Empirical Analysis Report

Clarisse Acknin & Gwendoline Hays-Valentin & Hugo Michel

Professor: Catherine Bruneau

Asset Pricing

-

M2 Finance Technology Data - Sorbonne School Of Economics

Université Paris 1 Panthéon-Sorbonne

Paris, France, Fall 2023

Contents

1 Empirical Analysis 1

The goal is to estimate the multi-beta relationship for a global stock index. To do so, we will use the APT model proposed by Ross. It can be seen as a generalization of the beta relationship. However the difference between the APT model and the CAPM is that Ross does not refer to market equilibrium. In the Ross model there is no arbitrage opportunity which means that it is not possible to become richer without expenses and risk taking.

APT is based on the basic idea that there are no arbitrage opportunities that last over time. In effect, an asset A that is as risky as asset B, but more profitable, would see its demand increase rapidly, until its profitability became equal to that of asset B, thus cancelling out any arbitrage opportunity.

The other basic assumption of APT is that the expected profitability of a stock can be modelled by a linear function of various macro-economic or sector-specific factors, weighted according to their impact on the stock by a specific beta coefficient.

These factors are diverse and can range from oil prices to US GDP, from European key rates to the exchange rate of a currency pair. These are all factors likely to influence the price of the asset under study.

1.1 Ross APT model

The model proposed by Ross, is based on a multi-factor model, where the returns of an asset are related to several macroeconomic factors. These factors could include inflation rates, interest rates, economic indicators, etc. The model assumes that the expected return on an asset is a linear function of these factors.

The period of study for the data we collected are from 2015 to 2019.

To select the exogenous and endogenous factors and implement the multi-beta model proposed by Ross we based our analysis on the research paper of Ross (1986)

Factors of the model are:

- I - Inflation : Log relative of US consumer Price Index (**FPCPITOTLZGUSA** : Inflation, consumer prices for the United States)

- TB - Treasury bill rate : end of period return on 1-month bills (**DTB4WK** : 4-Week Treasury Bill Secondary Market Rate, Discount Basis)
- LGB - Long-term government bonds : Return on LT government bonds (**DGS30**: Market Yield on U.S. Treasury Securities at 30-Year Constant Maturity, Quoted on an Investment Basis)
- IP - Industrial production : industrial production during month (**INDPRO**: The industrial production (IP) index measures the real output of all relevant establishments located in the United States)
- Baa - Low grade bond (**BAA**: Moody's Seasoned Baa Corporate Bond Yield Financial instruments are based on bonds with maturities 20 years and above.)
- $EWNY$ - return on equally weighted portfolio of NYSE listed stocks (**VADAX**: Invesco Equally-Wtd S&P 500 A)
- $VWNY$ - return on a value-weighted portfolio of NYSE listed stocks (**SP500**: SP 500 EQUAL WEIGHT INDEX)
- CG - Growth rate in real per capita consumption (**A794RX0Q048SBEA**: Real personal consumption expenditures per capita)
- OG - log relative of producer price index/crude petroleum series (**WTISPLC**:Spot Crude Oil Price: West Texas Intermediate)
- EI - Expected inflation: (**MICH**:Median expected price change next 12 months, Surveys of Consumers)

To collect the data mentioned above we use the FRED API and Yahoo Finance API

We includes also 3 Factors of the Fama French collected computed by researchers from Dartmouth University (Kenneth (2014)) . This dataset compute HML (High Minus low factors), SMB (Small minus Big factors) and $Mkt - RF$ (Market excess return). We will explain later in more details each factor of Fama French model.

Samples of factors collected from FRED API

Figure 1.1: Factors collected from FRED API

	I_factor	TB_factor	LGB_factor	IP_factor	Baa_factor	EWNY_factor	VWNY_factor	CG_factor	OG_factor	EI_factor	Mkt-RF	SMB	HML	RMW	CMA
2014-04-01	1.622223	0.051429	3.620952	101.8425	4.90	29.104841	1863.523333	0.001776	0.012599	3.2	0.87	0.65	-0.37	-0.18	-0.29
2014-05-01	1.622223	0.023333	3.517619	102.2594	4.76	29.115611	1864.263333	0.007778	0.001078	3.3	0.04	-0.18	-0.16	-0.50	-0.16
2014-07-01	1.622223	0.024286	3.420000	102.8163	4.73	30.493863	1947.087619	0.007778	-0.020796	3.3	0.74	0.43	-0.39	0.07	-0.10
2014-08-01	1.622223	0.024091	3.331818	102.6562	4.69	30.818151	1973.100000	0.007692	-0.068057	3.2	-0.32	-0.23	-0.05	0.09	0.16
2014-10-01	1.622223	0.011429	3.260000	102.9892	4.69	31.084981	1993.226190	0.007692	-0.094518	2.9	-1.39	-0.14	0.30	0.20	-0.07
2014-12-01	1.622223	0.041667	3.038333	103.6345	4.74	31.917171	2044.572105	0.009327	-0.217707	2.8	-0.89	-0.89	0.62	0.10	0.48
2015-04-01	0.118627	0.022727	2.626364	101.2440	4.48	32.897359	2079.990455	0.005937	0.138645	2.6	-0.38	0.34	0.44	-0.17	0.24
2015-05-01	0.118627	0.015909	2.585909	100.7830	4.89	33.284009	2094.862857	0.005080	0.088522	2.8	1.01	-0.31	-0.60	0.25	-0.10
2015-06-01	0.118627	0.013000	2.955000	100.4781	5.13	33.391662	2111.943500	0.005080	0.009280	2.7	0.17	-0.06	-0.22	0.24	-0.34
2015-07-01	0.118627	0.004091	3.111818	101.1052	5.20	33.126581	2099.283636	0.005080	-0.149114	2.8	0.61	-0.76	-0.03	0.21	0.08
2015-09-01	0.118627	0.033810	2.855714	100.6507	5.34	31.981528	2039.866190	0.004902	0.060882	2.8	-2.91	0.23	-0.50	-0.04	0.07
2015-10-01	0.118627	-0.000476	2.952857	100.1871	5.34	30.555512	1944.402381	0.004902	0.016271	2.7	0.13	-0.52	-0.06	-0.14	-0.15
2015-12-01	0.118627	0.068947	3.030000	98.9390	5.46	32.149797	2080.616500	0.001956	-0.123704	2.6	0.97	-0.63	0.25	-0.08	-0.11

As we can see we include also the 3 factors from Fama French model which are:

- Market risk premium which corresponds to the market excess return minus the risk free rate $(r_m - r_f)$
- SMB : Historic excess returns of small-cap companies over large-cap companies
- HML : Historic excess returns of value stocks (high book-to-price ratio) over growth stocks (low book-to-price ratio)

Compute derived factors

To compute the derived factors of the Ross APT model we proceed as follow:

- Monthly Growth IP: $MP(t) = log(IP(t)/IP_t - 1)$
- Annual growth IP: $YP(t) = log(IP(t)/IP_t - 12))$
- Annual growth IP: $E(I(t)) = ExpectedInflation$
- Unexpected Inflation: $UI(t) = I(t) - E(I(t)|t - 1)$
- Real interest ex post: $RHO(t) = TP(t - 1) - I(t)$

- Change in expected infla $DEI(t) = E(I(t+1)|t) - E(I(t)|t-1)$
- Risk premium: $UPR(t) = Baa(t) - LGB(t)$
- Term Structure: $UTS(t) = LGB(t) - TB(t-1)$

Figure 1.2: Sample of derived factors

	MP_derived_factor	YP_derived_factor	UI_derived_factor	RHO_derived_factor	DEI_derived_factor	UPR_derived_factor	UTS_derived_factor	EI_factor	Mkt-RF
2015-12-01	-0.012536	-0.028924	-2.581373	-0.119103	-0.1	2.430000	3.030476	2.6	0.97
2016-02-01	-0.000257	-0.033266	-1.338417	-1.192636	-0.1	2.481579	2.789474	2.5	-0.04
2016-03-01	-0.007335	-0.046032	-1.238417	-1.037373	0.2	2.507000	2.398789	2.7	2.34
2016-04-01	0.002809	-0.041665	-1.438417	-1.010583	0.1	2.105455	2.433545	2.8	0.64
2016-06-01	0.002643	-0.042261	-1.538417	-1.010674	-0.2	1.902381	2.376710	2.6	0.20
2016-07-01	0.001098	-0.047408	-1.338417	-1.040155	0.1	1.767727	2.230844	2.7	0.24
2016-08-01	-0.000816	-0.024887	-1.438417	-1.042947	-0.2	2.013000	2.008364	2.5	-0.16
2016-09-01	-0.000971	-0.021294	-1.238417	-1.003083	-0.1	2.048261	2.003239	2.4	0.03
2016-11-01	-0.003192	-0.021456	-1.138417	-1.003757	0.0	2.209500	2.242674	2.4	-0.68
2016-12-01	0.006953	-0.020725	-1.138417	-1.025583	-0.2	1.968000	2.626000	2.2	-0.36

Build the portfolio

To build the portfolio, we propose to select companies with different capitalization:

- **Large-Cap (Large Capitalization):** Market cap typically between \$10B-\$200B
- **Mid-Cap (Medium Capitalization):** Market cap typically between \$2B-\$10B
- **Small-Cap (Small Capitalization):** Market cap typically under \$2B

The composition of the portfolio is as follow:

Large-Cap

- McDonald's Corporation Common Stock (MCD)
- Amazon.com, Inc. Common Stock (AMZN)

Mid-Cap

- EMCOR Group, Inc. Common Stock (EME)
- Bank OZK Common Stock (OZK)

Smal-Cap

- First Financial Bancorp. Common Stock (FFBC)
- Douglas Emmett, Inc. Common Stock (DEI)

To compute the return, we consider that my portfolio is a equally portfolio return.

Figure 1.3: Portfolio composition with the return

Date	AMZN	DEI	EME	FFBC	MCD	OZK	return
2015-12-01	657.695496	467.035652	948.562462	296.976905	1853.377693	844.967262	-0.016744
2016-01-01	736.188499	508.266546	1036.999466	306.444738	2128.295326	892.780891	0.089610
2016-02-01	571.008501	419.743746	811.704830	231.999842	1854.394852	685.180885	-0.203325
2016-03-01	530.619999	408.554647	864.761814	245.358776	1957.997879	632.546017	0.000768
2016-04-01	629.611498	489.490295	1008.602448	301.150023	2230.913963	736.175743	0.180266
2016-05-01	644.273998	507.826395	978.213253	299.090362	2230.564278	720.696579	0.000433

1.2 Apply APT Ross Model: Estimating β coefficients

To apply the APT Model proposed by Ross we implement the following regression:

$$Portfolio_{return} = \alpha + \beta_1 MP(t) + \beta_2 YP(t) + \beta_3 E[I(t)] + \beta_4 UI(t) + \beta_5 RHO(t) + \beta_6 DEI(t) + \beta_7 URP(t) + \beta_8 UTS(t) + \beta_9 EI(t) + \beta_{10} Mkt - RF + \beta_{11} SMB + \beta_{12} HML$$

where:

- α : is the constant of the regression
- β_i : is the coefficient of each variables of the model will estimate by applying OLS.

Results of the regression

Figure 1.4: OLS Summary

OLS Regression Results

Dep. Variable:	return	**R-squared:**	0.598
Model:	OLS	**Adj. R-squared:**	0.283
Method:	Least Squares	**F-statistic:**	1.895
Date:	Fri, 17 Nov 2023	**Prob (F-statistic):**	0.130
Time:	00:13:09	**Log-Likelihood:**	32.157
No. Observations:	26	**AIC:**	-40.31
Df Residuals:	14	**BIC:**	-25.22
Df Model:	11		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.9553	0.659	2.967	0.010	0.542	3.369
MP_derived_factor	6.6422	4.192	1.584	0.135	-2.349	15.634
YP_derived_factor	-0.7179	1.838	-0.391	0.702	-4.661	3.225
UI_derived_factor	-0.1286	0.143	-0.898	0.384	-0.436	0.178
RHO_derived_factor	0.0922	0.140	0.658	0.521	-0.209	0.393
DEI_derived_factor	0.4488	0.244	1.840	0.087	-0.074	0.972
UPR_derived_factor	-0.1869	0.111	-1.685	0.114	-0.425	0.051
UTS_derived_factor	-0.1854	0.109	-1.701	0.111	-0.419	0.048
EI_factor	-0.4717	0.229	-2.060	0.059	-0.963	0.019
Mkt-RF	0.0752	0.032	2.325	0.036	0.006	0.145
SMB	0.0369	0.040	0.916	0.375	-0.049	0.123
HML	0.0011	0.041	0.027	0.979	-0.087	0.089

Omnibus:	1.380	**Durbin-Watson:**	2.682
Prob(Omnibus):	0.502	**Jarque-Bera (JB):**	0.418
Skew:	0.226	**Prob(JB):**	0.811
Kurtosis:	3.426	**Cond. No.**	939.

