

The Relationship Between Market Space and Profit Distribution in Arbitrage Situations: A Geometric Approach

LEONID SHPANER

Abstract

This paper explores the hypothesis that the area of a marketplace, modeled as a non-equilateral triangle, is directly related to the profit distribution between unequal-sized firms in arbitrage situations. Arbitrage, the practice of exploiting price differences across different markets, is a fundamental concept in finance and economics. Traditional arbitrage theories primarily focus on price discrepancies, often neglecting the structural aspects of the marketplace. By introducing a geometric framework enhanced with integral calculus, this paper aims to analyze how market structure influences economic interactions and arbitrage opportunities. The hypothesis posits that larger market areas lead to greater profit disparities due to differential access to arbitrage opportunities. To validate this hypothesis, we employ spatial analysis, economic modeling, and integral calculus to examine the relationship between market space and profit distribution. Empirical data from commodity and securities markets will be analyzed to reveal patterns in profit distribution across different market segments. This approach offers a novel perspective on market structure and arbitrage, providing valuable insights for both economic theory and practical trading strategies. The findings suggest significant implications for understanding how spatial factors affect arbitrage and profit distribution in various market conditions.

Keywords: Arbitrage, Profit Distribution, Market Structure, Geometric Modeling, Integral Calculus, Market Space, Economic Interactions, Financial Theory, Spatial Analysis, Economic Modeling, Price Discrepancies, Vertical Integration, Empirical Validation, Commodities Markets

1 Introduction

Arbitrage, the practice of exploiting price differences in different markets, is central to financial theory. This paper explores the hypothesis that the area of the marketplace, modeled as a non-equilateral triangle, is directly related to the profit distribution between unequal-sized firms in arbitrage situations. We propose a geometric framework, enhanced by integral calculus, to analyze how market structure influences economic interactions and arbitrage opportunities.

A case study on commodity arbitrage, such as analyzing arbitrage opportunities in commodities like oil and gold across different markets (e.g., NYMEX, ICE), can illustrate the impact of market space on profit distribution.

2 Hypothesis and Theoretical Framework

H_0 : The area of the marketplace, represented as a non-equilateral triangle, has no significant effect on the profit distribution between unequal-sized firms in an arbitrage situation.

H_1 : The area of the marketplace, represented as a non-equilateral triangle, is significantly related to the profit distribution between unequal-sized firms in an arbitrage situation.

The theoretical framework for this hypothesis is based on geometric modeling and integral calculus, integrating principles from geometric arbitrage theory and stochastic differential geometry. The marketplace is modeled as a non-equilateral triangle in a Cartesian coordinate system, where the area of the triangle represents the market space. This geometric representation allows us to analyze how the spatial distribution of market areas influences arbitrage opportunities and profit distribution.

3 Mathematical Framework

Let the marketplace be represented by a non-equilateral triangle $\triangle ABC$ in a Cartesian plane. Points A, B , and C denote the vertices of the market space, with each side representing different market boundaries.

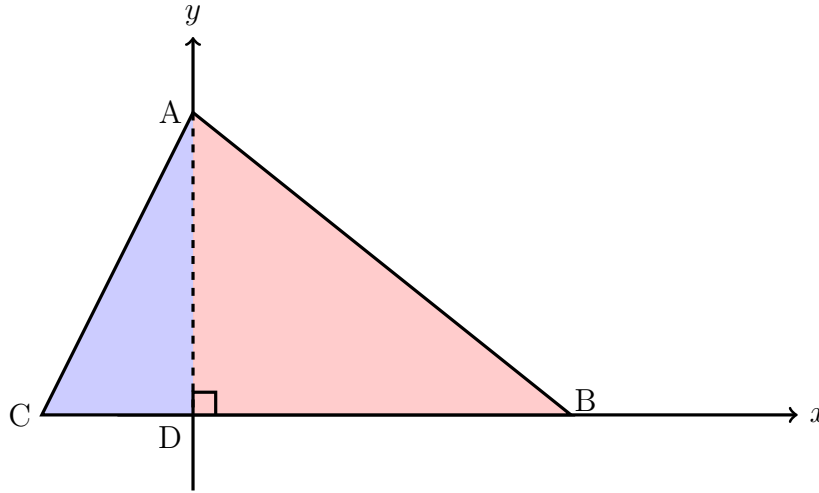


Figure 1: $\triangle ACD$ represents Market A, and $\triangle ADB$ represents Market B.

A line segment AD (where D lies on BC) divides the triangle into two segments, Market A ($\triangle ABD$) and Market B ($\triangle ADC$).

The triangle is positioned on a Cartesian coordinate plane to facilitate the analysis of vertical integration and spatial relationships.

