Exam Project

MSc Data Science
July 2022

UniMiB - Financial Market Analytics

Cerabolini Aurora, Comensoli Paolo, Mirko Tritella

INTRODUCTION

Introduction

In this group work we want to understand what are the structural characteristics that risk brings to real investment portfolios.

For this reason, we need to build real portfolios that are focused on a specific level and type of risk.

To investigate the level of risk of individual stocks, we rely on the Single Index Model, proposed by William Sharpe in 1963, and then proceed to group the securities based on common characteristics in order to study them.

The equation of the Single Index Model links the returns of a single stock with the returns of the market index: $r_i = \alpha_i + \beta_i(R_M) + e_i$

Benchmark Index

- Our Index is the **Nasdaq-100** (**NDX**): it includes 100 of the largest national (US) and international non-financial companies listed on the Nasdaq stock market by market capitalization.

- The Nasdaq-100 undergoes *quarterly* reviews that occur in March, June, September and December and become effective with the closing values on the third Friday of the month.

- To become part of this index, stocks must have a daily trading volume of at least 200,000 pieces, have a total average market capitalization equal to or greater than 0.1% of the average market capitalization of the Nasdaq-100 stocks and typically must have been listed for at least a couple of years.

Nasdaq-100 Historical Closing Prices
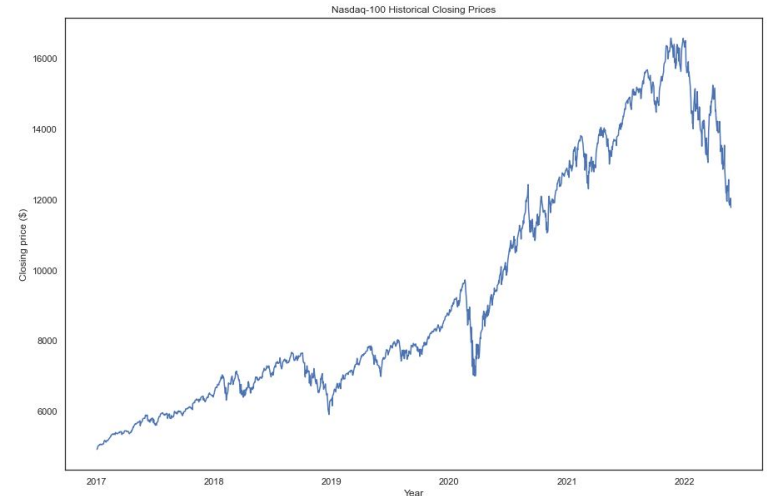
Dataset

For the construction of the dataset, we downloaded the daily closing prices of these securities and of all those securities that were part of the index starting from 2017.

We then managed the revisions of the securities by finding the incoming securities, the outgoing securities, and the effective date of exchange on the official Nasdaq website.

We used the Bloomberg terminal to download the data.

Survivorship Bias

When building the dataset we put attention to the survivorship bias:

- to create a dataset that takes into account the revisions that took place during the analysis period, we downloaded the historical series of daily prices even for those stocks that are no longer part of the index, but that are still within the U.S. stock market.

- By checking the date on which a certain change took place, indicated by X, between a security A (incoming) and B (outgoing), we set to #N/A all the cells of B starting from X, valued the cells of A starting from X and setting to #N/A all cells of A prior to X.

- In these 5 years that we are analyzing, it can happen that a stock was removed from the Nasdaq-100 and the US stock market. We also had to look up the prices of those stocks that are no longer listed. In this case we had to consider the sale price or what was obtained from a merge.

In this way we get 3 datasets:

- the dataset containing the closing prices of the stocks inside the index and the closing price of our benchmark index.

- the dataset with the closing prices of the securities removed from the index but still present in the U.S. stock market.

- for the stocks removed from both the Nasdaq-100 and the US stock market, we created a dataset with the price that the various shareholders received for each share

BUILDING PORTFOLIOS

Ranking & Selection

- With SIM we can decompose the returns of the stocks
- We can construct portfolios based on these components & properties
- We built 11 different portfolios with features like
  - Highest level or R-squared
  - Highest or lowest systematic risk
  - Highest or lowest idiosyncratic risk
  - Positive α_i and eventually significant
  - Absolute return and Excess return over total returns
- Steps:
  - Load 180-day window
  - Perform linear regression (versus NDX)
  - Extract the values
  - Select the best **10** stocks based on a selector

statsmodels
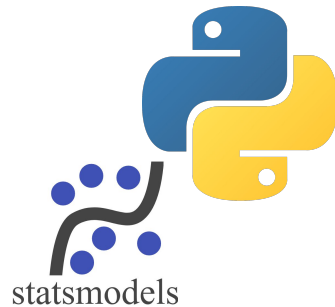
Results | Highest R-squared

- The higher the r-squared number, the more correlated the asset is to its benchmark
- Strong tech company presence