We can see that the value of the R^2 of our regression is 0.598. In this case, approximately 59.8% of the variance in the return can be explained by the independent variables in the model. In our case, the R^2 is not close to 1 we can say that the majority of our portfolio's movements can't be explained by the factors in our model.

Furthermore, the adjusted R-squared, which takes into account the number of predictors in the model is 0.283. It is a modified version of R-squared that adjusts for the number of predictors. The adjusted R-squared is lower than the R-squared, suggesting that not all the variables in the model contribute significantly.

The value of α is 1.9. Hence, given that $\alpha > 1$ we can say that the portfolio is outperforming, and the returns are higher than expected given its beta. Indeed a alpha significantly greater than 1, indicates a strong positive abnormal performance.

The coefficient for MP derived factor is 6.6422. This means that, holding other variables constant, a one-unit increase in "MP derived factor" is associated with an estimated increase of 6.6422 units in the dependent variable (which is "return" in our case). Conversely, a decrease in the MP derived factor is associated with a decrease in the portfolio return.

Otherwise, we observe that the values of the other β coefficients are between -1 and 1.

Based of the OLS results, regarding the statistical significance of the variable, we can see that the variables "const" (intercept), $Mkt - RF$ (p-value: 0.036) and EI (p-value: 0.059) factor appear to be statistically significant at a conventional significance level of 0.05. DEI (p-value: 0.087) derived factor and MP (p-value: 0.135) derived factor are marginally significant. The other variables are not statisticaly significant in this model.

Please refer to the following link to see the full code of this assignment: Empirical Analysis 1 - Multi Beta Ross model

1.3 Fama French 3 factors

The Fama-French Three-factor Model is an extension of the Capital Asset Pricing Model (CAPM). The Fama-French model aims to describe stock returns through three factors: market risk, the outperformance of small-cap companies relative to large-cap companies, and the outperformance of high book-to-market value companies versus low book-to-market value companies. The rationale behind the model is that high value and small-cap companies tend to regularly outperform the overall market.

Figure 1.5: Fama French Model

Get stocks data from API Yahoo Finance

To apply a Factors model we have to build a portfolio first. To constitute the portfolio, we propose to select companies with different capitalization

Market capitalization is typically expressed in terms of billions (B), millions (M), or even trillions (T) of dollars, depending on the size of the company. It categorizes companies into different size classes:

- Large-Cap (Large Capitalization): Market cap typically over $10 billions.
- Mid-Cap (Medium Capitalization): Market cap typically between $2 billions and 10 billions.
- Small-Cap (Small Capitalization): Market cap typically under $2 billions.

Market capitalization is an important metric for investors, analysts, and researchers as it provides insights into the size and scale of a company. It's widely used for categorizing stocks and understanding their relative risk and growth potential.

According these three size classes we select the following companies:

Large Cap stocks	**Mid Cap Stocks**	**Small Cap Stocks**
LVMH (MC.PA)	Arkema (AKE.PA)	OL (OLG.PA)
L'Oréal (OR.PA)	Klépierre (LI.PA)	Fnac Darty (FNAC.PA)
Hermès (RMS.PA)	FDJ (FDJ.PA)	Atos (ATO.PA)
Total Energies (TTE)	Teleperformance (TEP.PA)	Criteo (CRTO)
Sanofi (SNY)	Groupe SEB (SK.PA)	Rémy Cointreau (RCO.PA)
Dior (CDI.PA)	Scor (SRC.PA)	OVH Groupe (OVH.PA)
Schneider Electric (SU.PA)	SPIE (SPIE.PA)	Rubis (RUI.PA)
Air Liquide (AI.PA)	Rexel (RXL.PA)	Icade (ICAD.PA)
BNP Paribas (BNP.PA)	Alten (ATE.PA)	BIC (BB.PA)
Crédit Agricole (ACA.PA)	Colas (RE.PA)	Ubisoft (UBI.PA)

In summary our portfolio contains, 10 Large Cap stocks, 10 Mid-Cap stocks and 10 Small-Cap stocks

Regarding the period we chose the following period: from 2022-01-01 2023-09-01. We stop our analysis at 2023-09-01 because we don't have the interest rate for France after this date.

Regarding the frequency of data collection, we have chosen to concentrate on a daily frequency.

To collect the stocks of these companies mentioned in the table we use the Yahoo Finance API.

Example BIC stock data collected from Yahoo API

Figure 1.6: BIC stocks

Date	Open	High	Low	Close	Volume	Dividends	Stock Splits
2022-01-03 00:00:00+01:00	43.505770	44.112570	43.340278	43.855141	38866	0.0	0.0
2022-01-04 00:00:00+01:00	44.241287	44.903252	42.972524	43.505772	78283	0.0	0.0
2022-01-05 00:00:00+01:00	43.671260	44.866473	43.634484	44.094181	70955	0.0	0.0
2022-01-06 00:00:00+01:00	43.836752	44.259672	43.487381	44.094181	47330	0.0	0.0
2022-01-07 00:00:00+01:00	44.259673	44.351615	43.708038	44.314838	21717	0.0	0.0
...	...	...	...	...	...	...	...
2023-08-25 00:00:00+02:00	57.799999	58.250000	57.700001	58.150002	15661	0.0	0.0
2023-08-28 00:00:00+02:00	58.250000	58.799999	58.099998	58.500000	17023	0.0	0.0
2023-08-29 00:00:00+02:00	58.599998	59.000000	58.099998	58.750000	24606	0.0	0.0
2023-08-30 00:00:00+02:00	58.849998	59.500000	58.650002	59.049999	38847	0.0	0.0
2023-08-31 00:00:00+02:00	59.150002	59.200001	58.799999	59.000000	23641	0.0	0.0

426 rows × 7 columns

The next step is to compute the Market Capitalization, for each stocks over the period of time mentioned previously.

We use the following formula to compute the Market Capitalization:

$$MarketCap = TotalShares \times ValuePerShare$$

Figure 1.7: Sample of Market Cap of our portfolio

	Symbol	Market_cap
0	MC.PA	6.875861e+10
1	OR.PA	3.352710e+10
2	RMS.PA	2.714384e+10
3	TTE	2.783556e+10
4	SNY	2.384369e+10
5	CDI.PA	8.464171e+08
6	SU.PA	3.049931e+10
7	AI.PA	2.634739e+10

Import historical price for each stocks

We import historical stock price data to build our portfolio return.

Figure 1.8: Sample of stocks of our portfolio

Date	AI.PA	AKE.PA	ATE.PA	ATO.PA	BB.PA	BNP.PA	CA.PA	CDI.PA	CRTO	FDJ.PA	...
2022-01-03	135.460526	117.174904	154.256027	38.540001	43.855145	53.302734	15.186107	710.527283	38.880001	36.197689	...
2022-01-04	138.073303	118.437363	156.406631	38.180000	43.505772	55.088203	15.464016	722.693848	38.169998	36.281479	...
2022-01-05	140.317810	121.429863	155.820084	38.919998	44.094185	55.157883	16.250641	735.347046	36.099998	36.709743	...
2022-01-06	137.652435	119.699821	149.954834	38.770000	44.094185	55.906910	17.272783	702.253967	35.560001	36.048729	...
2022-01-07	137.161453	121.710396	147.217728	38.590000	44.314838	56.176903	17.291624	680.840881	35.610001	35.937004	...
...	...	...	...	...	...	...	...	...	...	...	...
2023-08-25	162.839996	93.320000	125.699997	6.648000	58.150002	58.310001	18.209999	742.500000	28.860001	32.580002	...
2023-08-28	165.300003	95.099998	127.800003	6.918000	58.500000	59.500000	18.235001	754.000000	28.629999	32.939999	...
2023-08-29	165.520004	97.000000	129.199997	7.068000	58.750000	59.900002	17.405001	764.000000	29.080000	33.320000	...
2023-08-30	166.240005	97.160004	130.000000	7.354000	59.049999	59.880001	17.530001	766.500000	29.450001	33.240002	...
2023-08-31	166.860001	96.639999	131.300003	7.762000	59.000000	59.709999	17.670000	758.500000	29.495001	33.400002	...

431 rows × 30 columns

To compute the daily return of each stocks return over time, we apply the following formula:

$$\text{Percentage Change} = \frac{\text{Price}_{\text{t}} - \text{Price}_{\text{t-1}}}{\text{Price}_{\text{t-1}}} \times 100\%$$

Where Price_{t} is the current stock price at time t and $\text{Price}_{\text{t-1}}$ is the previous stock price at time $t-1$.

Then, we compute the portfolio return.

To compute the portfolio return, we apply the following formula:

$$\text{Portfolio Return} = \sum_{i=1}^{n} w_i \cdot R_i$$

Where:

- n is the number of assets in the portfolio.
- w_i represents the weight of asset i in the portfolio.

- R_i is the return of asset i.

In our case, we consider our portfolio is equally weighted .

Figure 1.9: Return of our portfolio

Portfolio_return
0.009486
0.004433
-0.012526
-0.006029
-0.014416
...
0.000550
0.013374
0.007774
0.000409
0.003356

1.4 Compute the factors of the Fama French 3 factors model

Risk factors are the elements that make up a portfolio's performance. The stocks of the portfolio will have certain characteristics, such as volatility, size, strong financial fundamentals and so on. These characteristics determine the performance of individual stocks and, taken together, make up the factors in your portfolio. Factors are therefore measured by examining the combined characteristics of the assets in our portfolio.

There are two categories of risk in a portfolio:

- macro factors: interest rate, exchange rate, country, business sector.
- style factors: momentum, volatility, value and quality.

Endogeneous Risk (Systematic risk)

Systematic risk, often referred to as market risk, represents the portion of the total risk associated with an investment that cannot be eliminated through diversification. It is related to factors affecting the entire market or a specific market segment. In the context of the Fama-French model, systematic risk factors would include the Market Excess Return (Mkt-RF), Small Minus Big (SMB), and High Minus Low (HML) factors. These factors affect all assets in the market to varying degrees.

Exogenous Risk (Unsystematic risk)

Exogenous factors are external variables or factors that are not directly accounted for within the Fama-French model but can impact the financial markets and asset prices. Macro-economic indicators such as inflation rate, interest rate, and unemployment rate are all examples of exogenous factors. These macroeconomic variables can affect investor sentiment, economic conditions, and market behavior but are not explicitly included in the Fama-French Three-Factor Model.

Compute Small Minus Big (SMB)

The size premium represents the additional return that small-cap stocks are expected to provide over large-cap stocks. To estimate the size premium, we need data on small-cap and large-cap stocks' historical returns.

In fact, it exists different methods to compute the SMB factors.

For example researchers from Dartmouth University (Kenneth (2014)) compute SMB as the average return on the small portfolios minus the average return on the big portfolios. To do that researcher build 3 types of portfolio:

- Value portfolio by identifying undervalued stocks (i.e. priced lower than the broader market)
- Growth portfolio by identifying overvalued stocks (i.e. priced higher than the broader market)

- Neutral portfolio a mix of overvalued and undervalued stocks

Then each portfolio is splitted in Small and Big to a specific threshold from the average Market Capitalization for each stocks of a given period.

Finally, they apply the following formula to compute SMB:

$$SMB = 1/3(SmallValue + SmallNeutral + SmallGrowth) - 1/3(BigValue + BigNeutral + BigGrowth)$$

In our case, we don't apply the same approach, we will just compute the Average Market Capitalization for each stock of our portfolio and split each stock into Small and Big according to a given threshold. Then, the next step will be to compute the average return of the Small portfolio and the Big portfolio over the time. Finally, we will then compute the SMB factors by substract the average return of the big portfolio and small portfolio.

To do so that we have first to divide our portfolio in Small vs. Big Size company.

Common thresholds are in terms of percentiles, such as the lower 30% for Small stocks and the upper 70% for Big stocks.

To do that we focus on the Average Market Capitalization of each stock for our study period.