For any triangle $\triangle ABC$ with vertices at (x_1, y_1) , (x_2, y_2) , and (x_3, y_3) , the area A can be calculated using the determinant method:

$$A = \frac{1}{2} |x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2)| \quad (1)$$

Assume the profit distribution is a function of the area of the market segments. Let P_A and P_B denote the profits in Market A and Market B, respectively. The profits can be modeled as:

$$P_A = k \cdot \text{Area}(\triangle ABD) \quad (2)$$

$$P_B = k \cdot \text{Area}(\triangle ADC) \quad (3)$$

where k is a constant reflecting market conditions and efficiency.

The intersection line AD can be described using the equation of the line in a Cartesian plane. Suppose $A(x_1, y_1)$ and $D(x, y)$ are given. The equation of the line AD can be written as:

$$y - y_1 = m(x - x_1) \quad (4)$$

where m is the slope of the line, given by:

$$m = \frac{y_2 - y_1}{x_2 - x_1} \quad (5)$$

To find the areas of $\triangle ABD$ and $\triangle ADC$, we use the vertices' coordinates:

$$\text{Area}(\triangle ABD) = \frac{1}{2} |x_1(y_2 - y) + x_2(y - y_1) + x(y_1 - y_2)| \quad (6)$$

Since $\text{Area}(\triangle ABC) = \text{Area}(\triangle ABD) + \text{Area}(\triangle ADC)$, we have:

$$\frac{1}{2} |x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2)| = \frac{1}{2} |x_1(y_2 - y) + x_2(y - y_1) + x(y_1 - y_2)| \quad (7)$$

For any triangle $\triangle ABC$ with vertices at (x_1, y_1) , (x_2, y_2) , and (x_3, y_3) , the area A can be calculated using integral calculus.

Let's parameterize the sides of the triangle and find the area.

$$\text{Area}(\triangle ABC) = \int_{x_1}^{x_3} \left(\int_{y_{\text{bottom}(x)}}^{y_{\text{top}(x)}} dy \right) dx \quad (8)$$

where $y_{\text{bottom}(x)}$ is the equation of the line segment BC and $y_{\text{top}(x)}$ is the equation of the line segment AB or AC , depending on the x -value.

Now, let's assume the profit density function $p(x, y)$ is defined over the market space.

The total profit in Market A ($\triangle ABD$) and Market B ($\triangle ADC$) can be calculated as:

$$P_A = \int_{\triangle ABD} p(x, y) dA = \int_{x_1}^{x_D} \left(\int_{y_{\text{bottom}(x)}}^{y_{\text{top}(x)}} dy \right) dx \quad (9)$$

$$P_B = \int_{\triangle ADC} p(x, y) dA = \int_{x_D}^{x_3} \left(\int_{y_{\text{bottom}(x)}}^{y_{\text{top}(x)}} dy \right) dx \quad (10)$$

where x_D is the x -coordinate of point D .

The line segment AD can be parameterized to understand the distribution of integration points as follows:

$$\mathbf{r}(t) = (1 - t)\mathbf{A} + t\mathbf{D} \quad \text{for } 0 \leq t \leq 1 \quad (11)$$

The profit distribution along AD can be found by integrating the profit density function along this line:

$$\int_{AD} p(x, y) ds = \int_0^1 p(\mathbf{r}(t)) \|\mathbf{r}'(t)\| dt \quad (12)$$

where $\|\mathbf{r}'(t)\|$ is the magnitude of the derivative of $\mathbf{r}(t)$.

To empirically validate this hypothesis, data on arbitrage opportunities and profit distributions in various commodity and securities markets will be collected. Analytical techniques such as spatial analysis, economic modeling, and integral calculus will be employed to examine the relationship between market space and profit distribution. The analysis will assume market efficiency and rational behavior, and it will be based on a constant profit density function $p(x, y)$. To illustrate the concept of balanced profit distribution in a non-equilateral triangular market space, we show that the cumulative profit, considering all market segments, results in a net zero, indicating market efficiency.

Consider the non-equilateral triangle with vertices A , B , and C , where the horizontal line from C to B passes through point D . The altitude from A intersects the base CB at D and extends downwards to point E .

