Return and volatility interrelationships between Select Sector SPDR Funds and SP500 - An Exploratory Study

February 7, 2023

Kumar Shantanu

1 Abstract

The following paper examines the return and volatility interrelationships between the select sectors of S&P500 and presents the following stylized facts: 1. A higher monthly return correlation with S&P 500 also implies a higher volatility correlation with S&P 500 for select sector ETFs. 2. The Industrial Select Sector has the maximum correlation with S&P 500 for both the monthly returns and volatility with the market. It is followed by the The Consumer Discretionary Select Sector. 3. The Utilities Select Sector has lowest correlation with S&P 500 for both the monthly returns and volatility with the market. 4. The Energy Select Sector has relatively lower correlation with all the other sectors as well as the S&P 500.

2 Definitions

2.1 Select Sector SPDR Funds

The Select Sector SPDR Funds are considered to be a type of index fund that allows for the segregation of the benchmark S&P 500 index into distinct and individualized industry sectors. This enables investors to attain partial ownership in specified groups of industries, as represented by a particular Select Sector Index. The mechanism for investing in these funds is similar to that of purchasing a stock, whereby a single share provides access to any of the ten major industry sectors that are included within the S&P 500. It is important to note that the shares of the Select Sector SPDR Funds differ significantly from those of conventional mutual funds. [1]

The S&P 500 has the following select sectors ETFs: [1]

Sector	Ticker
The Consumer Discretionary Select Sector SPDR Fund	XLY
The Consumer Staples Select Sector SPDR Fund	XLP
The Energy Select Sector SPDR Fund	XLE
The Financial Select Sector SPDR Fund	XLF
The Health Care Select Sector SPDR Fund	XLV
The Industrial Select Sector SPDR Fund	XLI
The Materials Select Sector SPDR Fund	XLB
The Real Estate Select Sector SPDR Fund	XLRE

Sector	Ticker
The Technology Select Sector SPDR Fund	XLK
The Utilities Select Sector SPDR Fund	XLU

2.2 S&P 500

We will use S&P 500 index to benchmark each of these sectors to the overall market return. Esentially, we will compare the monthly returns of each of these sectoral ETFs to the market return to gather insights about correlatiobns and volatility.

2.3 Correlation Matrix

In order to define the correlation matrix, we first need to fundamentally define the concepts of variance and covariance for a vector of random variables.[2]

2.3.1 Variance

The variance is a statistical measure of dispersion that provides information regarding the extent to which the data values deviate from the mean.

The computation of variance involves the calculation of the average of the squared deviations between each data value and the mean. In mathematical terms, the variance is obtained by calculating the expected value of the squared deviation of each data point from the mean. The mean acts as a reference point, and the squared deviation from the mean represents the distance of each data point from the central tendency of the data distribution. Thus, the variance provides an assessment of the degree of spread or dispersion of the data values around the mean.

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})$$

2.3.2 Covariance

Covariance measures the variance between two variables. In the scenario where a dataset consists of two features, the aim is to describe the relationship between the variables. In this context, the concept of covariance serves as a useful tool to measure the inter-dependence between the two variables.

The calculation of covariance is a slight modification of the equation used to compute the variance. It involves the calculation of the variance between the two variables, providing a measure of the linear relationship between the variables. By calculating the covariance, it is possible to determine the degree to which the two variables move in tandem and to what extent changes in one variable are associated with changes in the other variable. The covariance is a scalar value that summarizes the linear relationship between the two variables and provides valuable information for understanding the structure of the data.

$$C_{x,y} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

2.3.3 Covariance Matrix

The covariance matrix is a square matrix that has a symmetric structure, with each feature represented along both its rows and columns. The diagonal elements of the matrix correspond to the variance of individual features, while the non-diagonal entries represent the covariance between the different features. The covariance matrix provides a comprehensive representation of the relationships between the variables and can be used to quantify the relationship between the different features.

$$C(x,y,z) = \begin{bmatrix} \text{var}_x & \text{covar}_{x,y} & \text{covar}_{x,z} \\ \text{covar}_{y,x} & \text{var}_y & \text{covar}_{y,z} \\ \text{covar}_{z,x} & \text{covar}_{z,y} & \text{var}_z \end{bmatrix}$$

Once these covariances are standardized by dividign them with respective standard deviations, it becomes the correlation matrix.

