)

for i,val in enumerate(zip(X,Y)):
    D[i] = int(val[1]-val[0])

from scipy.stats import t as t_dist

D_mean = D.mean()
s2 = ((D-D_mean)**2).sum()/(n-1)
t = np.sqrt(n) * D_mean / np.sqrt(s2)
print(t)
# Plotting the t-distribution
plt.figure(figsize=(12, 6))
x = np.linspace(-4, 4, 1000)

# Calculate the 5% probability threshold towards the right
threshold = t_dist.ppf(0.95, n-1)
plt.axvline(threshold, color='green', linestyle='--', label=f'5% threshold =
↪  {threshold:.2f}')
y = t_dist.pdf(x, n-1)

plt.plot(x, y, label=f't-distribution with n-1={n-1} degrees of freedom')
plt.axvline(t, color='red', linestyle='--', label=f't-statistic = {t:.2f}')
plt.title('t-distribution')
plt.xlabel('t')
plt.ylabel('Probability Density')
plt.legend()
plt.show()
```

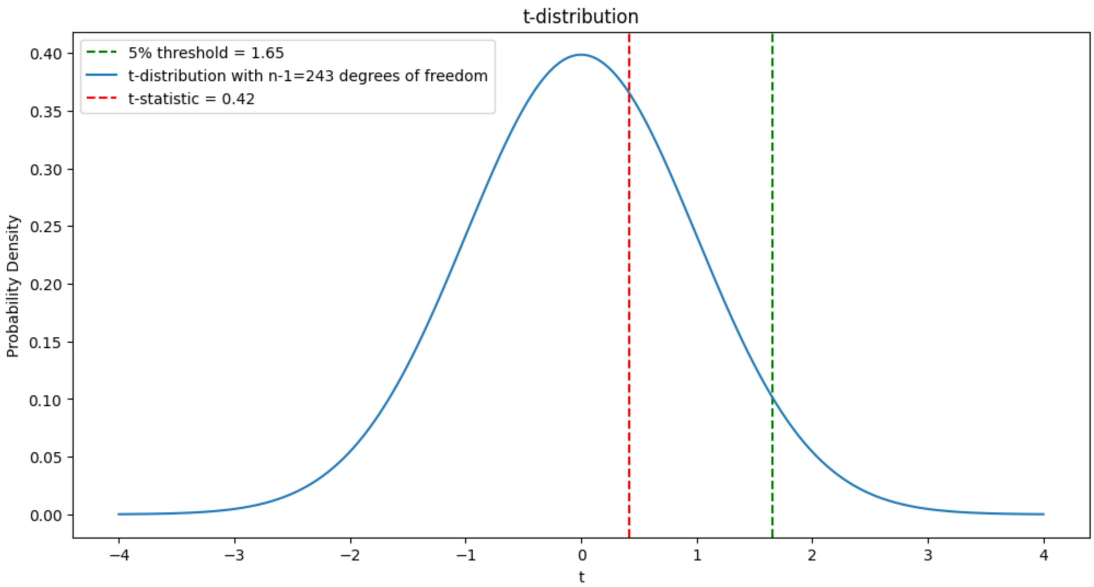

Figure 3.4: *Large Cap t-test*

Mid Cap

$$H_0 : \mu = 0 \quad \text{(Null Hypothesis)}$$
$$H_1 : \mu < 0 \quad \text{(Alternative Hypothesis)}$$