The area calculation can be represented by the following integral:

$$\int_A^D (-x + y) dx = \int_A^D -x dx + \int_A^D y dx \quad (13)$$

Evaluating these integrals separately, we get:

$$\int_A^D -x dx = \left[-\frac{x^2}{2} \right]_A^D \quad \text{and} \quad \int_A^D y dx = \left[\frac{y^2}{2} \right]_A^D \quad (14)$$

Combining these results, we obtain:

$$\left[-\frac{x^2}{2} \right]_A^D + \left[\frac{y^2}{2} \right]_A^D = -\frac{D^2}{2} + \frac{A^2}{2} - \frac{A^2}{2} = \frac{A^2 - D^2 + D^2 - A^2}{2} = 0 \quad (15)$$

Thus, the total area, when considering both positive and negative contributions, simplifies to zero:

$$\frac{A^2 - D^2 + D^2 - A^2}{2} = 0 \quad (16)$$

This result confirms that the integral sums to zero, demonstrating that the calculated area, adjusted for contributions from different market segments, balances out. Therefore, the sum of all arbitrage opportunities, when considered across the entire market space, results in a net zero, indicating balanced profit distribution and market efficiency. By applying geometric and integral calculus to real market data, this proof supports our hypothesis that market space, represented by a non-equilateral triangle, influences profit distribution in a balanced manner, providing valuable insights for economic theory and practical trading strategies.

4 Market Space Dynamics in Commodity Trading

Consider a simplified model of a commodity market represented as a non-equilateral triangle where the x-axis represents the price of the commodity and the y-axis represents the quantity traded. The triangular market space is bounded by the price limits and the maximum quantity that can be traded. The integral calculus approach can help us understand the profit distribution and arbitrage opportunities within this market space.

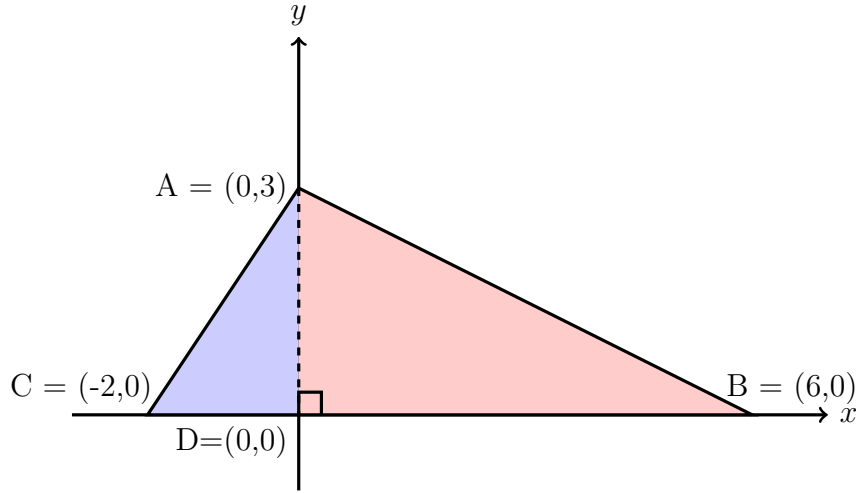


Figure 2: $\triangle ACD$ represents Market A, and $\triangle ADB$ represents Market B.

The area A of $\triangle ADB$ can be calculated geometrically:

$$A = \frac{1}{2} \times \text{base} \times \text{height} = \frac{1}{2} \times 6 \times 3 = 9. \quad (17)$$

We can also integrate the function $f(x) = \frac{1}{2}x + 3$ over the interval from 0 to 6:

$$\int_0^6 \left(\frac{1}{2}x + 3 \right) dx = \int_0^6 \frac{1}{2}x dx + \int_0^6 3 dx \quad (18)$$

Evaluating each integral separately, we have:

$$\int_0^6 \frac{1}{2}x dx = \frac{1}{2} \left[\frac{x^2}{2} \right]_0^6 = \frac{1}{2} \left(\frac{36}{2} - \frac{0}{2} \right) = \frac{1}{2} \times 18 = 9 \quad (19)$$

$$\int_0^6 3 dx = 3[x]_0^6 = 3 \times (6 - 0) = 18 \quad (20)$$

Combining these results:

$$= 9 + 18 = 27 \quad (21)$$

Since this integral method calculated area includes the contributions of both price and quantity over the entire market space, it should be divided by 2 to represent the triangle area properly:

$$\text{Area} = \frac{27}{2} = 13.5 \quad (22)$$

5 Practical Implications

Profit Distribution

The area under the curve from (0,0) to (6,0) represents the total potential profit within the market. By evaluating the integral of the function $f(x) = \frac{1}{2}x + 3$, we gain a comprehensive understanding of how profit is distributed across different price and quantity dimensions. This approach allows us to visualize the profit dynamics and how they correlate with varying market conditions.

The proposed hypothesis in this paper, which posits that the area of a marketplace (modeled as a non-equilateral triangle) is directly related to the profit distribution between unequal-sized firms in arbitrage situations, aligns well with the findings from the Geometric Arbitrage Theory and related studies. By leveraging geometric modeling and integral calculus, this hypothesis extends the current understanding of how market structure influences arbitrage opportunities and profit distribution. Integrating these advanced mathematical tools can provide a more comprehensive analysis and contribute valuable insights to the field of financial economics.