Figure 1.10: Small vs. Big Size

	Symbol	Market_cap	Size
18	RE.PA	1.121089e+07	Small
19	FR.PA	6.638143e+07	Small
25	OVH.PA	3.534701e+08	Small
22	ATO.PA	3.762253e+08	Small
20	OLG.PA	3.762253e+08	Small
28	BB.PA	6.518869e+08	Small
27	ICAD.PA	6.973080e+08	Small
5	CDI.PA	8.464171e+08	Small
17	ATE.PA	1.034755e+09	Small
26	RUI.PA	1.208405e+09	Small
15	SPIE.PA	1.307413e+09	Small
14	SK.PA	1.399866e+09	Small
12	FDJ.PA	1.519934e+09	Small
23	CRTO	1.911234e+09	Small
24	RCO.PA	2.550520e+09	Small
10	AKE.PA	3.316239e+09	Small
21	FNAC.PA	3.736239e+09	Small
16	RXL.PA	3.884010e+09	Small
29	UBI.PA	4.004375e+09	Small
11	LI.PA	4.341193e+09	Small
13	TEP.PA	7.434295e+09	Small
9	CA.PA	9.694501e+09	Small
4	SNY	2.384369e+10	Big
7	AI.PA	2.634739e+10	Big
2	RMS.PA	2.714384e+10	Big
3	TTE	2.783556e+10	Big
6	SU.PA	3.049931e+10	Big
1	OR.PA	3.352710e+10	Big
8	BNP.PA	4.153781e+10	Big
0	MC.PA	6.875861e+10	Big

Then we compute the return of the Small size portfolio and Big portfolio size.

Finlay we take the difference of the return of the both portfolio to compute the SMB factor.

Figure 1.11: SMB factor

Portfolio_return	Small_portfolio_return	Big_portfolio_return	SMB
0.009486	0.008700	0.011648	-0.002948
0.004433	0.003938	0.005795	-0.001857
-0.012526	-0.010278	-0.018708	0.008429
-0.006029	-0.006058	-0.005950	-0.000108
-0.014416	-0.013472	-0.017012	0.003541
...	...	...	...
0.000550	-0.000525	0.003506	-0.004031
0.013374	0.013394	0.013318	0.000076
0.007774	0.007418	0.008754	-0.001336
0.000409	0.000973	-0.001143	0.002116
0.003356	0.008315	-0.010281	0.018596

Compute HMB factors (High Minus Big)

HMB represents the difference in returns between high book-to-market (value) and low book-to-market (growth) portfolios of stocks. The book-to-market ratio identifies undervalued or overvalued securities by taking the book value and dividing it by the market value. The ratio determines the market value of a company relative to its actual worth.

To compute the Book-to-Market (B/M) ratio in the Fama-French model, we need the historical data for both book value and market value of equity for a set of stocks. The B/M ratio is calculated as the ratio of book value to market value.

$BookToMarketRatio = BookValueEquity/MarketValueEquity$

Then according to a specific threshold we can split the portfolio into Small and Big portfolio and compute HMB factor.

However, in our case it's complicate to collect automatically the data of the Book To Market Ratio for each 30 stocks of our portfolio over the time for a given period.

According to the researcher of Kenneth (2014) we can compute HML factors as the average return on the two value portfolios minus the average return on the two growth portfolios.

$$HML = 1/2(SmallValue + BigValue) - 1/2(SmallGrowth + BigGrowth)$$

As reminder, value portfolio corresponds to set of undervalued stocks and growth portfoloio to a set of overvalued stocks.

We chose to apply the approach of Dartmouth researcher to compute the HML factors. To do that let's consider the three following portfolios:

- Small Value Portfolio : [BIC, OVH, Criteo]
- Big Value Portfolio : [LVMH, Air Liquide, Total Energies]
- Big Growth Value Portfolio : [Hermès, Sanofi, Crédit Agricole]

Please note that the construction of these portfolio is totally subjective.

Figure 1.12: HML factor

Small_value_port_return	Small_growth_port_return	Big_value_port_return	Big_growth_port_return	HML
-0.018165	0.004500	0.014057	0.003221	-0.005915
-0.024819	0.004591	0.012667	0.021192	-0.018968
-0.016500	-0.003594	-0.015315	0.004535	-0.016378
-0.011988	-0.008848	-0.000803	-0.008411	0.002234
0.006496	0.012995	-0.005420	-0.020556	0.004318
...	...	...	...	...
-0.002457	-0.004397	0.004102	0.004941	0.000550
0.003080	0.013975	0.012405	0.008668	-0.003579
0.015334	0.007218	0.008546	-0.009181	0.012922
0.001628	-0.017889	-0.001283	0.001170	0.008533
0.008788	0.016925	-0.009114	-0.008530	-0.004361

Then, we import risk free rate of France collected from FRED website and compute the Market Excess Return by subtract the portfolio return with the risk free rate.

Finally, the consolidated dataset we have construct for applying Fama French 3 Factors is as follow:

Figure 1.13: Dataset Fama French 3 Factors

	rf	Portfolio_return	SMB	HML	Market_excess_return
2022-01-04	0.31	0.009486	-0.002948	-0.005915	-0.300514
2022-01-05	0.31	0.004433	-0.001857	-0.018968	-0.305567
2022-01-06	0.31	-0.012526	0.008429	-0.016378	-0.322526
2022-01-07	0.31	-0.006029	-0.000108	0.002234	-0.316029
2022-01-10	0.31	-0.014416	0.003541	0.004318	-0.324416
...	...	...	...	...	...
2023-08-25	3.11	0.000550	-0.004031	0.000550	-3.109450
2023-08-28	3.11	0.013374	0.000076	-0.003579	-3.096626
2023-08-29	3.11	0.007774	-0.001336	0.012922	-3.102226
2023-08-30	3.11	0.000409	0.002116	0.008533	-3.109591
2023-08-31	3.11	0.003356	0.018596	-0.004361	-3.106644

430 rows × 5 columns

Applying Fama French Model

Developed by Eugene Fama and Kenneth French in the early 1990s, this model extends the CAPM model by adding the SMB, or size factor, and the HML, or value factor. The acronym for value, HML, stands for high minus low, while the acronym for size, SMB, stands for small minus big. This is because small stocks tend to outperform large stocks, so that small stock returns minus large stock returns essentially represent the small size premium. Value versus growth, or up versus down, on the other hand, is more cyclical. For example, during the crisis, value stocks outperformed, but during bull phases, growth stocks sometimes outperform.

We will apply a fama french 3 factors.

As reminder the three factors are :

- Market Excess Return
- SMB (Small Minus Big)
- HML (High Minus Low)

The Fama-French Three-Factor Model Formula

The mathematical representation of the Fama-French three-factor model is:

$$R_p = r_f + \beta_1(r_m - r_f) + \beta_2(SMB) + \beta 3(HML) + \epsilon$$

- R_p : exepected return of the portfolio
- $(r_m - r_f)$: Market risk premium
- SMB : Historic excess returns of small-cap companies over large-cap companies
- HML : Historic excess returns of value stocks (high book-to-price ratio) over growth stocks (low book-to-price ratio)
- ϵ : risk

The result of our Fama model is as following:

Figure 1.14: Fama French 3 Factors Summary

```
                            OLS Regression Results
==============================================================================
Dep. Variable:       Portfolio_return   R-squared:                       1.000
Model:                            OLS   Adj. R-squared:                  1.000
Method:                 Least Squares   F-statistic:                 6.858e+28
Date:                Mon, 30 Oct 2023   Prob (F-statistic):               0.00
Time:                        21:19:22   Log-Likelihood:                 14569.
No. Observations:                 430   AIC:                        -2.913e+04
Df Residuals:                     425   BIC:                        -2.911e+04
Df Model:                           4
Covariance Type:            nonrobust
=========================================================================================
                            coef    std err          t      P>|t|      [0.025      0.975]
-----------------------------------------------------------------------------------------
const                  8.713e-16   6.29e-17     13.853      0.000    7.48e-16    9.95e-16
Market_excess_return      1.0000   1.92e-15   5.21e+14      0.000       1.000       1.000
SMB                   -2.831e-15   2.96e-15     -0.957      0.339   -8.65e-15    2.98e-15
HML                   -4.739e-15   2.38e-15     -1.992      0.047   -9.42e-15   -6.35e-17
rf                        1.0000   1.92e-15   5.22e+14      0.000       1.000       1.000
==============================================================================
Omnibus:                       23.235   Durbin-Watson:                   0.238
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               11.347
Skew:                           0.189   Prob(JB):                      0.00344
Kurtosis:                       2.300   Cond. No.                         461.
==============================================================================
```

We can conclude that:

- The Market Excess Return is 1.00 and is statistically significant ($0.0 < 0.05$).
- The SMB is -2.831e-15 and is not statistically significant ($0.339 > 0.05$).

- The HML is -4.739e-15 and is statistically significant ($0.047 < 0.05$).
- The rf is 1.00 and is statistically significant ($0.00 < 0.05$).

We note that market excess return and risk free rate are statistically significant. According to Fama French's results, we would expect the HML and SMB variables to be significant in explaining portfolio returns. This may be due to the method we employed, which is largely based on very strong assumptions.

Indeed, the studies conducted by Fama and French revealed that the model could explain more than 90% of diversified portfolios' returns. Similar to the CAPM, the three-factor model is designed based on the assumption that riskier investments require higher returns.

1.5 Extension of the Fama French 3 factors model with Macro-economics factors (exogenous factors)

These three factors (Mkt-excess, SMB, HML) are endogenous to the model, meaning they are explicitly included and analyzed as part of the model. They are used to explain differences in stock returns beyond what can be attributed to the overall market return (market risk).

We will extend the model by adding exogenous factors (macroeconomic indicators) such inflation rate, interest rate, and unemployment rate are typically considered exogenous or external factors. The Fama-French model primarily focuses on factors that are endogenous to the financial markets and the specific stocks or portfolios being analyzed.

Even if Fama-French model provides a specific framework for analyzing the returns of stocks in relation to the identified factors within its scope, these macroeconomic factors can indeed have a significant impact on investment decisions and market dynamics.

To collect exogenous data we use the FRED API. We collect data from Consumer Price index and Unemployment Rate of France because we are dealing with French stocks price.

Figure 1.15: Macro Economic factors

consumer_price_index	unemployment_rate
108.12	7.3
108.12	7.3
108.12	7.3
108.12	7.3
108.12	7.3
...	...
117.71	7.4
117.71	7.4
117.71	7.4
117.71	7.4
118.89	7.3

The result of Fama French model is the following:

Figure 1.16: Fama French model with macro economic variables

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     Portfolio_return   R-squared:                       1.000
Model:                          OLS   Adj. R-squared:                  1.000
Method:               Least Squares   F-statistic:                 1.331e+26
Date:              Mon, 30 Oct 2023   Prob (F-statistic):               0.00
Time:                      22:17:52   Log-Likelihood:                 12638.
No. Observations:               408   AIC:                        -2.526e+04
Df Residuals:                   401   BIC:                        -2.523e+04
Df Model:                         6
Covariance Type:          nonrobust
================================================================================
=========
                          coef    std err          t      P>|t|      [0.025
0.975]
--------------------------------------------------------------------------------
---------
const               -1.994e-14    4.7e-14     -0.424      0.671   -1.12e-13
7.24e-14
Market_excess_return    1.0000   3.56e-14   2.81e+13      0.000       1.000
1.000
SMB                 -2.082e-15   5.54e-14     -0.038      0.970   -1.11e-13
1.07e-13
HML                 -4.817e-15   4.44e-14     -0.108      0.914   -9.22e-14
8.25e-14
rf                      1.0000   3.56e-14   2.81e+13      0.000       1.000
1.000
consumer_price_index  1.38e-16   4.38e-16      0.315      0.753   -7.23e-16
1e-15
unemployment_rate   -5.846e-16   3.81e-15     -0.154      0.878   -8.06e-15
6.9e-15
==============================================================================
Omnibus:                     88.578   Durbin-Watson:                   0.000
Prob(Omnibus):                0.000   Jarque-Bera (JB):               21.110
Skew:                         0.231   Prob(JB):                     2.61e-05
Kurtosis:                     1.986   Cond. No.                     1.51e+04
==============================================================================
```

We can observe that the two economic indicators we have added do not explain our portfolio returns because they are not statistically significant. Indeed the p-value of the two macro-economic variables consumer price index and unemployment rate exceed 0.05. Moreover, as with the 3-factor Fama French model (without taking macroeconomic variables into account), we find that the Market excess return and RF factors remain statistically significant.

1.6 Fama French 5 factors

In the latest version of the Fama-French model, two new factors have been added: RMW, or profitability, and CMA, or investment. These additional factors have been the subject of numerous studies.

To perform the Fama French Factors Model with 5 factors, we will use the following dataset from researcher of the Dartmouth University (Kenneth (2014)) and we will perform a Fama French model using this 5 factors on our portfolio we build previously.

The Fama French factors build by the researchers of the university of Dartmouth are constructed using the 6 value-weight portfolios formed on size and book-to-market.

SMB (Small Minus Big) is the average return on the three small portfolios minus the average return on the three big portfolios :

$$SMB = 1/3(SmallValue + SmallNeutral + SmallGrowth) - 1/3(BigValue + BigNeutral + BigGrowth)$$

HML (High Minus Low) is the average return on the two value portfolios minus the average return on the two growth portfolios :

$$HML = 1/2(SmallValue + BigValue) - 1/2(SmallGrowth + BigGrowth)$$

RMW (Robust Minus Weak) factors represents the returns of companies with high operating capacity versus those with low operating capacity. More precisely in our case RMW is the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios.

$$RMW = 1/2(SmallRobust + BigRobust) - 1/2(SmallWeak + BigWeak)$$

CMA (Conservative Minus Aggressive) factor represents the returns of companies with aggressive investments versus those that are more conservative. More precisely in our case is the average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios.

$$CMA = 1/2(SmallConservative + BigConservative) - 1/2(SmallAggressive + BigAggressive)$$

Rm-Rf, the excess return on the market, value-weight return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ that have a CRSP share code of 10 or 11 at the beginning of month t, good shares and price data at the beginning of t, and good return data for t minus the one-month Treasury bill rate (from Ibbotson

Associates).

Figure 1.17: Dataset Fama 5 Factors

Date	Mkt-RF	SMB	HML	RMW	CMA	RF	Portfolio_return
2022-01-04	-0.29	-0.07	3.62	0.63	1.47	0.000	0.009486
2022-01-05	-2.28	-0.94	2.60	1.41	1.22	0.000	0.004433
2022-01-06	0.00	0.38	1.75	-0.55	-0.02	0.000	-0.012526
2022-01-07	-0.48	-0.96	2.02	-0.12	0.85	0.000	-0.006029
2022-01-10	-0.15	-0.30	-0.28	-0.04	-0.38	0.000	-0.014416
...	...	...	...	...	...	...	...
2023-07-25	0.25	-0.23	-0.79	0.47	-0.41	0.022	0.000803
2023-07-26	0.02	0.87	1.03	-0.35	0.65	0.022	-0.003927
2023-07-27	-0.74	-0.80	0.27	0.38	0.14	0.022	0.009114
2023-07-28	1.14	0.41	-0.33	-0.75	-0.40	0.022	-0.009196
2023-07-31	0.26	0.99	-0.10	-0.93	0.05	0.022	-0.006808

394 rows × 7 columns

We apply the following linaer regression:

$$R_p = r_f + \beta_1(r_m - r_f) + \beta_2(SMB) + \beta 3(HML) + \beta 4(RMW) + \beta 5(CMA) + \epsilon$$

The result of the linear regression is the following:

Figure 1.18: OLS Fama 5 factors results