3 Technical Methodology

The data is taken from AlphaVantage API. 'Polars' library is used to clean and manipulate data rather than Pandas because of its higher speed and efficiency compared to Pandas. Seaborn and Matblotlib are used for visualization.

3.1 Importing libraries, defining functions and classes

[ ]:
```
import requests
import polars as pl
from datetime import datetime
from dataclasses import dataclass

class getDatafromAPI ():
  """
  This class is used to get data from the Alpha Vantage API
  """
  def __init__(self) -> None:
     self._api_key = "M5VLV21IZ3B7QQDX"

  def map_dict_to_list(self, dict_obj: dict, ticker: str) -> dict:
    result = {}
    keys = list(dict_obj.keys())
    result['ticker'] = [ticker]*len(keys)
    result['date'] = [datetime.strptime(key, '%Y-%m-%d') for key in keys]
    for key in dict_obj[keys[0]]:
      result[key] = [dict_obj[k][key] for k in keys]
    return result

  def cleanData (getFunction):
      def cleaner (self, symbol, function):
```

```
            ticker, data = getFunction(self, symbol, function)
            df = pl.from_dict(self.map_dict_to_list(data, ticker)).sort('date').
 ↪with_columns(pl.exclude(['date', 'ticker']).cast(pl.Float32))
            df = df.with_columns([
                (pl.col('5. adjusted close') / pl.col('5. adjusted close').
 ↪shift(1) - 1).alias('monthlyReturn')
            ]).drop_nulls()
            return df
        return cleaner

    @cleanData
    def getAlphaVantageData (self, symbol, function):
        url = f'https://www.alphavantage.co/query?
 ↪function={function}&symbol={symbol}&apikey={self._api_key}'
        r = requests.get(url)
        response = r.json()
        return (response['Meta Data']['2. Symbol'], response['Monthly Adjusted␣
 ↪Time Series'])

@dataclass
class dataWarehouse():
    """
    This class is used to store the data from the API as a polars dataframe.
    """
    pass

getter = getDatafromAPI()
warehouse = dataWarehouse()
```

3.2 Reading the select sectors tickers from the csv file

[ ]:
```
# Read a two column CSV file into dictionary
lines = open('select_sectors.csv').read().splitlines()
select_sectors = {key:value for value, key in [line.split(",") for line in␣
 ↪lines]}
select_sectors
```

[ ]:
```
{'XLY': 'The Consumer Discretionary Select Sector SPDR Fund',
 'XLP': 'The Consumer Staples Select Sector SPDR Fund',
 'XLE': 'The Energy Select Sector SPDR Fund',
 'XLF': 'The Financial Select Sector SPDR Fund',
 'XLV': 'The Health Care Select Sector SPDR Fund',
 'XLI': 'The Industrial Select Sector SPDR Fund',
 'XLB': 'The Materials Select Sector SPDR Fund',
 'XLRE': 'The Real Estate Select Sector SPDR Fund',
 'XLK': 'The Technology Select Sector SPDR Fund',
 'XLU': 'The Utilities Select Sector SPDR Fund'}
```

3.3 Getting the data

The following code will get the data from the API and store it in the data warehouse. There is a function to handle the API limit.

```
[ ]: import time
tickers = list(select_sectors.keys())
tickers.append('SPY')
def handleProblematicTicker(ticker, wait=61):
  print(f"Unable to retrieve ticker:{ticker}")
  print(f"Retrying after {wait} seconds") #API has a limit of 5 calls per minute
  time.sleep(wait)
  setattr(warehouse, ticker, getter.getAlphaVantageData(ticker,␣
 ↪'TIME_SERIES_MONTHLY_ADJUSTED'))
```

```
[ ]: for ticker in tickers:
    print(f"Retrieving ticker:{ticker}")
    try:
        setattr(warehouse, ticker, getter.getAlphaVantageData(ticker,␣
 ↪'TIME_SERIES_MONTHLY_ADJUSTED'))
    except:
        handleProblematicTicker(ticker)
```

```
Retrieving ticker:XLY
Unable to retrieve ticker:XLY
Retrying after 61 seconds
Retrieving ticker:XLP
Retrieving ticker:XLE
Retrieving ticker:XLF
Retrieving ticker:XLV
Retrieving ticker:XLI
Unable to retrieve ticker:XLI
Retrying after 61 seconds
Retrieving ticker:XLB
Retrieving ticker:XLRE
Retrieving ticker:XLK
Retrieving ticker:XLU
Retrieving ticker:SPY
Unable to retrieve ticker:SPY
Retrying after 61 seconds
```