The plot is attached. The code is almost similar, so it's omitted. Here, we find that the t-statistic is **0.55** which is way greater than the **5%** threshold to the left.

Clearly, there's not enough evidence to reject the Null Hypothesis.

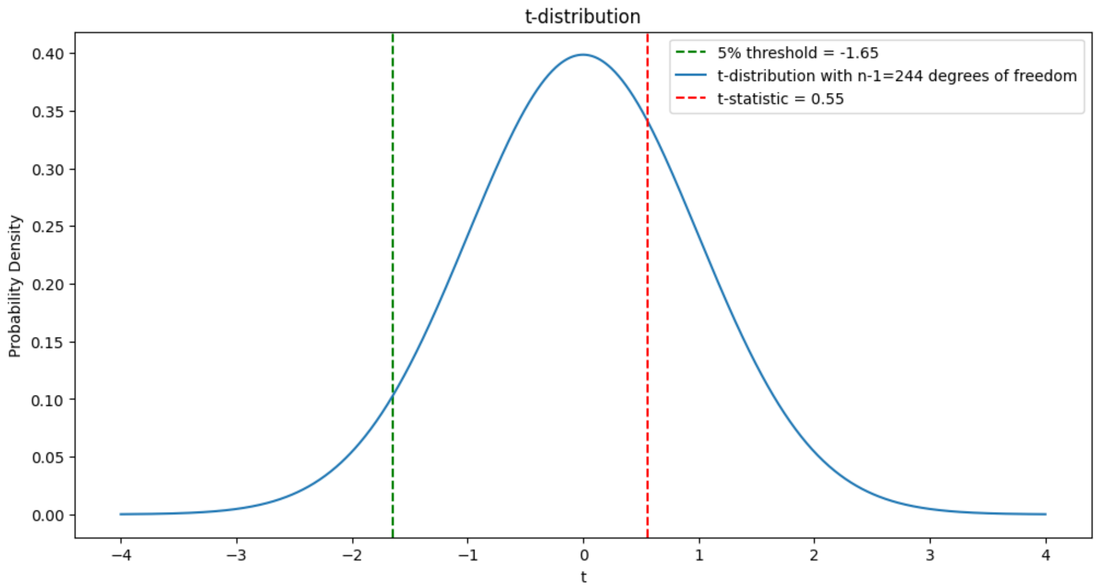

Figure 3.5: *Mid Cap t-test*

3.2 Question #2

Question

Was the economic behavior of the top large-cap companies more volatile during the **UPAII** regime over **BJPI** regime whereas the opposite is true about the top mid-cap companies?

3.2.1 Setup

Using the previous notations, let

$$[\{Z_{G(R_1,n)}\}_n, \{Z_{G(R_2,n)}\}_n]$$

represent the **Averaged Adjusted Price Deltas** (AAPD) in both the regimes. We plot histograms of these:

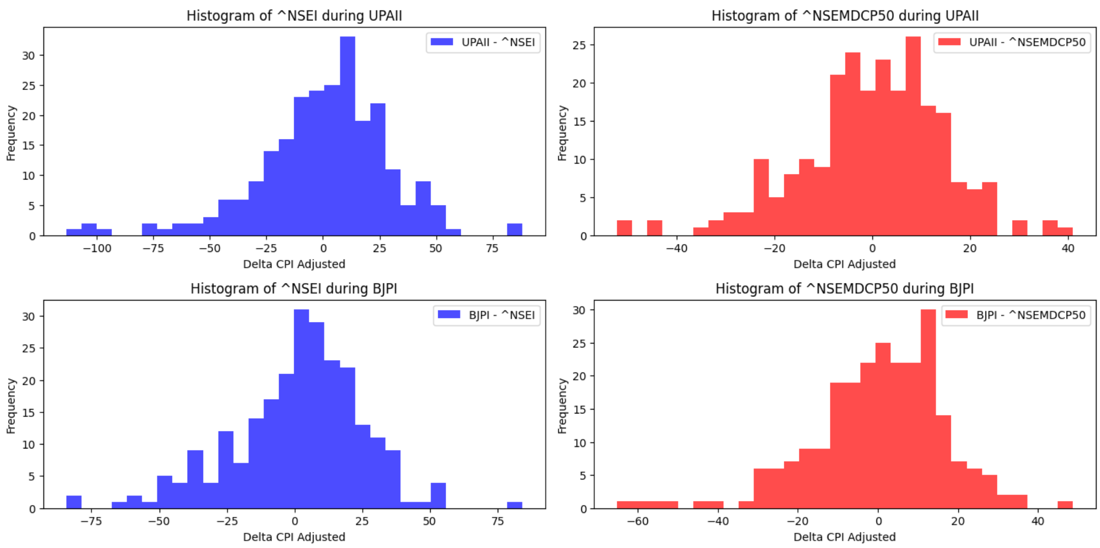

Figure 3.6: *AAPD histogram*

Now, we model $Z_{G(R_i,n)} \sim N(\mu_i, \sigma_i^2)$ with observations being **IID** upon varying n.

3.2.2 Statistical Analysis

We know that, then,