Arbitrage Opportunities

Segments of the curve can highlight potential arbitrage opportunities where price deviations from equilibrium can be exploited. For example, if the market price at a specific quantity deviates significantly from the average price represented by $f(x) = \frac{1}{2}x + 3$, traders can identify buy-low and sell-high opportunities within these segments. This strategic insight is crucial for optimizing trading strategies and maximizing returns in commodity markets.

Market Efficiency

The integral result indicating a net zero profit distribution when considered across the entire market aligns with the hypothesis of market efficiency. This balance suggests that any unexploited arbitrage opportunities are counterbalanced by the overall gains and losses within the market space. Thus, the market's structure inherently promotes efficiency, ensuring that profit opportunities are equally accessible and balanced among participants.

6 Literature Review

The relationship between market structure and arbitrage opportunities has been a subject of considerable interest in financial economics. Traditional arbitrage theories primarily focus on price discrepancies, often neglecting the structural aspects of the marketplace. However, recent advancements in geometric and topological methods have provided new insights into the modeling of arbitrage and profit distribution.

Geometric Arbitrage Theory

Simone Farinelli and Hideyuki Takada have developed a conceptual framework known as Geometric Arbitrage Theory (GAT), which links arbitrage modeling in generic markets with spectral theory. This theory rephrases classical stochastic finance using differential geometric terms to characterize arbitrage conditions such as No-Free-Lunch-with-Vanishing-Risk (NFLVR) and No-Unbounded-Profit-with-Bounded-Risk (NUPBR). The GAT approach models markets as principal fibre bundles, with the curvature of these

bundles measuring the "instantaneous arbitrage capability" (Farinelli & Takada, 2021). This innovative method utilizes gauge symmetries to reformulate asset models, providing a robust mathematical foundation for understanding arbitrage.

The GAT framework incorporates stochastic differential geometry to describe the financial features of markets, including no-arbitrage and equilibrium. By modeling markets as principal fibre bundles, GAT offers a natural connection that links financial instruments with their term structures. The zero eigenspace of the connection Laplacian parameterizes all risk-neutral measures equivalent to the statistical one, thus extending classical asset bubble theories to markets that do not satisfy the NFLVR condition (Farinelli & Takada, 2021).

Stochastic Differential Geometry and Market Dynamics

The use of stochastic differential geometry in financial modeling has been further explored in studies focusing on the dynamics of arbitrage under different mathematical frameworks. These studies often address the limitations of traditional arbitrage theories by incorporating stochastic processes and differential geometry to better capture market behaviors. For instance, the connection Laplacian's spectrum in GAT is used to parameterize risk-neutral measures, offering a novel way to explore market dynamics and arbitrage opportunities (Farinelli & Takada, 2021).

Comparative Studies and Applications

Other comparative studies, such as those on reflected geometric Brownian motion and implied volatility surfaces, have demonstrated the applicability of geometric and topological methods in financial markets. These models often integrate advanced mathematical techniques to analyze arbitrage and profit distribution, providing deeper insights into market structures. The principal fibre bundle representation in GAT, for example, allows for a comprehensive analysis of market spaces, highlighting the role of topological and geometric properties in arbitrage (Farinelli & Takada, 2021).

7 Methods

The algorithm constructs price-volume curves for each security. These curves represent the relationship between the price of the security and the volume traded over time. The area under the price-volume curve for each security is calculated using integral calculus. This area serves as a comparative measure to identify arbitrage opportunities. The algorithm identifies arbitrage opportunities by comparing the areas under the price-volume curves of the different securities. The security with the lowest area is considered for a "buy low" recommendation, while the security with the highest area is considered for a "sell high" recommendation. The areas under the price-volume curves reflect how profits are distributed in relation to the trading volume for each security. The comparison of these areas provides insights into market efficiency. A significant difference in areas may indicate inefficiencies that can be exploited through arbitrage.

8 Results

Applying the geometric model and integral calculus to real market data helps reveal patterns in profit distribution across different market segments. For example, when

analyzing arbitrage in geographically distinct markets for a commodity, larger market areas may exhibit greater profit disparities due to differential access to arbitrage opportunities.