```
                            OLS Regression Results
==============================================================================
Dep. Variable:       Portfolio_return   R-squared:                       0.375
Model:                            OLS   Adj. R-squared:                  0.365
Method:                 Least Squares   F-statistic:                     38.65
Date:                Mon, 30 Oct 2023   Prob (F-statistic):           9.46e-37
Time:                        21:19:22   Log-Likelihood:                 1273.1
No. Observations:                 394   AIC:                            -2532.
Df Residuals:                     387   BIC:                            -2504.
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0005      0.001     -0.550      0.583      -0.002       0.001
Mkt-RF         0.0054      0.000     12.494      0.000       0.005       0.006
SMB            0.0002      0.001      0.268      0.789      -0.002       0.002
HML            0.0032      0.001      4.133      0.000       0.002       0.005
RMW           -0.0010      0.001     -1.185      0.237      -0.003       0.001
CMA           -0.0023      0.001     -1.912      0.057      -0.005    6.47e-05
RF             0.0481      0.065      0.735      0.463      -0.081       0.177
==============================================================================
Omnibus:                       26.248   Durbin-Watson:                   2.234
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               92.117
Skew:                           0.062   Prob(JB):                     9.93e-21
Kurtosis:                       5.366   Cond. No.                         230.
==============================================================================
```

- The Mkt-RF is 0.0054 and is statisticaly significant (2.50e-30 $<$ 0.05)
- The SMB is 0.0002 and is not statisticaly significant (0.789 $>$ 0.05)
- The HML is 0.0032 and is statisticaly significant (4.395e-05 $<$ 0.05)
- The RMW is -0.0010 and is not statisticaly significant (0.236 $>$ 0.05)
- The CMA is -0.0023 and is not statisticaly significant (0.057 $>$ 0.05)
- The RF is 0.0481 and is not statisticaly significant (0.463 $>$ 0.05)

We note that the market risk premium and HML are statistically significant as their p-values are less than 5% (i.e. 0.05). In addition as the p-value of RF is close the 0.05 significance level, we can say that the RF factor is marginally significant.

Overall this result is in line with Fama and French's finding that value companies beat growth companies because the HML variable is statistically significant on portfolio performance.

This means that the portfolio represents a statistically positive exposure (0.0032) of high-value stocks (value companies), compared with the return on value stocks of growth companies.

1.7 Conclusion

We don't achieve to find the same result than Fama and French namely the fact that Fama and French could explain more than 90% of diversified portfolios' returns

In fact the hard part of this assignment is finding new factors and adopting a methodology to compute endogenous risk factors such as SMB and HML factors, testing them and explaining why they make economic sense.

As conclusion we can say that Fama French, are subject to several limitations:

- **Factor Selection and Specification:** The Fama-French model includes three or five factors: the market risk factor (MKT), the size factor (SMB), and the value factor (HML). The problem is that these factors might not be the only relevant ones and that alternative or additional factors may improve the model's performance.
- **Data Period Sensitivity:** The Fama-French model is based on historical data, and the choice of the data period can impact the results. Different time periods may lead to different factor loadings and risk premium, making it challenging to generalize the model across various market conditions.
- **Assumption of Factor Risk Premiums:** The Fama-French model assumes that the risk premiums for the size and value factors remain constant over time. However, these premiums may vary, and some critics argue that they might be influenced by economic conditions and market sentiment.
- **Behavioral Factors:** The Fama-French model primarily relies on rational risk-based explanations for asset pricing. Critics argue that the model does not explicitly incorporate behavioral factors, such as investor sentiment and behavioral biases, which could influence stock prices.
- **Market Efficiency Assumption:** The Fama-French model assumes market efficiency, meaning that all relevant information is reflected in stock prices. However, if markets are not perfectly efficient, the model may not capture certain anomalies or mispricings.

In conclusion, our results are open to interpretation, and there's no reason why we can't find better factors to use.

Please refer to the following links to see the full code of this assignment:

- Empirical Analysis 1 - Fama French Model

2 Empirical Analysis 2

In this report, we will focus on the analysis of stock index prices and macroeconomic indicators, which are crucial for understanding market dynamics and economic trends. Stock indices, like Microsoft (MSFT) and the NASDAQ Composite Index (NDX), are not just aggregates of stock prices; they are reflections of market sentiment, economic health, and investor behavior. Additionally, the inclusion of macroeconomic data, specifically the Producer Price Index by Commodity for Stage of Processing: Intermediate Energy Goods (WPUSOP2910, PPI), provides a broader economic perspective, linking market performance with underlying economic factors. The objective of our study is to explore the complex relationships between these selected stock indices and the macroeconomic series. We aim to identify and analyze patterns, correlations, and potential cointegration relationships among these time-series data. By analyzing the relationships between 'MSFT', 'NDX', and 'WPUSOP2910', we aim to provide valuable insights into the interconnected nature of stock markets and the economy. The report will navigate through the intricate relationships between stock indices and a key macroeconomic indicator, commencing with the rationale for the chosen study period and data frequency, proceeding to the analysis of fundamental series and their stationarity, and culminating in the application of cointegration and error correction models to unveil the latent equilibrium dynamics.

2.1 Period and Frequency Justification

2.1.1 Period : 2005-2013

The period chosen for our analysis, spanning from 2005 to 2013, is particularly significant due to the inclusion of the 2008 financial crisis. This period encompasses not only the crisis itself but also provides a substantial buffer of data before: from 2005 to 2007 offers insights into the behavior of stock indices and the Producer Price Index in a relatively stable economic environment, and after the event: following the crisis, up to 2013, are equally important as they represent a period of recovery and normalization. Such a time frame allows for a comprehensive study of the indices and economic indicators under different economic conditions: pre-crisis stability, the crisis period, and the post-crisis recovery phase. Analyzing data from this period provides critical insights into how stock indices and economic indicators behave under stress and can help in identifying early warning signs of financial instability.

2.1.2 Frequency : montlhy

As for the frequency of the data, we have chosen monthly averages. Monthly data strikes an optimal balance between capturing short-term fluctuations vital during turbulent periods like the 2008 crisis and smoothing out the noise that daily data might present using also log-values of prices. This frequency is sufficient to observe the medium to long-term trends and reactions of the markets to macroeconomic changes, while also being detailed enough to pinpoint significant events and shifts in investor sentiment.

2.2 Representative Fundamental Series

2.2.1 Series plotted

Because the price differences between the different studied financial series spectrum is so large we have chosen to first make a global graph and then an individual graph for each them so as not to encounter scaling issue. We can see that NASDAQ and MSFT seems to have same kind of variations. In the second graph, all three series fluctuate around the zero line, which indicates that the differentiation has well removed any trend the data may have had. The series now appear to be stationary, as they fluctuate within a range without any discernible trend over time. The convergence and divergence of the series at different points in time could suggest periods where the stock prices and the PPI are more or less aligned with each other (price of initial products). For instance, if the lines converge, this may indicate a period where the stock market performance and producer prices were moving in a similar direction, possibly reflecting broader economic trends.

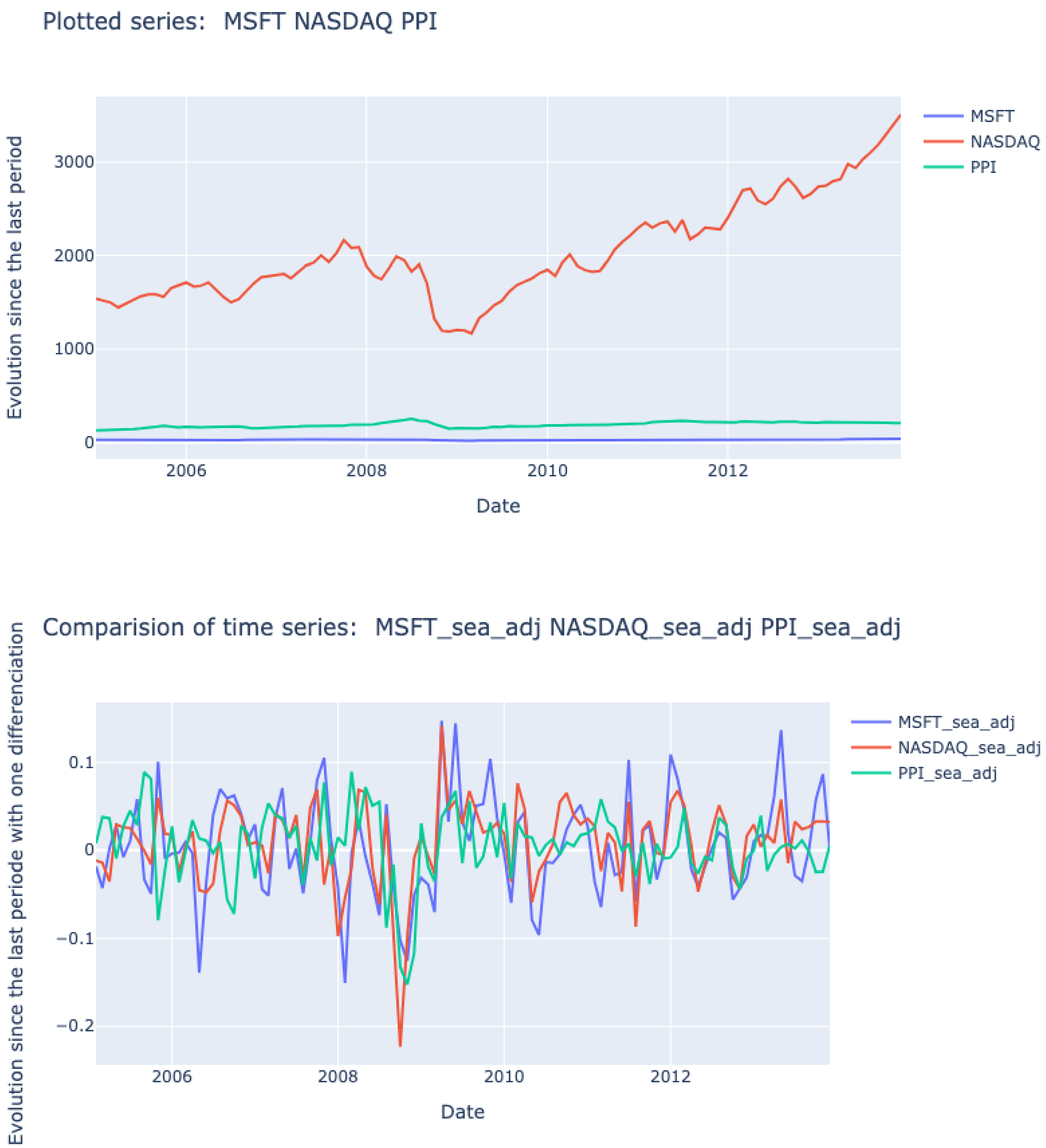


Figure 2.1: Graph of level data and differenciated data

Figure 2.2: Graph of NASDAQ and Microsoft level series

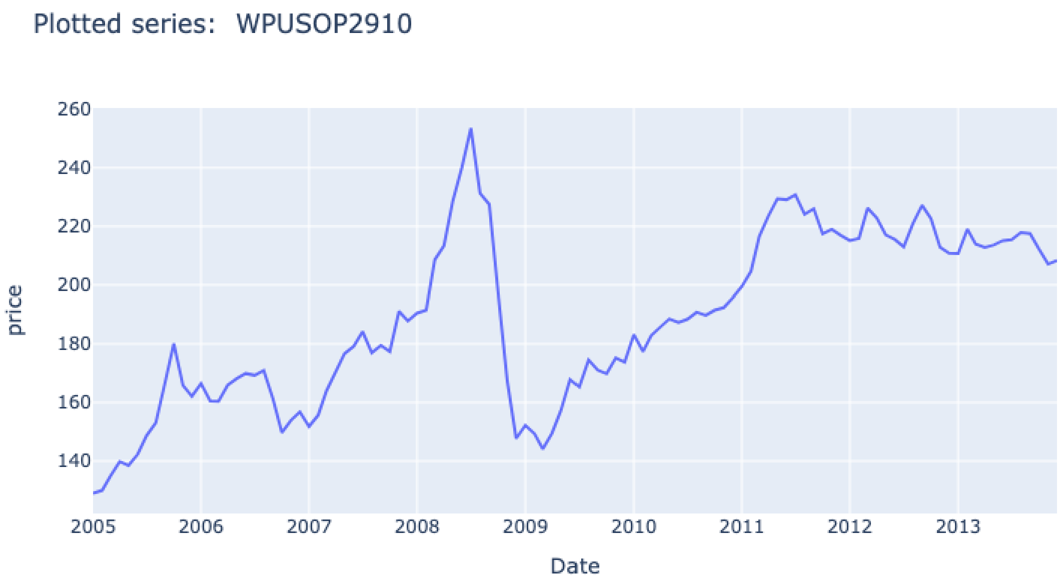


Figure 2.3: PPI plotted

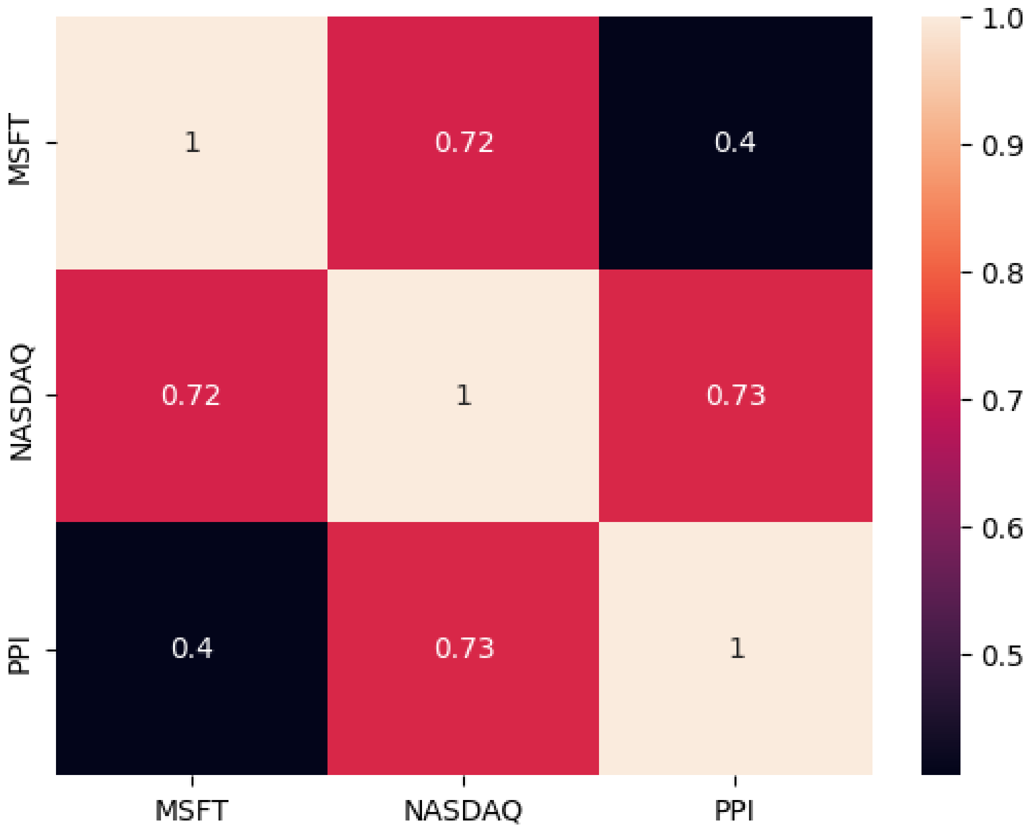


Figure 2.4: Correlation matrix

The first graph, plotting the NASDAQ index, shows a general upward trend in the log price over the examined period. Notably, there is a significant dip around 2008, which aligns with the global financial crisis that led to substantial declines in stock values. Following this dip, the NASDAQ index recovers and continues to rise, indicating a period of growth after the crisis. The bottom graph displays the log price of MSFT. Similar to the NASDAQ index, MSFT also experiences a drop around 2008, which is consistent with the overall market trend during the financial crisis. The recovery for MSFT after 2008 appears to be robust, with the log price reaching higher levels towards the end of the period. This could indicate that Microsoft's stock not only recovered from the crisis but also experienced growth exceeding pre-crisis levels. Both graphs illustrate the volatility inherent in stock prices, with peaks and troughs representing periods of growth and decline. The log transformation aids in highlighting proportional changes, which are more meaningful in financial analysis than absolute changes. The movement of MSFT seems to mirror the general trend of the NASDAQ index, which could suggest that Microsoft's stock performance is somewhat correlated with the broader tech-dominated NASDAQ index. However, MSFT seems to have more pronounced fluctuations compared to the overall index, which may be attributed to company-specific news, product releases, and other factors affecting an individual stock more than an index. In summary, these graphs suggest a potential relationship between the performance of a large technology company's stock and the broader technology-focused stock index.

We choose to make a correlation matrix to see if they are as correlated as it seems. There is a correlation of more than 70% between financial series but also between Producer Price Index and NASDAQ and a quite more than 40% between Producer Price Index and MSFT.

2.2.2 Verification of All Series as I(1)

To determine whether the series are stationary at order 1, we perform an Augmented-Dickey Fuller test.The "order 1" in I(1) indicates that the series becomes stationary after differencing once. Differencing is a transformation applied to the time series where we subtract the current value from the previous value. This process can help to eliminate trends and allow us to see similar evolutions between trends.

To examine whether the series are stationary, we apply ADF test.

Let's recall the test hypotheses of ADF test:

- **Null Hypothesis** (H_0): $\gamma = 0$:Presence of a unit root (the series is not stationary);
- **Alternative Hypothesis** (H_1): γ <0: Non-presence of a unit root (the series is stationary).

Table 2.1: ADF Test with Raw Data

Column	ADF Statistic	P-Value	Stationary
MSFT	-1.95801	0.305282	Not Stationary
NASDAQ	0.418065	0.982158	Not Stationary
PPI	-2.20222	0.20544	Not Stationary

Table 2.2: ADF Test with One Differentiation Data

Column	ADF Statistic	P-Value	Stationary
MSFT_diff	-7.45093	5.67342e-11	Stationary
NASDAQ_diff	-7.28568	1.46104e-10	Stationary
PPI_diff	-5.42106	3.05876e-06	Stationary

The results of the ADF test suggest that the series have a unit root when they are not differentiated, where the test statistics and p-values (>0.05) show that we cannot reject the null hypothesis of the presence of a unit root. We then differentiate our series once and repeat the test. The ADF test's t-statistics and p-values (<0.01) indicate that we can reject the null hypothesis of non-stationarity. The series are therefore stationary at the 1% level and series are integrated of order 1.

2.3 Gordon-Shapiro Model and Its Development

In the context of the series 'MSFT','^ NDX', and Producer Price Index by Commodity for Stage of Processing: Intermediate Energy Goods, the Gordon-Shapiro model can be used to establish a long-term equilibrium relationship. This is based on the assumption that these series are influenced by similar market and economic factors and thus, should move together in the long run. The Gordon-Shapiro model, or the Dividend Discount Model (DDM), can be integrated into this framework. This model posits that the value of a stock is equal to the present value of all its future dividends. In a cointegration and

VECM context, we could consider the stock price, dividends, and other economic factors as time series that potentially have a long-term equilibrium relationship.

2.3.1 Long-Term Equation and Structural Break Tests

We've tried to create a variable that models the shock (in this case, the 2008 crisis) starting on July 1, 2007 and ending on July 1, 2007. We then performed cointegration tests, without success. So we decided not to pursue the analysis further. The cointegrating relationship is a long-term equilibrium relationship between the variables, which remains stable despite short-term fluctuations. If the variables deviate from this equilibrium relationship, the error correction term guides the variables back to their long-term equilibrium relationship. The cointegration relationship is given by:

$$\mathrm{MSFT} = 1 - 0.0016 \times \mathrm{NASDAQ} + 0.1271 \times \mathrm{PPI} \tag{2.1}$$

This means that if MSFT, NASDAQ and PPI deviate from this relationship, the error correction term (ec1) becomes non-zero, indicating an imbalance. The coefficients in the VECM equation associated with ec1 (-0.2186 for MSFT, -0.1555 for NASDAQ and 2.8533 for PPI) represent the speed at which each variable returns to equilibrium after a shock. For example, if ec1 is non-zero, MSFT's price will adjust by -0.2186 * ec1 in the next period to return to equilibrium.

2.3.2 Error Correction Equation Estimation