The Real Estate Select Sector SPDR Fund is removed from the list of tickers because it has data only from 2015.

```
[ ]: tickers.remove('XLRE') # This ticker has data beginning from 2015
```

The following code creates columns for monthly returns for each ticker and joins it to the dataframe.

```
[ ]: data = getattr(warehouse, tickers[0]).select(['date','monthlyReturn'])
for ticker in tickers[1:]:
  data = data.join(getattr(warehouse, ticker).select(['date',␣
 ↪'monthlyReturn']), on='date', how='left', suffix=f'_{ticker}')
data = data.with_columns(pl.col('monthlyReturn').
 ↪alias(f'monthlyReturn_{tickers[0]}')).select(pl.exclude(['monthlyReturn']))
```

Similarly, The following code creates columns for the volatility of monthly returns for each ticker and joins it to the dataframe. Note that the volatility is measured by 24-month rolling standard deviation.

```
[ ]: # Calculate volatility of monthly returns
for ticker in tickers:
  data = data.join(data.select(['date',f'monthlyReturn_{ticker}']).
 ↪with_columns(pl.col(f'monthlyReturn_{ticker}').rolling_std(24).
 ↪alias(f'rollingStd_{ticker}')).select(['date',f'rollingStd_{ticker}']).
 ↪select(['date',f'rollingStd_{ticker}']), on="date")
```

The following code separates the volatility and returns into two dataframes.

```
[ ]: AdjClose =  getattr(warehouse, tickers[0]).select(['date','5. adjusted close'])
for ticker in tickers[1:]:
  AdjClose = AdjClose.join(getattr(warehouse, ticker).select(['date', '5.␣
 ↪adjusted close']), on='date', how='left', suffix=f'_{ticker}')
AdjClose = AdjClose.with_columns(pl.col('5. adjusted close').alias(f'5.␣
 ↪adjusted close_{tickers[0]}')).select(pl.exclude(['5. adjusted close']))
```

```
[ ]: volatility = data.select(['date']+[f"rollingStd_{ticker}" for ticker in␣
 ↪tickers]).drop_nulls()
```

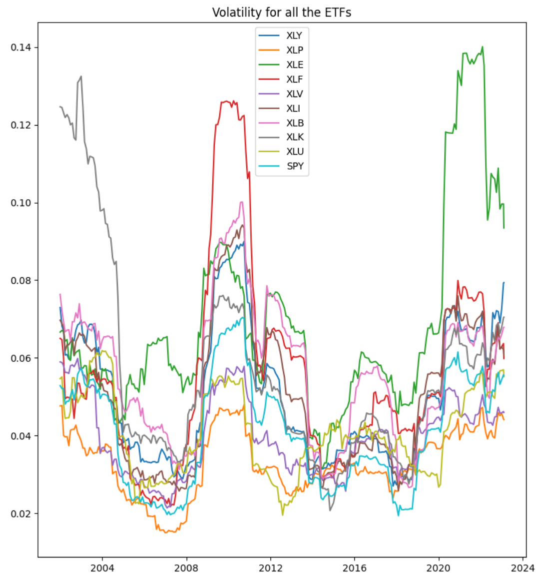
```
[ ]: monthlyReturns = data.select(['date']+[f"monthlyReturn_{ticker}" for ticker in␣
 ↪tickers]).drop_nulls()
```