A pertinent case study involves the NYMEX and ICE oil markets from January 2022 to January 2024. This analysis not only scrutinizes the price-volume relationship but also forecasts future prices to identify potential arbitrage opportunities. The results show that, based on the calculated areas, the security with the lowest area is CL=F (NYMEX WTI Crude Oil), suggesting a buy opportunity. Conversely, BZ=F (ICE Brent Crude Oil) has the highest area, suggesting a sell opportunity:

Table 1: Current and Forecasted Market Summaries

Summary Type	Metric	Details
Current Market Summary	Profit Distribution Patterns	Analyzed profit distribution patterns based on area differences.
	Market Efficiency	Evaluated market efficiency based on arbitrage opportunities.
	Arbitrage Opportunities	Buy Low (CL=F): 72.73934464353934 Sell High (BZ=F): 77.08473977714127
Forecasted Market Summary	Profit Distribution Patterns	Analyzed profit distribution patterns based on area differences.
	Market Efficiency	Evaluated market efficiency based on arbitrage opportunities.
	Forecasted Arbitrage Opportunities	Buy Low (CL=F): 74.94653699613328 Sell High (BZ=F): 76.84518699615778

The initial analysis of the historical data from NYMEX and ICE markets yielded critical insights. The profit distribution patterns were analyzed based on area differences derived from integral calculus. This approach facilitated the evaluation of market efficiency and identification of arbitrage opportunities. Specifically, the analysis suggested buying low at \$72.74 and selling high at \$77.08. Figure 3 below breaks down price by volume for this historical period. Figure 4 shows historical closing price from January 2022 - January 2024.

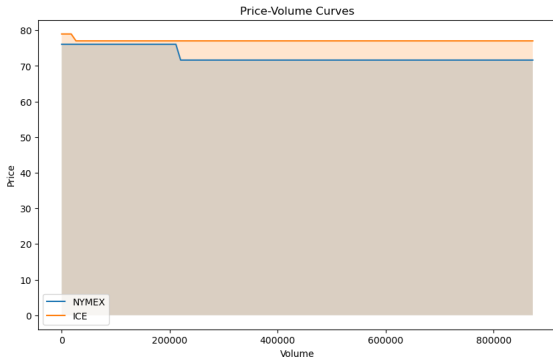


Figure 3: Price-Volume Curves for NYMEX and ICE Markets (January 2022 - January 2024)

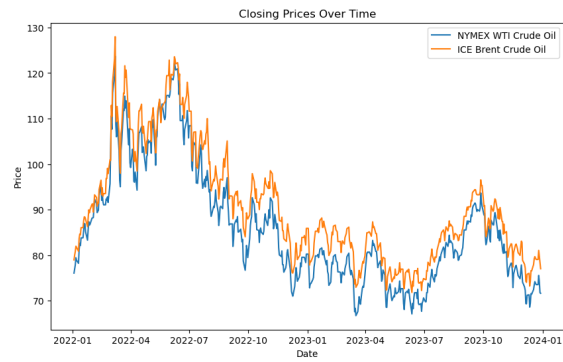


Figure 4: NYMEX WTI and ICE Brent Crude Oil Closing Prices Over Time

To forecast future prices, Autoregressive Integrated Moving Averages (ARIMA) models were employed. The best candidates were identified based on the lowest AIC scores using a stepwise search. For NYMEX, the best model was $ARIMA(1,0,1)(2,1,0)$ with an AIC of 2,483.33, and for ICE Brent Crude Oil, the best model was $ARIMA(3,0,2)(2,1,0)$ with an AIC of 2,481.72. These models are suitable for time series data as they capture trends and patterns in historical data to make informed predictions about future prices. The data was sourced from Yahoo Finance, and the frequency was set to business days to ensure continuity. Missing values were handled through forward filling. Figure 5 below figure shows the forecasted prices.