To achieve this, we employ advanced statistical techniques, including the Johansen Cointegration Test and the Vector Error Correction Model (VECM). These methods are particularly effective for time-series data analysis, allowing us to investigate long-term relationships and short-term dynamics among the series.The Cointegrated Augmented Dickey Füller (CADF) test enables one to test for cointegration between two-time series and the presence of unit root. The Johansen Test can be used to check for cointegration between a maximum of 12-time series. The Johansen Cointegration Test, as explained in Ahlgren (2010), helps us determine whether a group of non-stationary series, such as stock prices, are cointegrated, meaning they have a long-term equilibrium relationship. Concerning cointegration, before applying the VECM model we perform a cointegration test (Johansen) to find out how many variables are cointegrated with each other:

- Null Hypothesis H_0: The number of cointegration vectors is equal to or less than r
- Alternative Hypothesis H_1: The number of cointegration vectors is equal to or greater than r

where, r denotes the rank of the cointegration matrix, representing the number of linearly independent cointegration vectors.

Table 2.3: Johansen Cointegration Test

	Eigenvalues	Test Statistics	Critical Values (90%, 95%, 99%)
3 cointegrated variables	0.15207782	29.8187421	[27.0669, 29.7961, 35.4628]
2 cointegrated variables	0.09596783	12.33230136	[13.4294, 15.4943, 19.9349]
1 cointegrated variable	0.01533336	1.63792603	[2.7055, 3.8415, 6.6349]

For the first eigenvalue, the test statistic (29.8187421) is greater than the critical value (29.7961), so we reject the null hypothesis of no cointegration at this level. For the second eigenvalue, the test statistic (12.33230136) is less than the critical value (15.4943), so we cannot reject the null hypothesis of at most 1 cointegrating relationship at this level. For the third eigenvalue, the test statistic (1.63792603) is less than the critical value (3.8415), so we cannot reject the null hypothesis of at most 2 cointegrating relationships at this level.

We conclude that only the three variables have a cointegration relationship. The variables have neither a two-by-two cointegration relationship nor a single one with itself.

This means that while pairs of variables or individual variables do not share a cointegrating relationship, the trio as a whole moves together in the long term. This finding aligns with the general unpredictability of stock prices, which while subject to individual fluctuations, tend to exhibit interconnectedness over an extended period, thereby maintaining a form of equilibrium as a collective.

The VECM builds on this understanding by providing a framework to model the short-term dynamics while accounting for the long-term cointegration. Through the VECM, we'll analyze how the variables adjust and revert to their long-term equilibrium following short-term deviations.

2.3.3 VECM results

We have seen that there is a cointegration relationship between the three variables. So we have to apply the VECM model. From the table summarizing the model we can say that we have three equations, the following:

$$\begin{aligned}\text{MSFT equation:}\quad \text{MSFT} &= 0.2525 \times L1.\text{MSFT} + 0.0028 \times L1.\text{NASDAQ} \\ &+ 0.0025 \times L1.\text{PPI} - 0.0563 \times \text{ec1}\end{aligned} \tag{2.2}$$

$$\begin{aligned}\text{NASDAQ equation:}\quad \text{NASDAQ} &= 0.5226 \times L1.\text{MSFT} + 0.2958 \times L1.\text{NASDAQ} \\ &+ 0.5003 \times L1.\text{PPI} - 0.0016 \times \text{ec1}\end{aligned} \tag{2.3}$$

$$\begin{aligned}\text{PPI equation:}\quad \text{PPI} &= -1.5906 \times L1.\text{MSFT} + 0.0446 \times L1.\text{NASDAQ} \\ &+ 0.2030 \times L1.\text{PPI} + 0.4821 \times \text{ec1}\end{aligned} \tag{2.4}$$

L1.correspond to one period lag (one month in our study).

The VECM results suggest that there is cointegration among the three variables, which implies that they share a long-run equilibrium relationship.

Figure 2.5: VECM Results

	coef	**std err**	**z**	**P> |z|**	**0.975]**
L1.MSFT	0.2525	0.125	2.026	0.043	0.497
L1.NASDAQ	0.0028	0.002	1.297	0.195	0.007
L1.PPI	0.0025	0.018	0.139	0.890	0.038
L1.MSFT	0.5226	7.644	0.068	0.945	15.505
L1.NASDAQ	0.2958	0.131	2.265	0.024	0.552
L1.PPI	0.5003	1.115	0.449	0.654	2.686
L1.MSFT	-1.5906	0.584	-2.722	0.006	-0.445
L1.NASDAQ	0.0446	0.010	4.464	0.000	0.064
L1.PPI	0.2030	0.085	2.381	0.017	0.370
ec1	-0.0563	0.035	-1.621	0.105	0.012
ec1	-0.3931	2.131	-0.184	0.854	3.784
ec1	0.4821	0.163	2.959	0.003	0.801
beta.1	1.0000	0.000	0.000	0.000	1.000
beta.2	-0.0016	0.003	-0.455	0.649	0.005
beta.3	-0.1271	0.037	-3.396	0.001	-0.054

From the results table and the equations provided, we can see that each equation represents the relationship between one of the variables (MSFT, NASDAQ, PPI) and its lagged value, the lagged values of the other variables, and an error correction term (ec1). The error correction term (ec1) represents the speed at which the variables return to equilibrium after a shock.

The MSFT equation suggests that Microsoft's stock price is positively related to its own lagged price and also slightly to the lagged NASDAQ index and the lagged Producer Price Index (PPI). The NASDAQ equation indicates that the NASDAQ index is positively related to the lagged MSFT, its own lagged value, and the lagged PPI. The PPI equation shows that the PPI is negatively related to the lagged MSFT stock price, positively to the lagged NASDAQ index, and to its own lagged value, along with a positive relationship to the error correction term.

However, we can note that only the PPI equation shows all coefficients as significant at the 10% threshold, which suggests that the PPI equation is reliable for making inferences at this level of significance. The lack of significance in other coefficients implies that the MSFT and NASDAQ equations might not be reliable predictors of these variables based on the past month's lagged values alone, or at least not at the 10% significance level.

In conclusion, the VECM analysis indicates a long-term equilibrium relationship among MSFT, NASDAQ, and PPI, with the PPI being responsive to deviations from equilibrium. The results provide insights into the dynamics of these variables over time, but caution should be exercised when using MSFT and NASDAQ equations for predictions, as not all their coefficients are significant at the chosen level of confidence.

2.4 Conclusion

In conclusion, our empirical application centered on the fundamental analysis of a stock index, with a focus on establishing a long-term equilibrium relationship among the selected variables. We carefully selected the period from 2005 to 2013 to provide a comprehensive understanding of market behavior under different economic conditions, without specifically considering structural breaks during the 2008 crisis.

Our analysis indicated that the selected series, including log prices of the selected stocks series such as MSFT, NDX, and the Producer Price Index for Intermediate Energy Goods

(WPUSOP2910), are integrated of order 1 (I(1)), making them suitable for time-series analysis. While we did not directly employ the Gordon-Shapiro model, we used its insights to guide our approach in modeling a long-term equation that captures potential relationships among these variables.
In addition to our cointegration analysis, we acknowledge the importance of forecasting. As we look to the future, we aim to enhance our predictive models and economic forecasts by incorporating the elastic net technique, which blends both the strengths of Ridge and Lasso regression. As discussed by Ashfaq et al. (2021), machine learning regressors have shown promising results in stock market predictions. This innovative approach promises a more nuanced model selection and complexity, enabling a more precise capture of the long-term linkages between stock indices and macroeconomic indicators.

In summary, our study emphasizes the importance of establishing long-term equilibrium relationships among financial and economic variables. While we did not delve into the specific structural breaks during the 2008 crisis, our analysis provides valuable insights into the potential long-term connections between these variables. These findings have implications for understanding market dynamics, forecasting economic trends, and offering valuable insights for investors and policymakers.

Please refer to the following links to see the full code of this assignment:

- Empirical Analysis 2 - Cointegration Theory

3 Empirical Analysis 3

To begin our study, we'll focus on implementing the approach detailed in the document "Pricing and Simulation Method." This approach centers on determining the value of derivatives in scenarios where traditional closed-form equations fall short. Our discussion is divided into two main parts:

We'll start by briefly examining the "Put Lookback Option" as described in the "Pricing and Simulation Method" document. This option, which is influenced by the lowest historical value of the underlying asset, poses a unique challenge for valuation due to its path-dependent characteristics. We'll look at how simulation methods are utilized to estimate its fair price, and the idea is to well understand the Monte Carlo model.

Following this, we shift our attention to a different type of derivative, the "Call Shout Option." This option grants the holder the right to 'shout,' or secure a certain profit at a chosen point during the option's lifetime, without immediate exercise. The valuation of this option is more complex due to the additional strategic choice it offers. We'll discuss adapting simulation methods from the context of Put Lookback Options to handle this new challenge, underlining the importance and flexibility of simulation techniques in the pricing of complex financial instruments.

3.1 Put Lookback Option: Application of the Method

3.1.1 Some definitions

What is a Put Lookback Option?

A put lookback option is a type of financial option that allows the buyer to sell an asset at its lowest price during the life of the option. Unlike standard put options which are exercised at a fixed strike price determined in advance, lookback options offer additional flexibility by allowing the buyer to benefit from the most advantageous price reached by the asset during the option period. The price of a put lookback option is complex to determine because it depends not only on the current price of the underlying asset but also on its minimum price during the entire option period.

Understanding Closed Formulas in Financial Mathematics

In financial mathematics, a closed formula is an explicit equation that provides a direct means to calculate the price of a derivative or the value of a financial instrument. Such formulas are highly prized for their ability to yield immediate results without necessitating complex numerical methods or extensive computations.

A classic example of a closed formula is the Black-Scholes formula for pricing European call and put options. Closed formulas are derived from continuous-time models and typically involve elements like exponential, logarithmic, and trigonometric functions, which are well-understood and easily computable.

The existence of a closed formula depends on the complexity of the financial model and the assumptions underlying it. Closed formulas are possible when the model is mathematically tractable and all the parameters are known with certainty. However, in many real-world scenarios, the assumptions required to derive a closed formula do not hold, or the model itself is too complex, leading to situations where closed formulas cannot be applied. In such cases, numerical methods such as the Monte Carlo simulation or binomial trees are used to approximate the value of the financial instrument. These are the models we are going to test in the second part of this study.

3.1.2 Method Used to Estimate the Fair Price

To calculate the fair price of a lookback put option, the method involves discounting the risk-neutral expectation of the difference between the maximum price of the underlying asset during the option period and its price at maturity, as explained by the following formula:

Fair Price of the Put Lookback Option

The fair price of the put lookback option is given by the following equations:

$$G_0 = e^{-rT}\mathbb{E}_0^*\left[\max_{t\in[0,T]}(P_t) - P_T\right] \tag{3.1}$$

This can also be represented as:

$$G_0 = e^{-rT}\mathbb{E}_0^*\left[\max_{t\in[0,T]}(P_t)\right] - e^{-rT}\mathbb{E}_0^*(P_T) \tag{3.2}$$

And further simplified to:

$$G_0 = e^{-rT}\mathbb{E}_0^*\left[\max_{t\in[0,T]}(P_t)\right] - P_0 \tag{3.3}$$

where G_0 is the fair price of the put option, e^{-rT} is the discount factor, $\mathbb{E}_0^*$ represents the risk-neutral expected value, $\max_{t\in[0,T]}(P_t)$ is the maximum price of the underlying asset over the period $[0, T]$, P_T is the price of the underlying asset at maturity, and P_0 is the initial price of the underlying asset.

Model for the Behavior of the Underlying Asset Price

The behavior of the underlying asset price is modeled by the following stochastic differential equation in a risk-neutral world:

$$dP_t^* = rP_t^*dt + \sigma P_t^*dW_t^* \tag{3.4}$$

where:

- dP_t^* represents the infinitesimal change in the risk-neutral price of the asset at time t,
- r is the risk-free rate, which is assumed to be constant over time,
- dt is an infinitesimal increment of time,
- σ is the volatility of the asset's returns, assumed to be constant,
- P_t^* is the risk-neutral price of the asset at time t,
- dW_t^* is the increment of a Wiener process (or Brownian motion) under the risk-neutral measure, representing the random component of the asset price change.

This equation accounts for the asset price movements over time, factoring in both the constant risk-free rate of return (r) and the asset's constant volatility (σ). The term $\sigma P_t^*dW_t^*$ captures the uncertainty or risk inherent in the asset price's trajectory.

To compute the risk-neutral expectation necessary for ascertaining the put option's fair value, a simulation-based method is recommended. This involves:

1. Generating simulated values for the asset price following the dynamics described by the equation.
2. Estimating the mathematical expectation from these simulated values.

The process includes simulating numerous potential price paths for the asset, determining the maximum price reached in each path over the specified period, and computing the average of these maximum values. This average, once discounted at the risk-free rate, yields the estimated fair price of the put option.

3.1.3 The Monte-Carlo Method

This section describes the application of the Monte Carlo method to the pricing of financial options. For the put lookback option, the method is used to simulate multiple potential scenarios for the price of the underlying asset over a specified period to estimate the option price. The Apple ticker from Yahoo Finance is selected as the underlying asset.