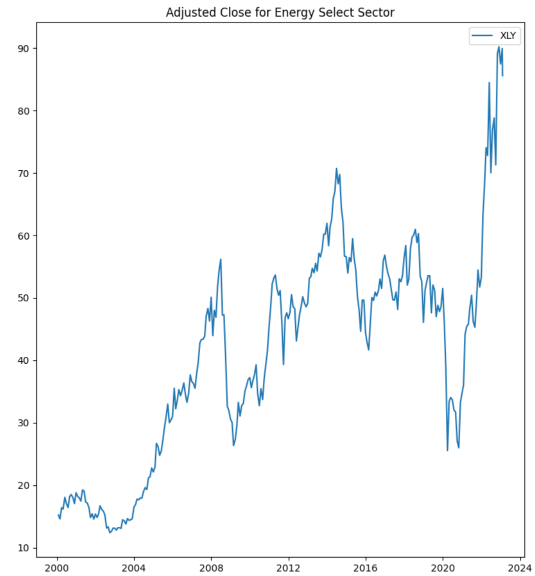
4 Insights from Data

Next, we plot how the volatility of the monthly returns has changed over time.

```
[ ]: import matplotlib.pyplot as plt
fig, ax = plt.subplots(1,2,figsize=(20,10))
ax[0].plot(volatility['date'], volatility.select(pl.exclude('date')))
ax[1].plot(monthlyReturns['date'], AdjClose.select('5. adjusted close_XLE'))
ax[0].legend(tickers)
ax[1].legend(tickers)
ax[0].set_title('Volatility for all the ETFs')
ax[1].set_title('Adjusted Close for Energy Select Sector')
```

```
[ ]: Text(0.5, 1.0, 'Adjusted Close for Energy Select Sector')
```