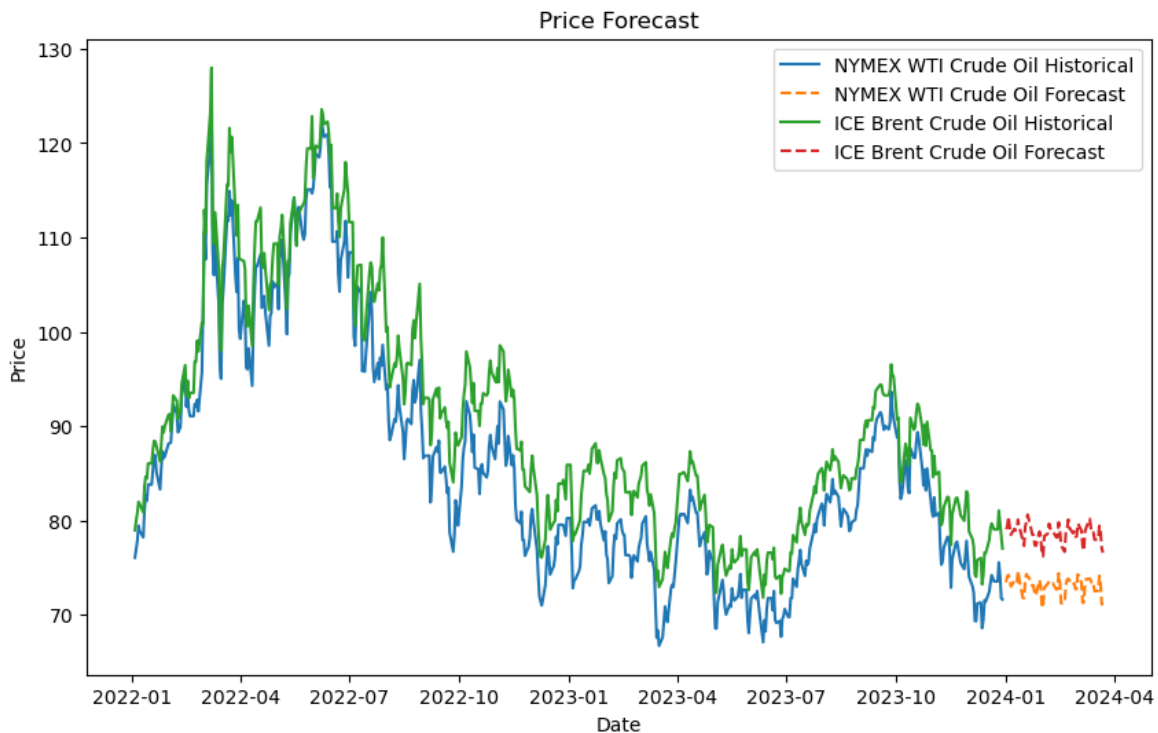


Figure 5: NYMEX WTI and ICE Brent Crude Oil 60 Day ARIMA Forecast

Traders can use the identified buy low and sell high recommendations to potentially

profit from price discrepancies between the two. The analysis provides insights into the relationship between trading volume and price movements, which can be valuable for understanding market dynamics.

By incorporating the price-volume relationship through the area calculation, the algorithm provides a systematic approach to identify arbitrage opportunities. This relationship indirectly influences the buy/sell decisions by serving as a basis for comparison between different securities.

9 Limitations

Several limitations need to be acknowledged. First, the model simplifies market dynamics, potentially overlooking some complexities of real-world markets. Second, the assumption that the market space can be accurately represented as a non-equilateral triangle may not always hold true.

The effectiveness of this method heavily relies on the availability and quality of high-resolution price and volume data. Inconsistent or incomplete data could undermine its effectiveness. The method involves complex calculations, which might require significant computational resources, especially when applied to multiple securities or large datasets. Markets are dynamic, and the introduction of a new arbitrage identification method could lead to changes in market behavior as participants adapt to and potentially exploit the new approach.

10 Conclusion

This geometric approach offers new insights into market structure and arbitrage opportunities by representing market space as a non-equilateral triangle. It highlights the dynamics between price and volume, uncovering hidden arbitrage opportunities across different segments. Integrating real market data bridges theoretical frameworks with practical trading, with areas under price-volume curves indicating market efficiency and varying arbitrage opportunities. This method identifies profitable trading points and illustrates how market space shapes profit distribution patterns.

The 60-day forecast extends this analysis into future market dynamics using the ARIMA model, which captures trends and patterns in time series data. This allows for advanced evaluation of arbitrage opportunities, predicting future prices with notable accuracy. The forecast suggests buying low at 74.95 and selling high at 76.85, indicating stability in arbitrage potential. Future research could refine this model by incorporating advanced machine learning techniques and exploring broader applications in finance, such as the impact of external economic factors and high-frequency trading data for more granular insights.

References

- [1] Farinelli, S., & Takada, H. (2021). Can you hear the shape of a market? Geometric arbitrage and spectral theory. *Axioms*, 10(4), 242. <https://doi.org/10.3390/axioms10040242>