Why this method ?

The Monte Carlo method is particularly well suited to lookback options due to the path-dependent nature of these options. It allows the dynamics of financial asset prices to be modelled realistically and takes account of market uncertainty and volatility. In addition, this approach offers great flexibility and can be adapted to various types of complex derivatives.

Application from the paper

We establish fundamental parameters such as the initial price of the Apple ticker $P_0 = 190.64$ (extracted from Yahoo Finance), the risk-free interest rate $r = 0.05$, the asset's volatility $\sigma = 0.22$ (also taken from Yahoo Finance for the Apple ticker), the time to maturity $T = 1$ year, the number of simulations $N = 1,000,000$, and the number of time steps $n = 252$ which corresponds to the number of trading days per year.

For each simulation, a price trajectory of the asset is generated, simulating the price at regular intervals up to the maturity. At each interval, the price is modified according to a random draw from a normal distribution to reflect the price movement's uncertainty. The asset price S is updated at each time step dt using the equation:

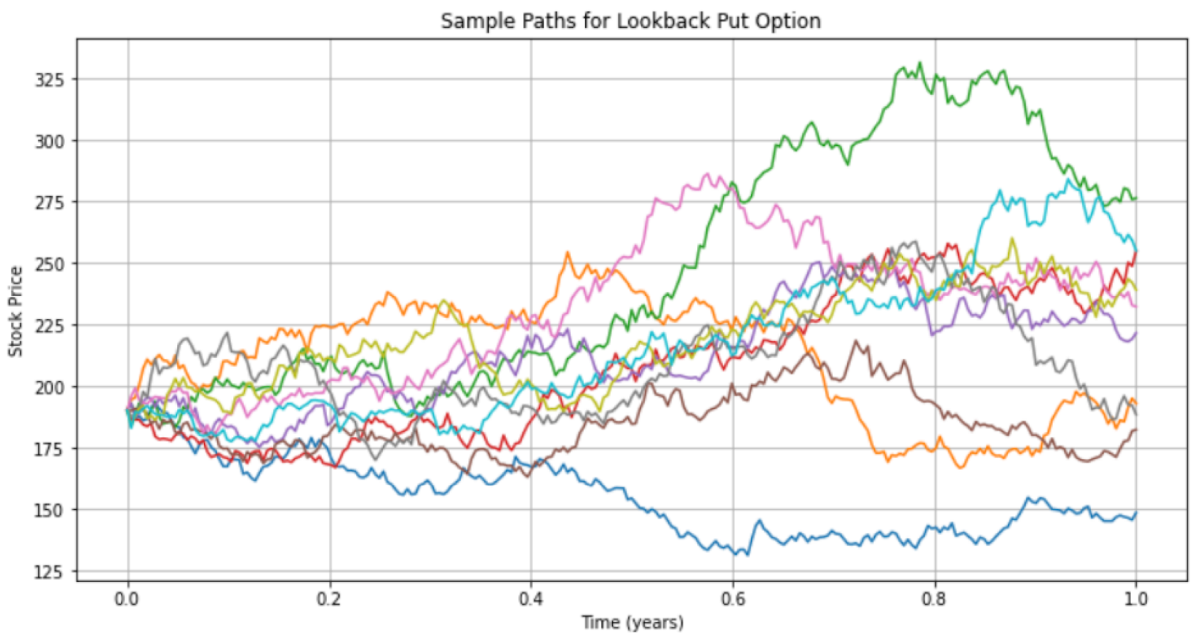
$$S = S \times \exp(u), \tag{3.5}$$

where u is a value randomly drawn from a normal distribution with mean $(r - 0.5 \times \sigma^2) \times dt$ and standard deviation $\sigma\sqrt{dt}$. This approach to generating u mirrors the dynamics of a Geometric Brownian Motion (GBM).

The GBM model here simulates various possible paths for the price evolution of the underlying asset. Each simulation is a potential price movement scenario, with each line representing a possible path for the price of the underlying asset over time, based on the specified parameters.

For each simulated path, the maximum asset price reached during the period is noted. This is crucial for determining the payoff of the put lookback option.

Figure 3.1: Simple Paths for lookback Put Option



This graph provides a visualization of how the asset price evolves over time across different simulations, encapsulating the uncertainty and volatility inherent in the financial markets. Lookback options are distinguished by their path-dependent nature, meaning that their value is contingent not just on the current price of the underlying asset but also on its historical prices. The chart exemplifies this dependency by depicting the various paths the asset price might take and the corresponding maximum prices reached. It illustrates the impact of changes in input parameters such as volatility and interest rates on the asset price trajectories, offering an intuitive understanding of the option's sensitivity to market conditions.

The estimated price of the put lookback option on the Apple ticker, derived from these simulations, is **$39.04**. This figure represents the theoretical price an option buyer would need to pay to acquire this lookback put option under the given simulated market conditions. It encapsulates the present value of the potential profits that the buyer could realize by strategically exercising the option throughout its lifespan, considering the minimum price of the underlying asset during this time frame.

3.2 Pricing of Shout Option using Other simulations Methods

3.2.1 What is a Shout Option?

Shout options are a type of European options where the holder has the right to 'shout' to their counterparty once during the life of the contract. At expiration time, the holder receives the greater of the traditional European option payoff or the intrinsic value at the time of the shout. In the valuation procedure, it is indeed necessary to investigate the optimal shouting policy and determine the critical stock price at which it is optimal to shout (Kwok, 2003). We are going to work on a Call Shout Option.

Example of a Shout Option

Let us consider an example with the following option parameters, maintaining the same ticker, N, r, steps, and n as in our previous model:

- Ticker: AAPL
- Strike Price: $185
- Current Stock Price (S_0): $180
- Option Premium: $11 (or $1,100 per contract)

- Expiration Date: 3 months later ($T = \frac{1}{4}$ year)
- Risk-Free Interest Rate (r): 0.05
- Number of Simulations (N): 100
- Number of Time Steps in Monte Carlo Simulation (n): 252
- Number of Steps in Binomial Tree (steps): 30
- The trader is allowed to shout once during the term of the option.

The trader's break-even point is at \$196 (\$185 strike price + \$11 premium). Suppose one month after purchase, AAPL stock trades at \$193. The trader decides to shout, securing an intrinsic value of \$8 (\$193 - \$185). This guarantees that the trader will not lose the entire premium, retaining at least \$8 of it.

Scenarios After the Shout

3.2.1.1 If the price drops below \$193 and remains there until expiry:

- The trader receives the locked-in intrinsic value of \$8.
- They incur a loss of \$3 per contract (\$11 premium - \$8 intrinsic value), which is \$300 per contract.

3.2.1.2 If the price of AAPL rises to \$205 at expiration:

- The option's intrinsic value is \$20 (\$205 - \$185).
- The trader collects \$20 per contract, or \$2,000, despite having shouted for \$8 in intrinsic value.
- The profit is \$9 per contract (\$2,000 collected - \$1,100 premium), or \$900 per contract.

Determining a Fair Price

We will analyze a Call Shout option with AAPL as the underlying asset. To determine the fair price of this shout call option, the process involves considering the option's value if it is exercised ("shouted") at date t with the underlying asset valued at S_t on that date.

$$\max(0; S_{\max} - S_t) + (S_T - K)$$

3.2.2 Inefficacy of Closed Formulas

Closed formulas, while efficient for plain vanilla options, fail to capture the nuanced behaviors of more complex derivatives. For example, the Black-Scholes model assumes a continuous-time framework and does not account for path-dependent options like lookback or shout options, where the payoff is contingent on the asset price history. In the context of financial option pricing, the traditional Black-Scholes approach, although revolutionary for its ability to provide a closed formula for standard European options, shows its limitations when it comes to more complex and path-dependent products, such as shout options. The latter, offering the holder the unique right to 'shout' or lock in an intrinsic value before maturity, introduce a dimension of strategic choice that is not captured by the standard Black-Scholes assumptions.

Black Scholes Model, critics and results

Moreover, according to (Hull, 2009, chapter 26, section 26.12, page 637), the value of the option in the event of a shout at time t is thus the present value of $S_t - K$ (the amount received at maturity T), plus the value of a European call option with exercise price S_t. This value can be obtained using the Black-Scholes formulas. Therefore, we will attempt to estimate the price of the option using this initial method to compare it with non-closed form methods to understand the benefits of the other methods.

In the context of the provided estimate for the AAPL shout call option price using the Black & Scholes model, we see a value of approximately **\$21.90**. This price reflects the current market valuation while incorporating the strategic advantage the holder gains by having the opportunity to lock in profits before the expiration date, an aspect not available in standard European options.

The model exhibit significant limitations when dealing with complex options like shout options. These limitations largely stem from the model's simplified assumptions, like constant volatility and neglect of dividends, which may not always reflect market realities, especially for derivative products offering additional flexibility like shout options. Moreover, Black-Scholes fails to incorporate behavioral factors and unforeseen market events which can influence the optimal timing for exercising a shout option - a critical decision to maximize the option's value for the investor.

Necessity of Open Formulas and Simulation Methods

Given the limitations of models like Black-Scholes, the need arises to explore open formulas and simulation methods. Partial Differential Equations (PDEs) offer a more robust approach to model the complex dynamics of shout options. They allow for the incorporation of variables such as volatility changes and strategic decisions of option holders. Moreover, simulation methods, such as Monte Carlo simulations or binomial tree models, are particularly suited for studying shout options. These methods simulate a wide range of possible market scenarios, thus providing a more accurate estimation of the option's value while accounting for strategic decisions like early exercise. These approaches offer increased flexibility and precision, essential for accurately assessing the value of shout options and other complex derivative products.

3.2.3 The Monte Carlo Method adapted to the Call Shout Option

We employed the same methodology for the shout call option as we did for the put lookback option, with a few notable differences. The primary distinction in our approach and coding for the two simulations lies in the manner in which the payoff is calculated for each simulated price path. For lookback options, the focus is solely on the maximum price achieved. In contrast, for shout options, the calculation includes an early exercise decision, represented by the 'shout value', which is assessed and updated at each step of the simulation.

Formula for the Shout Call Option Payoff

The formula for calculating the payoff of a shout call option at each step in the simulation is given by:

$$\text{Payoff} = \max(\text{shout value, final intrinsic value}) \tag{3.6}$$

The 'Shout Value' is defined as the best possible payoff if the option is exercised (shouted) at any point before maturity. For a shout call option, it is calculated as:

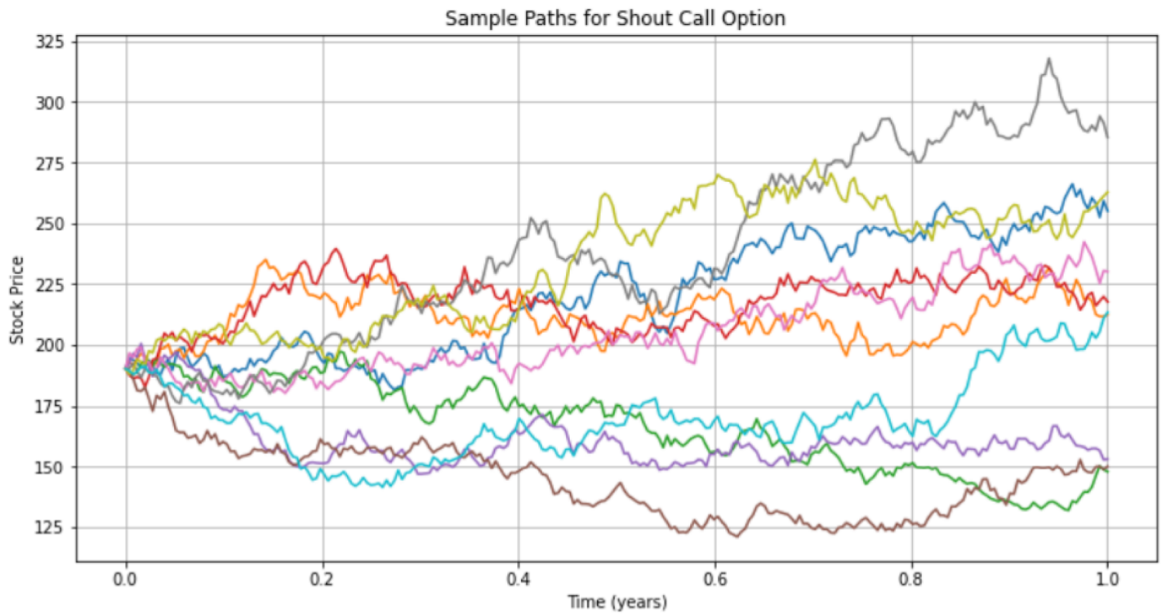
$$\max(0, S_t - K) \tag{3.7}$$

where S_t is the price of the underlying asset at step t. This value is continually updated to reflect the best possible payoff until that moment. The 'Final Intrinsic Value' represents the intrinsic value of the option at maturity if it has not been exercised beforehand. For a shout call option, it is defined as:

$$\max(0, S_T - K) \tag{3.8}$$

where S_T is the asset price at maturity.

Figure 3.2: Simple Paths for Shout Call option



In this context, K represents the strike price of the option, which in our example is taken as \$190.64, the initial price of the Apple ticker. For a call option, if the underlying asset's price (S_t or S_T) is above K, the option possesses positive intrinsic value and is considered "in the money."

The estimated fair price of the Shout Call option on the Apple ticker, derived from these simulations, is **\$39.11**.

Limit of the Monte Carlo model

According to the paper "Shout Options: A Framework for Pricing Contracts Which Can Be Modified by the Investor" (Windcliff et al., 2001), traditional Monte Carlo methods face specific limitations when valuing shout options. Shout options are unique in that they allow investors to modify the contract or secure part of the intrinsic value before maturity. While Monte Carlo methods are effective in simulating a range of market scenarios, they struggle to efficiently handle the optimization component inherent in these contracts. Primarily designed to generate asset price trajectories and calculate an average of outcomes, these methods are not well-suited to the conditional and strategic decisions crucial for determining the optimal exercise time for a shout option. This inability to dynamically integrate optimal decisions at each trajectory point is a major limitation of Monte Carlo methods in accurately valuing shout options, where decisions are influenced by complex factors such as market movements and volatility forecasts. Furthermore, the one of main limitation of Monte Carlo method is this method is very expensive in terms of computing power and time.

3.2.4 The Binomial and Trinomial Tree Simulation

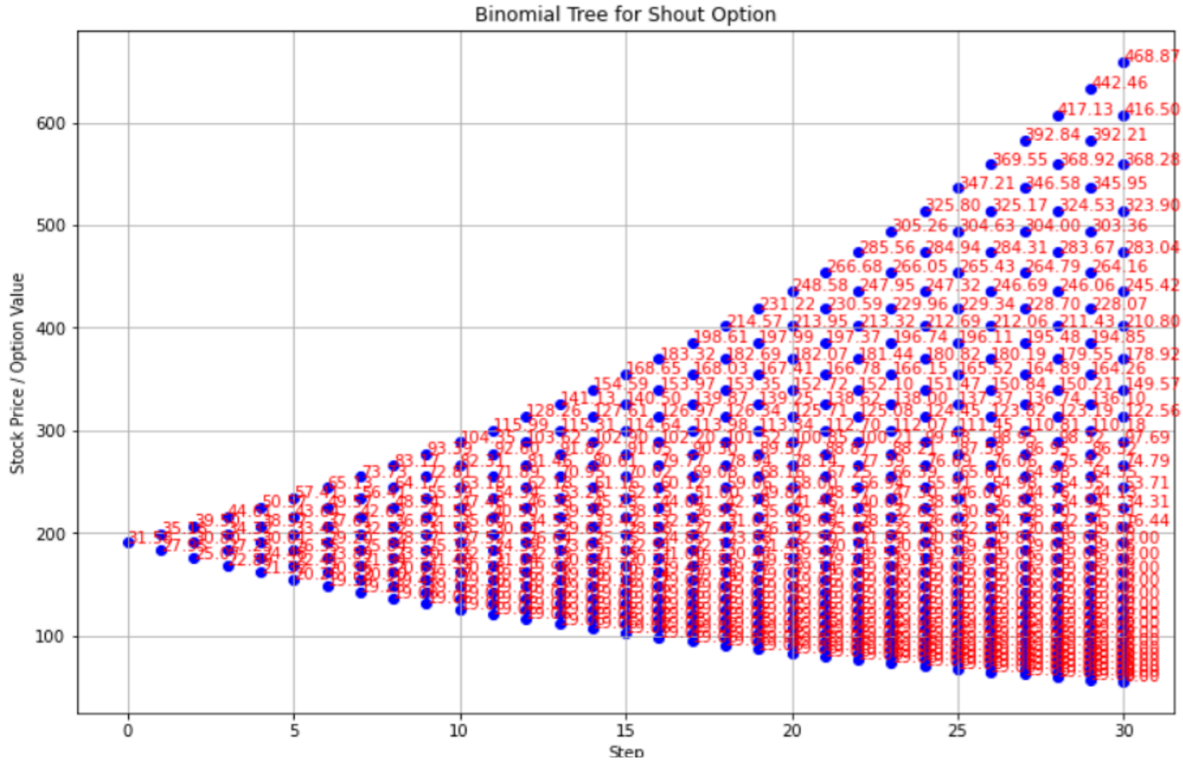
Shout options can be valued in the usual way in a binomial or trinomial price tree for the underlying asset. Working from the terminal date towards the initial date, the value of the option in the event of a call and its value in the event of a put are calculated at each node; the value retained is of course the higher of the two. The procedure for valuing a shout option thus appears relatively similar to that for traditional American options.

The Binomial Tree

The binomial tree models the potential future movements of the stock price over a set number of discrete time steps leading up to the option's expiration. At each step, the stock price can move up by a factor of u or down by a factor of d, with $u = e^{\sigma\sqrt{dt}}$ and $d = \frac{1}{u}$. The probability of an upward move is p, calculated as $\frac{e^{r \cdot dt} - d}{u - d}$.

For the shout option, at each node, we evaluate the option to "shout," or lock in the intrinsic value, compared to holding the option. The shout value at any node is the maximum of the current shout value and the intrinsic value at that node for a call $\max(0, S_t - K)$.

Figure 3.3: Binomial Tree for Shout Option

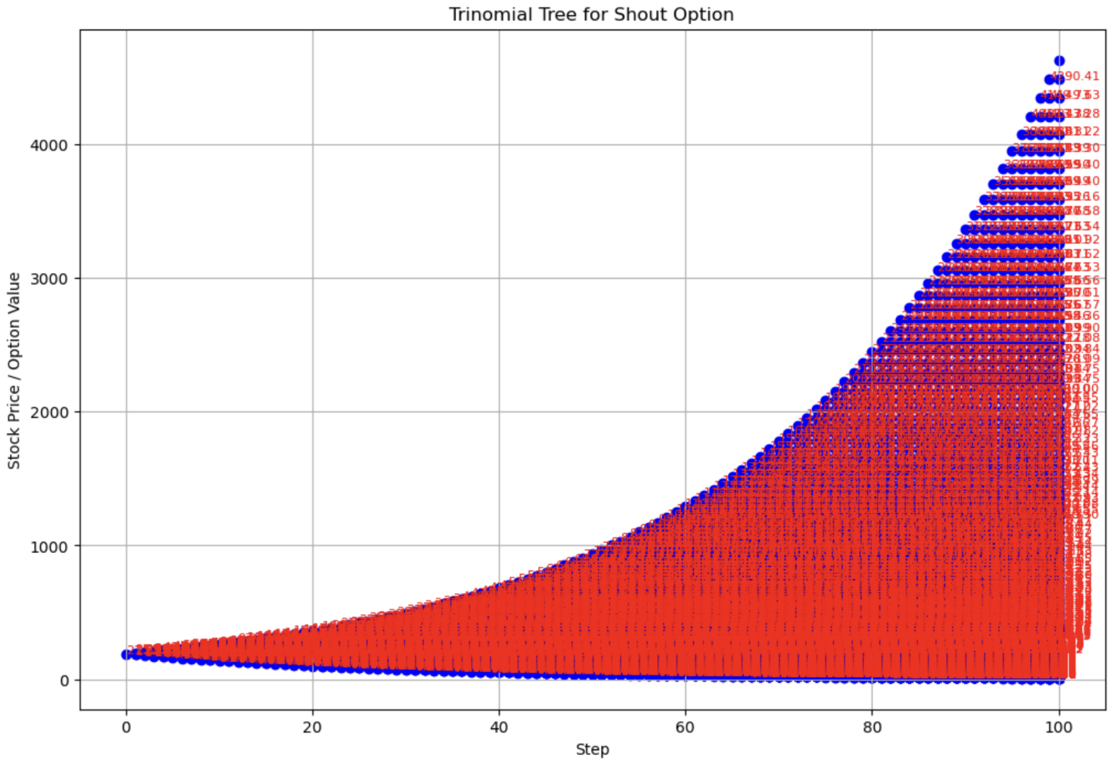


The graph associated with the binomial tree visually represents the evolution of the option values at each node, overlaid with the possible stock prices at each step of the tree. Each node displays the stock price and the corresponding option value, based on whether the option is exercised or held at that point. The final option price is the value at the root of the tree, which represents the present value of expected payoffs, discounted back at the risk-free rate. The binomial tree effectively illustrates all possible paths the stock price can take and the decisions to "shout" or hold the option at each step, leading to the valuation of the shout option. The Binomial Tree Shout Call fair Price for AAPL is **$21.65**.

The Trinomial Tree

The trinomial tree model extends the binomial approach by introducing a third possible movement at each step, where the stock price remains unchanged. In this model, at each time step, the stock price can either move up by a factor of u, move down by a factor of d, or stay the same. This additional movement adds complexity but also provides a more nuanced representation of stock price movements, potentially making the model more accurate, especially for options with longer durations or more volatile underlying assets. The trinomial tree thus incorporates an additional probability pm for the stock price remaining constant, alongside the probabilities of upward and downward movements. The Trinomial Tree Shout Call fair Price for AAPL is also **$21.67**.

Figure 3.4: Trinomial Tree for Shout Option