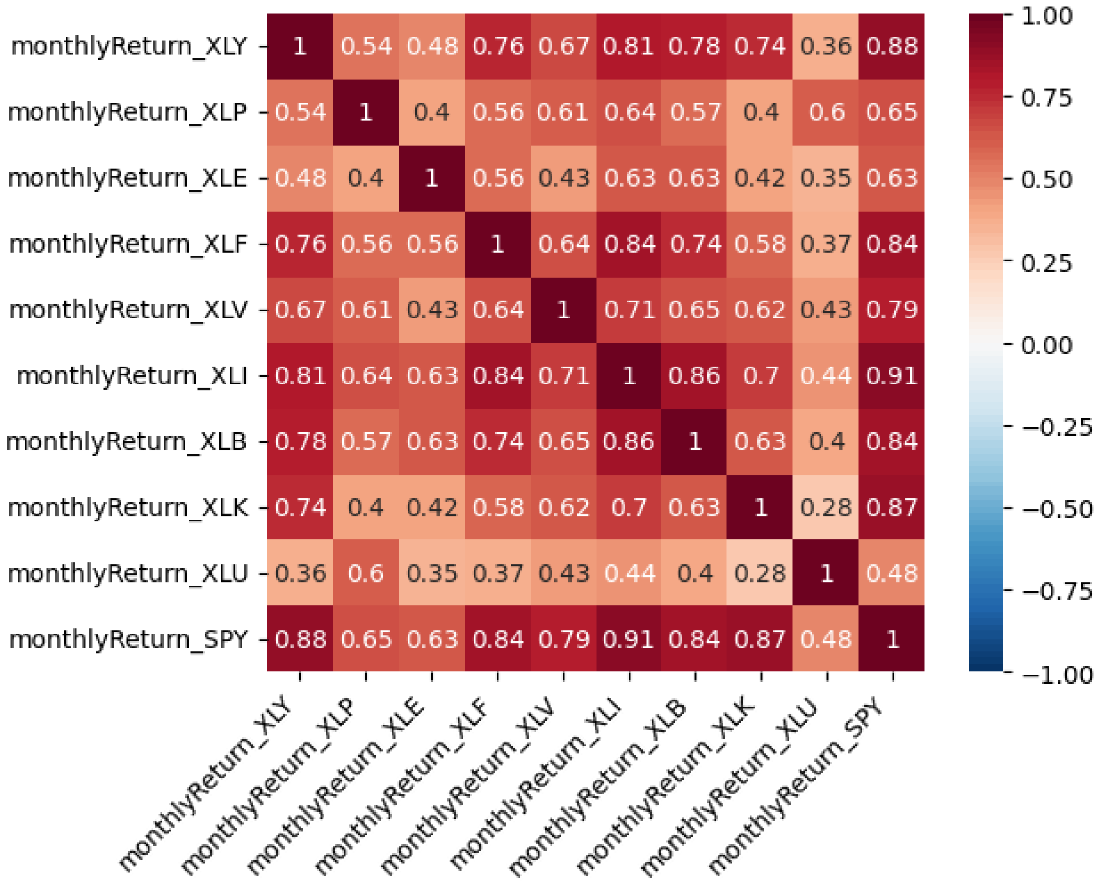


The first apparent thing is that there is a increase in volatility in the last 2 years. We find a phenomenal increase in volatility for the Energy Select Sector during the recent times. There is also an increase in the prices of energy select sector.

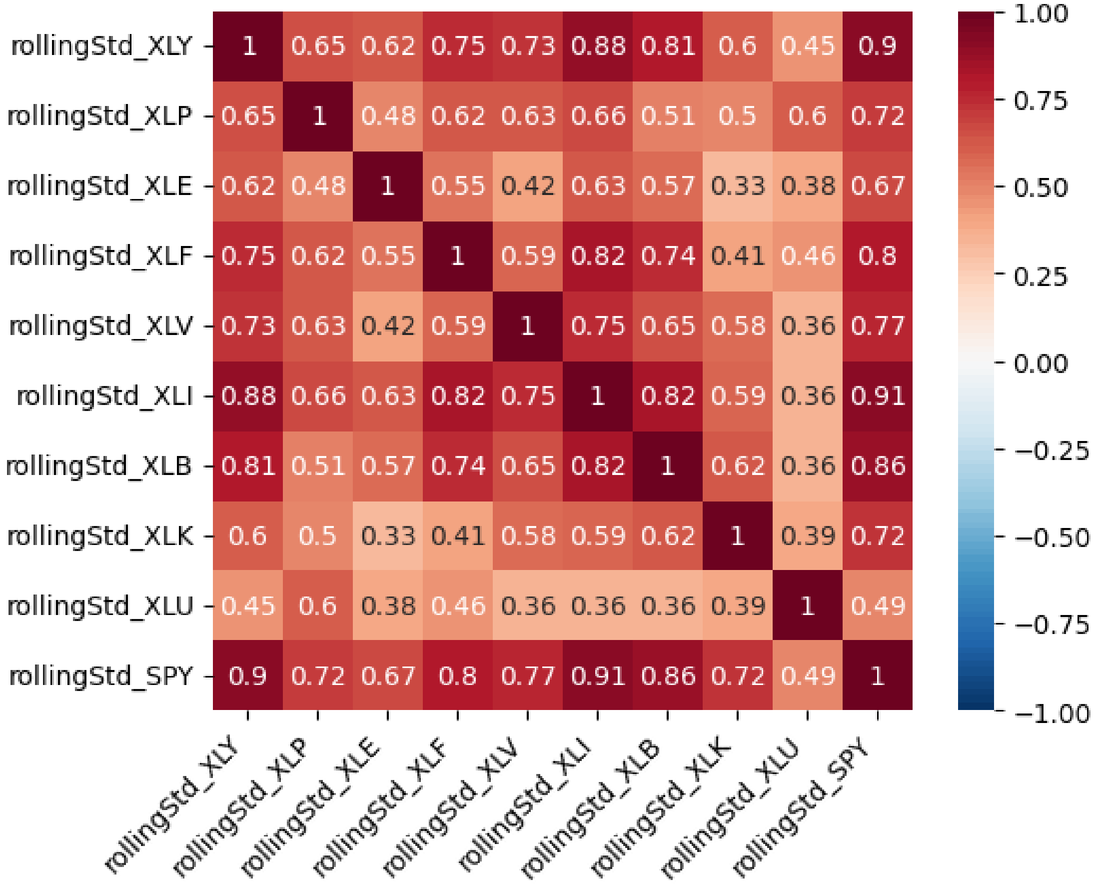
Now let us examine the correlations by plotting a corelation heatmap.