$$S_i^2 = \frac{1}{N-1} \cdot \sum_n \left(Z_{G(R_i,n)} - \text{mean}_n(Z_{G(R_i,n)})\right)^2 \sim \sigma_i^2 \cdot \chi^2(N-1)$$

We know $Z_{G(R_i),n} \sim N(\mu_i, \sigma_i^2)$. Now, it is fair to assume that $S_1 \perp S_2$.

$$H_0 : \sigma_1 = \sigma_2 \quad \text{(Null Hypothesis)}$$

Consider the following statistic:

$$F = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F(N-1, N-1)$$

Under the null hypothesis, $\frac{S_1^2}{S_2^2} = F \sim F(N-1, N-1)$.

Large Cap

$$H_0 : \sigma_1 = \sigma_2 \quad \text{(Null Hypothesis)}$$
$$H_1 : \sigma_1 > \sigma_2 \quad \text{(Alternative Hypothesis)}$$

The plot and the code is attached. Here, we find that the F-statistic is **1.47** which is, infact, greater than the **5%** threshold to the (right) of **1.24**!

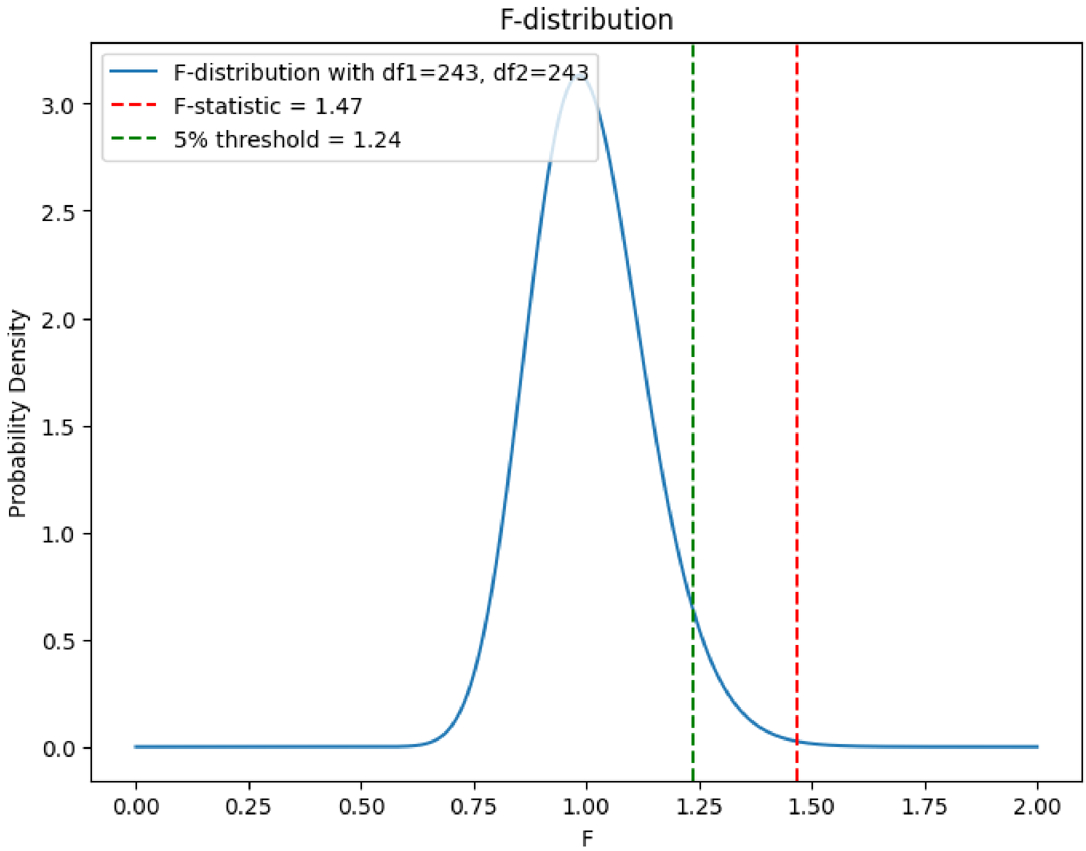

Figure 3.7: *Large Cap f-test*

There is enough evidence to reject the Null Hypothesis!

Listing 3.4: *Large Cap f-test*