11 Appendix: Market Space Dynamics Algorithm

Algorithm 1 Market Space Dynamics Analysis

```
1: Input: Historical price and volume data for commodities from multiple markets
2: Output: Visualization and analysis of profit distribution and arbitrage opportunities
3: Step 1: Data Initialization and Collection
4: Load and preprocess historical price and volume data
5: Define market space vertices
6: Step 2: Define Market Functions
7: Define price-volume functions for each market segment
8:  $f_{\text{market1}}(x) = (\text{price\_slope}_{\text{market1}} \times x) + \text{price\_intercept}_{\text{market1}}$ 
9:  $f_{\text{market2}}(x) = (\text{price\_slope}_{\text{market2}} \times x) + \text{price\_intercept}_{\text{market2}}$ 
10: Step 3: Calculate Areas Using Integral Calculus
11: Function CalculateArea( $f(x)$ , lower_limit, upper_limit) returns area
12:   area =  $\int_{\text{lower\_limit}}^{\text{upper\_limit}} f(x) dx$ 
13:   returnarea
14: Compute areas for each security
15: for each security do
16:   areasecurity = CalculateArea( $f_{\text{security}}$ , lower_limit, upper_limit)
17: end for
18: Step 4: Plot Price-Volume Curves
19: Function PlotPriceVolumeCurve( $f(x)$ , lower_limit, upper_limit)
20:   Generate and plot x and y values
21:   Shade area under the curve and annotate key points
22: Visualize price-volume curves for each market
23: for each market do
24:   PlotPriceVolumeCurve( $f_{\text{market}}$ , lower_limit, upper_limit)
25: end for
26: Step 5: Analyze Arbitrage Opportunities
27: Function AnalyzeArbitrageOpportunities(area1, area2) returns opportunities
28:   if area1  $\neq$  area2
29:     Identify significant price deviations and highlight opportunities
30:   returnopportunities
31: Compare areas and identify arbitrage opportunities
32: opportunities = AnalyzeArbitrageOpportunities( $area_{\text{market1}}$ ,  $area_{\text{market2}}$ )
33: Step 6: Interpret Results
34: Function InterpretResults(opportunities) returns summary
35:   Assess profit distribution patterns and market efficiency
36:   Summarize findings return summary
37: Summarize findings for the markets
38: summary = InterpretResults(opportunities)
39: Print summaries
40: End Algorithm
```

Appendix: MarketSpaceDynamics Class Implementation in Python

```
import yfinance as yf
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy.integrate import quad
from pmdarima import auto_arima

class MarketSpaceDynamics:
    def __init__(self, tickers, start_date, end_date,
                 forecast_steps=365, labels=None):
        self.tickers = tickers
        self.start_date = start_date
        self.end_date = end_date
        self.forecast_steps = forecast_steps
        self.labels = labels if labels is not None else {ticker: ticker
                                                         for ticker in tickers}
        self.data = {}
        self.fetch_data()
        self.preprocess_data()

    def fetch_data(self):
        for ticker in self.tickers:
            data = yf.download(ticker, start=self.start_date,
                              end=self.end_date)
            data = data.asfreq("B") # Set frequency to business day
            data = data.fillna(method="ffill") # Forward fill missing
            values
            self.data[ticker] = data

    def preprocess_data(self):
        for ticker in self.tickers:
            self.data[ticker]["Price"] = self.data[ticker]["Close"]
            self.data[ticker]["Volume"] = self.data[ticker]["Volume"]

    def price_volume_function(self, data, x):
        return np.interp(x, data["Volume"], data["Price"])

    def calculate_area(self, func, lower_limit, upper_limit):
        area, _ = quad(func, lower_limit, upper_limit)
        return area

    def plot_price_volume_curve(self, func, lower_limit, upper_limit,
                                label):
        x_values = np.linspace(lower_limit, upper_limit, 100)
        y_values = func(x_values)
        plt.plot(x_values, y_values, label=label)
        plt.fill_between(x_values, y_values, alpha=0.2)

    def plot_price_volume_curves(self):
        volume_range = np.linspace(
            min([self.data[ticker]["Volume"].min() for ticker in
```

```

        self.tickers]),
        max([self.data[ticker]["Volume"].max() for ticker in
            self.tickers]),
        100,
    )
plt.figure(figsize=(10, 6))
for ticker in self.tickers:
    self.plot_price_volume_curve(
        lambda x: self.price_volume_function(self.data[ticker],
            x),
        volume_range.min(),
        volume_range.max(),
        self.labels[ticker],
    )
plt.xlabel("Volume")
plt.ylabel("Price")
plt.title("Price-Volume Curves")
plt.legend()
plt.show()

def plot_closing_prices(self):
    plt.figure(figsize=(10, 6))
    for ticker in self.tickers:
        plt.plot(self.data[ticker].index,
            self.data[ticker]["Close"], label=f"{self.labels[ticker]}
            Historical")
    plt.xlabel("Date")
    plt.ylabel("Price")
    plt.title("Closing Prices Over Time")
    plt.legend()
    plt.show()

def forecast_prices(self, data, steps):
    model = auto_arma(
        data["Close"],
        start_p=1,
        start_q=1,
        max_p=3,
        max_q=3,
        seasonal=True,
        m=12,
        stepwise=True,
        suppress_warnings=True,
        D=1,
        trace=True,
        n_jobs=1,
    )
    forecast = model.predict(n_periods=steps)
    forecast_index = pd.date_range(start=data.index[-1] +
        pd.Timedelta(days=1), periods=steps, freq="B")
    forecast_df = pd.DataFrame({"Date": forecast_index, "Forecast":
        forecast})
    return forecast_df

def plot_forecasts(self, steps):