[ ]:
```
import seaborn as sns

def draw_correlation_heatmap(data, title, removeDate=True):
    data = data.select(pl.exclude('date')).pearson_corr()
    ax = sns.heatmap(
        data,
        vmin=-1, vmax=1, center=0,
        cmap=sns.color_palette("RdBu_r", 100),
        square=True,
        annot=True,
    )

    ax.set_yticklabels(
        data.columns,
        rotation=360,
    )


    ax.set_xticklabels(
        data.columns,
        rotation=45,
        horizontalalignment='right')
```

```
[ ]: draw_correlation_heatmap(monthlyReturns, "Correlation of monthly returns")
```




```
[ ]: draw_heatmap(volatility.select(pl.exclude('date')).pearson_corr(), "Volatility␣
     ↪Correlation")
```



We find that: 1. A higher monthly return correlation with S&P 500 also implies a higher volatility correlation with S&P 500 for select sector ETFs. 2. The Industrial Select Sector has the maximum correlation with S&P 500 for both the monthly returns and volatility with the market. It is followed by the The Consumer Discretionary Select Sector. 3. The Utilities Select Sector has lowest correlation with S&P 500 for both the monthly returns and volatility with the market. 4. The Energy Select Sector has relatively lower correlation with all the other sectors as well as the S&P 500.

4.1 Citations

1. "THE SELECT SECTOR SPDR® TRUST." Https://www.Sec.Gov/, 30 Sept. 2017, www.sec.gov/. Accessed 5 Feb. 2023.

2. Reichel, L. (n.d.). Random Vectors and the Variance–Covariance Matrix. Retrieved February 7, 2023, from http://www.math.kent.edu/~reichel/courses/monte.carlo/alt4.7d.pdf

Leverage and non-linearity are two risk characteristics that can exacerbate the impact of financial market movements on investment performance.

The use of borrowed funds to improve the possible return on an investment is referred to as leverage. This implies that an investor may manage a greater position in an asset with less cash, thereby boosting returns if the investment performs well. However, if the investment does not perform well, leverage can multiply losses and increase the danger of default. Leverage enables investors to control a greater position in an asset with less capital. If the investment performs well, this can multiply prospective benefits since the investor can make a higher profit on a lesser investment. For example, if an individual utilizes leverage to buy $100,000 in stock using $10,000 of their own money and $90,000 borrowed, a 10% gain in stock price results in a 100% return on investment.

Leverage, on the other hand, multiplies losses. If the investment underperforms and the stock's value falls, the investor would lose not just their $10,000 investment but also the $90,000 they borrowed. As a result, the investor may be obliged to sell their stake at a loss in order to fulfill margin calls or repay borrowed funds. A minor fall in the value of the assets in a highly leveraged venture can result in huge losses, potentially leading to bankruptcy.

Leverage can be defined mathematically as the ratio of debt to equity, or the amount of borrowed funds used to finance investments. The formula for leverage is:

Leverage = Total Debt / Total Equity

For example, if a company has $10 million in total debt and $20 million in total equity, its leverage ratio would be:

Leverage = $10 million / $20 million = 0.5

This means that the company has used 50% borrowed funds to finance its investments.
A diagram that shows the concept of leverage can be a simple bar graph, where the total debt and total equity are represented by the height of two bars. The higher the ratio of debt to equity, the higher the leverage and the greater the risk of financial instability.

Leverage may therefore dramatically increase an investment's risk, and it is critical for investors to recognize and manage their leverage exposure. To restrict leverage and lessen the danger of financial instability, financial institutions and regulators have developed different measures such as margin requirements. Leverage and its possible implications are taken seriously by financial authorities and regulators, as leverage may compound risk and lead to financial instability. As a result, they have developed a number of steps to restrict leverage and mitigate the dangers connected with it. Among these measures are:

1. Margin Requirements: Financial institutions must set aside a percentage of their investment as collateral to protect themselves from potential losses. This decreases the risk of default by limiting the amount of leverage that may be employed in an investment.
2. Risk-Based Capital Requirements: Regulators require financial institutions to maintain a certain level of capital to cover possible losses, which is defined by the riskiness of their assets. This ensures that financial institutions can sustain losses in the case of market downturns or other unfavorable events.
3. Stress Testing: Regulators perform stress tests on a regular basis to assess financial institutions' capacity to absorb possible losses in the event of a market downturn or other adverse event. This aids in identifying and addressing any flaws in the financial system, reducing the risk of instability.
4. Restrictions on Proprietary Trading: Some authorities have implemented restrictions on proprietary trading by financial institutions, which refers to the practice of trading in financial markets using the institution's own funds. This minimizes the amount of leverage employed by financial institutions and the danger of instability.
5. Disclosure Requirements: Financial institutions are required by regulators to publish information about their leverage, risk exposure, and other financial parameters, which serves to boost transparency and strengthen market discipline.

These are just a few examples of the measures taken by financial authorities and regulators to limit leverage and reduce its associated risks. By implementing these measures, authorities aim to promote stability in the financial system and reduce the risk of systemic failures.

Non-linearity, on the other hand, refers to a non-proportional relationship between changes in an investment's underlying assets and its returns. This means that a little change in the underlying assets' value might cause a disproportionately significant change in the investment's returns. Non-linearity may be found in a variety of financial instruments, including options, futures, and some derivatives. Consider a call option, which grants the holder the right to purchase an underlying asset at a specific price (the strike price) within a given time period. The value of a call option is determined by the difference between the underlying asset's current price and the strike price. A small change in the price of the underlying asset can result in a large change in the value of the call option, creating non-linearity.

Non-linearity may also be observed in financial products that monitor the performance of numerous underlying assets, such as exchange-traded funds (ETFs) or stock portfolios. If the underlying assets are not completely linked, modest changes in one asset's value can cause huge changes in the portfolio's returns, resulting in non-linearity. When many underlying assets are not perfectly linked, minor changes in the value of one asset can result in huge changes in the portfolio's returns due to the way the assets interact with one another. If the assets are positively correlated, or move in the same direction, then a slight increase in the value of one asset is likely to be matched by a comparable increase in the value of another asset, and the total impact on the portfolio's returns is likely to be minimal. If the assets are negatively correlated, which means they tend to move in opposing directions, a little gain in the value of one asset may be accompanied by a considerably bigger decline in the value of another, resulting in non-linearity.

Furthermore, the weighting of each item in the portfolio might influence the amount of nonlinearity. If a certain asset constitutes a big amount of the portfolio, then modest changes in its value will have a greater influence on the portfolio's returns, resulting in non-linearity. This may be especially troublesome if the underlying assets have varying levels of volatility, which means they might suffer substantial changes in value, since this can raise the portfolio's total volatility and compound the impact of non-linearity. Non-linearity in stock or ETF portfolios occurs when the underlying assets are not perfectly linked and can result in substantial variations in the portfolio's returns owing to asset interactions and the weighting of each asset in the portfolio. This emphasizes the need of carefully considering the underlying assets and their linkages when building portfolios.

To represent non-linearity mathematically, one can use functions or equations that show the relationship between inputs and outputs. For example, a simple non-linear equation is the quadratic function:

$y = ax^2 + bx + c$

where y is the output, x is the input, and a, b, and c are constants. This equation shows how the output changes disproportionately with respect to the input, as x increases.
A diagram that represents non-linearity can be a graph of the function, where the x-axis represents the input and the y-axis represents the output. The graph will show a curved shape, indicating that the output changes non-linearly with respect to the input.

In order to mitigate the risk of instability, financial authorities monitor non-linearity in financial markets and take efforts to prevent the use of financial instruments with excessive non-linearity. 1 Increasing capital requirements: To decrease risk exposure, regulators have increased the amount of capital that financial institutions must retain. This helps to guarantee that institutions have enough capital to sustain losses in the case of a market collapse.

1. Leverage restrictions: Regulators have also restricted the amount of leverage that financial organizations can utilize. This reduces the magnitude of potential losses in the case of market volatility.
2. Improving risk management techniques: Regulators have also urged financial institutions to strengthen their risk management processes, such as stress testing and scenario analysis, in order to analyze the effect of market disruptions.
3. Outright prohibition of certain financial products: In some circumstances, authorities have outright prohibited certain financial instruments judged overly hazardous, such as naked credit default swaps.
4. Increasing openness: Regulators have also attempted to increase financial market transparency by demanding better disclosure of information about financial products and their underlying hazards.