```
For the Trinomial Tree on the Shout Call Price for AAPL, we estimated: 21.67214021686045
```

3.2.5 Conclusion

This study, focused on the valuation of shout options, particularly Call Shout Options, highlights the unique challenges and opportunities they present in the field of financial derivatives. The analysis has demonstrated the effectiveness of various simulation methods, such as Monte Carlo simulations and binomial/trinomial tree models, in estimating the fair price of these options.

Shout options stand out due to their inherent "shout" feature, which allows locking in intrinsic value before maturity. This attribute adds a strategic dimension not present in standard European options, necessitating advanced valuation methods beyond traditional closed-form formulas like Black-Scholes.

The study sheds light on the effectiveness of simulation methods in capturing the nuanced

Method	Strengths	Limitations	Estimated Price
Monte Carlo	Dynamic modeling, adaptability	Complexity in optimization, computational cost	\$39.11
Binomial Trees	Simplicity, easy to understand	Less precise for certain option types, slower convergence for American options	\$21.65
Trinomial Trees	Better convergence properties, handles more complex option types effectively	More complex than binomial model, higher computational demand	\$21.67

Table 3.1: Comparative Table of Valuation Methods

behaviors of shout options. The Monte Carlo method, for example, excels in modeling the dynamics of financial asset prices and adapts well to the path-dependent nature of these options. However, it has limitations in efficiently handling the optimization component inherent in shout options.

The shout feature offers a strategic advantage to the holder, allowing them to lock in profits under favorable market conditions. This characteristic makes the valuation sensitive to market movements and volatility forecasts.

The study noted price discrepancies between different models, underscoring the importance of choosing an appropriate model based on the specific characteristics of the option and market conditions.

Given the limitations of closed-form models like Black-Scholes for complex options, the study emphasizes the need for open formulas and diverse simulation methods. These methods as binomial/trinomial tree models, offer increased flexibility and precision, essential for accurately assessing the value of shout options and other complex derivative products.

Looking ahead, partial differential equations represent a promising path for the valuation of shout options, offering the ability to integrate complex features and dynamic market behaviors. The examination of alternative methods mentioned in the "Shout Options: A

Framework for Pricing Contracts Which Can Be Modified by the Investor" (Windcliff et al., 2001), such as PDE-based approaches, enhances our understanding of financial options and contributes to the development of more robust financial models and strategies.

Please refer to the following links to see the full code of this assignment:

- Empirical Analysis 3 - Simulation method

References

Ahlgren, N. (2010). Testing for cointegration between international stock prices. *International Stock Prices Journal*.

Ashfaq, N., Nawaz, Z., and Ilyas, M. (2021). A comparative study of different machine learning regressors for stock market prediction.

Hull, J. C. (2009). *Options, Futures, and Other Derivatives*. Pearson Education.

Kenneth (2014). Description of fama french.

Kwok, Y. K. (2003). *Mathematical Models of Financial Derivatives*. Springer, Singapore, 2 edition.

Ross (1986). Economic forces and the stock market.

Windcliff, H., Forsyth, P., and Vetzal, K. (2001). Shout options: a framework for pricing contracts which can be modified by the investor. *Journal of Computational and Applied Mathematics*, 134:213–241.