```
from scipy.stats import f as f_dist

NSEI = {}
for regime in deltas_cpi_grouped:
    NSEI[regime] = deltas_cpi_grouped[regime]["^NSEI"].values
n = len(NSEI["UPAII"])

X = NSEI["UPAII"]
Y = NSEI["BJPI"]

S1_sq = 1/(n-1) * ((X - X.mean())**2).sum()
S2_sq = 1/(n-1) * ((Y - Y.mean())**2).sum()
F = S1_sq/S2_sq
# Calculate the 5% probability threshold for the F-distribution
f_threshold = f_dist.ppf(0.95, n-1, n-1)

# Plotting the F-distribution
plt.figure(figsize=(8, 6))
x = np.linspace(0, 2, 1000)
y = f_dist.pdf(x, n-1, n-1)

plt.plot(x, y, label=f'F-distribution with df1={n-1}, df2={n-1}')
plt.axvline(F, color='red', linestyle='--', label=f'F-statistic = {F:.2f}')
plt.axvline(f_threshold, color='green', linestyle='--', label=f'5% threshold =
↪   {f_threshold:.2f}')
plt.title('F-distribution')
plt.xlabel('F')
plt.ylabel('Probability Density')
plt.legend(loc="upper left")
plt.show()
```

Mid Cap

$$H_0 : \sigma_1 = \sigma_2 \quad \text{(Null Hypothesis)}$$

$$H_1 : \sigma_1 < \sigma_2 \quad \text{(Alternative Hypothesis)}$$

The plot is attached. The code is almost similar, so it's omitted. Here, we find that the F-statistic is **0.78** which is, infact, lesser than the **5%** threshold (left) of **0.81**!

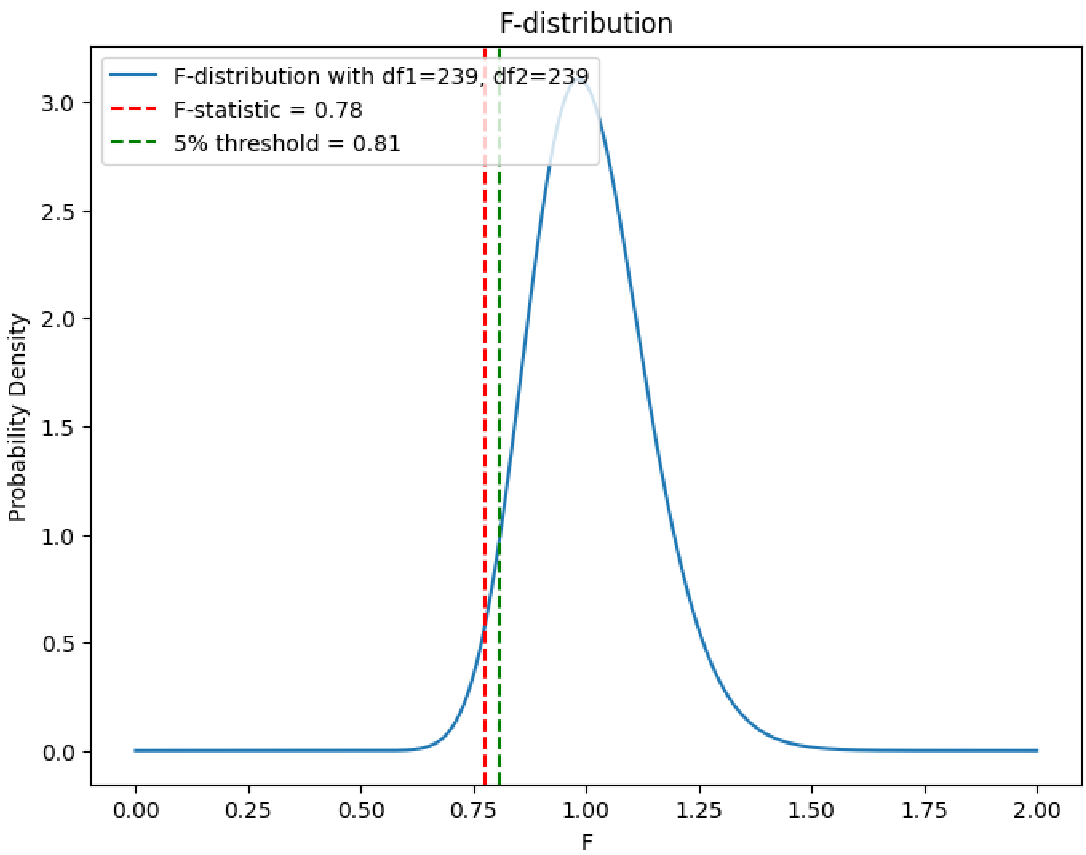

Figure 3.8: *Mid Cap f-test*

There is enough evidence to reject the Null Hypothesis!

4

Conclusion

We reach four major conclusions.

I. There is **not enough** evidence to claim that the **BJPI** regime is more favourable for the top large-cap companies than the **UPAII** regime.

II. There is **not enough** evidence to claim that the **BJPI** regime is less favourable for the top mid-cap companies than the **UPAII** regime. However,

III. There is **enough** evidence to claim that the economic behavior of the top large-cap companies was more volatile during the **UPAII** regime over the **BJPI** regime.

IV. There is **enough** evidence to claim that the economic behavior of the top mid-cap companies was more volatile during the **BJPI** regime over the **UPAII** regime.

This page intentionally left blank.

Appendices

This page intentionally left blank.

A

Appendix A

You can find the Jupyter Notebook associated with this project in my GitHub repository named **stat-3-project**.

This page intentionally left blank.