Real-life Applications

The papers discuss the challenges of regulating non-linearity and leverage in finance, and argue that reducing leverage is a critical step in mitigating systemic risk. The authors also argue that regulation should focus on limiting the use of non-linear financial products and addressing the non-linearities that can contribute to systemic risk, such as the interplay between banks and sovereigns.

Acharya and Richardson's (2009) book "The Causes of the Financial Crisis" investigates the impact of non-linearity and leverage in the 2008 financial crisis. The interaction between non-linear financial products, such as mortgage-backed securities, and high levels of leverage in the financial system, according to the author, contributed to the crisis by amplifying the effects of losses and creating a feedback loop of declining asset values and tightening credit. On top of that, banks had failed to meet regulatory capital requirements. First, they had temporarily put assets, like as securitized mortgages, in offbalancesheet organizations in order to avoid holding hefty capital buffers against them. Second, the capital laws permitted banks to cut the amount of capital they kept against assets that remained on their balance sheets, as long as those assets were AAArated tranches of securitized mortgages. Thus, by repackaging mortgages into mortgagebacked securities, whether kept on or off their balance sheets, banks lowered the amount of capital required against their loans, multiplying their ability to issue loans by a factor of ten. However, the main result of this regulatory arbitrage was to concentrate the risk of mortgage defaults in the banks, rendering them insolvent when the housing bubble burst.

Atif Mian, Amir Sufi, and Francesco Trebbi wrote "Leverage Cycles and Financial Crises" (2013). Using data on US banks from 1920 to 2006, this research investigates the association between leverage and financial crises. The authors discover that leverage cycles, or periods of fast growth followed by quick reduction in leverage, are a typical feature of financial crises and that the accumulation of leverage can enhance the severity of financial crises.

Viral V. Acharya, Lasse H. Pedersen, Thomas Philippon, and Matthew Richardson wrote "Leverage and Systemic Risk" (2011). The link between leverage and systemic risk, or the danger that a collapse in one financial institution would spread to others and disrupt the whole financial system, is investigated in this research. The authors discover that financial institutions with high levels of leverage are more likely to go bankrupt and contribute to systemic risk, and that the impact of leverage on systemic risk is non-linear, with greater effects at higher levels of leverage.

Eduardo Schwartz's "The Dark Side of Optionality: Optimal Financial Contracts with Hidden Actions" (2010). The influence of non-linear financial contracts, such as options, on financial stability is investigated in this research. The author contends that the optionality built in these contracts might lead to concealed actions that disrupt the financial system, and that this risk can be mitigated by restricting their usage or putting greater capital requirements on institutions that utilize them.

Hélène Rey's "Dynamic Interactions between Banks and Sovereigns: The Case of Systemic Risk" (2015). This study looks at how banks and sovereigns interact to contribute to systemic risk, or the danger that a collapse in one financial institution would spread to others, destabilizing the whole financial system. The author contends that dynamic interactions between banks and sovereigns, such as the employment of banks to fund government debt, can lead to non-linearities in the financial system, amplifying the consequences of losses and contributing to systemic risk.

"Systemic Risk and Financial System Regulation," by Anat Admati and Martin Hellwig (2013). This study examines the significance of leverage and systemic risk in the financial system, as well as the difficulties in controlling these risks. According to the authors, excessive levels of leverage raise the risk of systemic instability, and lowering leverage is a vital step in minimizing systemic risk. They also suggest that rather than just forcing financial firms to maintain additional capital, regulation should be geared to minimize the use of leverage.

Claudia Buch, Martin Hellwig, and Isabel Schnabel wrote "Macroprudential Policy in the Aftermath of the Crisis" (2017). This article examines the lessons learned from the 2008 financial crisis and the policy responses that followed for the creation of macroprudential policy. According to the authors, macroprudential policy should focus on restricting the use of leverage in the financial system and tackling nonlinearities that might contribute to systemic risk, such as the interaction between banks and sovereigns.