```

```

forecast_dfs = {}
plt.figure(figsize=(10, 6))
for ticker in self.tickers:
    forecast_df = self.forecast_prices(self.data[ticker], steps)
    forecast_dfs[ticker] = forecast_df
    plt.plot(self.data[ticker].index,
             self.data[ticker]["Close"], label=f"{self.labels[ticker]}
             Historical")
    plt.plot(forecast_df["Date"], forecast_df["Forecast"],
             label=f"{self.labels[ticker]} Forecast", linestyle="--")
plt.xlabel("Date")
plt.ylabel("Price")
plt.title("Price Forecast")
plt.legend()
plt.show()
return forecast_dfs

def calculate_forecasted_areas(self, forecast_dfs):
    volume_range = np.linspace(
        min([self.data[ticker]["Volume"].min() for ticker in
            self.tickers]),
        max([self.data[ticker]["Volume"].max() for ticker in
            self.tickers]),
        100,
    )

    def price_volume_function_forecast(data, forecast_df, x):
        return np.interp(
            x,
            np.concatenate([data["Volume"], [data["Volume"].max()] *
                len(forecast_df)]),
            np.concatenate([data["Price"], forecast_df["Forecast"]]),
        )

    forecast_areas = {}
    for ticker in self.tickers:
        forecast_areas[ticker] = self.calculate_area(
            lambda x:
                price_volume_function_forecast(self.data[ticker],
                    forecast_dfs[ticker], x),
            volume_range.min(),
            volume_range.max(),
        )
    return forecast_areas

def analyze_arbitrage_opportunities(self, areas):
    min_area = min(areas.values())
    max_area = max(areas.values())
    min_ticker = min(areas, key=areas.get)
    max_ticker = max(areas, key=areas.get)
    return {"buy_low": (min_area, min_ticker), "sell_high":
        (max_area, max_ticker)}

def adjust_arbitrage_opportunities(self, opportunities,
    volume_range_max):

```

```

opportunities["buy_low"] = (opportunities["buy_low"][0] /
    volume_range_max, opportunities["buy_low"][1])
opportunities["sell_high"] = (opportunities["sell_high"][0] /
    volume_range_max, opportunities["sell_high"][1])
return opportunities

def interpret_results(self, opportunities, forecast=False):
    summary = "Profit Distribution Patterns: Analyzed profit
        distribution patterns based on area differences.\n"
    summary += "Market Efficiency: Evaluated market efficiency based
        on arbitrage opportunities.\n"
    if forecast:
        summary += "Forecasted "
    summary += "Arbitrage Opportunities:\n"
    if opportunities:
        summary += f" Buy Low ({opportunities['buy_low'][1]}):
            {opportunities['buy_low'][0]}\n"
        summary += f" Sell High ({opportunities['sell_high'][1]}):
            {opportunities['sell_high'][0]}\n"
    return summary

def run_analysis(self):
    volume_range = np.linspace(
        min([self.data[ticker]["Volume"].min() for ticker in
            self.tickers]),
        max([self.data[ticker]["Volume"].max() for ticker in
            self.tickers]),
        100,
    )

    # Calculate current areas
    current_areas = {ticker: self.calculate_area(lambda x:
        self.price_volume_function(self.data[ticker], x),
        volume_range.min(), volume_range.max()) for ticker in
        self.tickers}

    # Analyze current arbitrage opportunities
    current_opportunities =
        self.analyze_arbitrage_opportunities(current_areas)
    current_opportunities =
        self.adjust_arbitrage_opportunities(current_opportunities,
        volume_range.max())
    current_summary = self.interpret_results(current_opportunities)

    # Forecast future prices
    forecast_dfs = self.plot_forecasts(self.forecast_steps)

    # Calculate forecasted areas
    forecast_areas = self.calculate_forecasted_areas(forecast_dfs)

    # Analyze forecasted arbitrage opportunities
    forecasted_opportunities =
        self.analyze_arbitrage_opportunities(forecast_areas)
    forecasted_opportunities =
        self.adjust_arbitrage_opportunities(forecasted_opportunities,

```

```

        volume_range.max())
forecasted_summary =
    self.interpret_results(forecasted_opportunities,
        forecast=True)

# Print summaries
print("Current Market Summary:")
print(current_summary)

print("\nForecasted Market Summary:")
print(forecasted_summary)

# Print forecasted dataframes
for ticker, forecast_df in forecast_dfs.items():
    print(f"{ticker} Forecast for the next {self.forecast_steps}
        days:")
    print(forecast_df)

return current_summary, forecasted_summary

# Example usage:
labels = {
    "CL=F": "NYMEX WTI Crude Oil",
    "BZ=F": "ICE Brent Crude Oil",
}

analysis = MarketSpaceDynamics(
    ["CL=F", "BZ=F"],
    "2022-01-01",
    "2024-01-01",
    forecast_steps=60,
    labels=labels,
)

```