	Dates	Title	r2	specific_risk	beta	alpha	alpha_significance	absolute_returns	systematic_risk	total_risk
0	2/10/2017	FB	0.600861	0.000038	1.161801	0.000889	0.059219	0.387506	0.336740	0.336778
1	2/10/2017	GOOG	0.599868	0.000036	1.132423	-0.000289	0.528493	0.169747	0.319926	0.319962
2	2/10/2017	AMAT	0.587449	0.000097	1.801867	0.000219	0.769658	0.392228	0.809985	0.810081
3	2/10/2017	GOOGL	0.583270	0.000037	1.105663	-0.000318	0.492564	0.159295	0.304984	0.305021
4	2/10/2017	MSFT	0.575311	0.000029	0.958420	-0.000041	0.919458	0.180242	0.229162	0.229191
5	2/10/2017	ADBE	0.542516	0.000044	1.111706	0.000852	0.093291	0.370990	0.308327	0.308371
6	2/10/2017	AAPL	0.534622	0.000057	1.244978	0.000287	0.618008	0.295418	0.386683	0.386740
7	2/10/2017	TXN	0.509606	0.000055	1.166210	-0.000413	0.466200	0.153995	0.339301	0.339356
8	2/10/2017	AVGO	0.506455	0.000113	1.655574	-0.000084	0.917055	0.308999	0.683799	0.683912
9	2/10/2017	LRCX	0.504303	0.000124	1.723359	0.000782	0.356067	0.478151	0.740940	0.741063

NDX VS MAX R-SQUARED PORTFOLIO
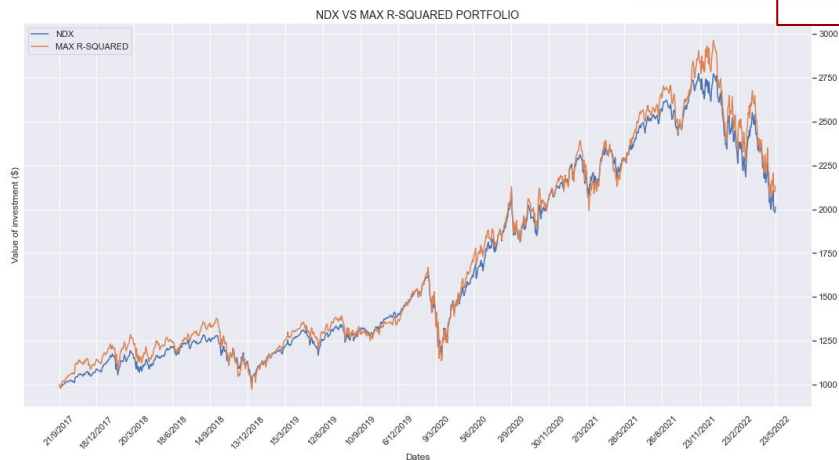
Results | Absolute return

	Dates	Title	r2	specific_risk	beta	alpha	alpha_significance	absolute_returns	systematic_risk	total_risk
0	2/10/2017	VRTX	0.153709	0.000553	1.540778	0.002283	0.203283	0.712586	0.592259	0.592812
1	2/10/2017	NVDA	0.241062	0.000547	2.026196	0.001128	0.526473	0.599815	1.024223	1.024770
2	2/10/2017	ATVI	0.242589	0.000272	1.434711	0.001624	0.196800	0.573247	0.513524	0.513796
3	2/10/2017	TSLA	0.145605	0.000441	1.331995	0.001574	0.325138	0.544164	0.442626	0.443067
4	2/10/2017	JD	0.333218	0.000223	1.621815	0.001252	0.271072	0.542991	0.656197	0.656420
5	2/10/2017	ISRG	0.200560	0.000096	0.755998	0.001893	0.012045	0.488818	0.142585	0.142681
6	2/10/2017	LRCX	0.504303	0.000124	1.723359	0.000782	0.356067	0.478151	0.740940	0.741063
7	2/10/2017	PYPL	0.346526	0.000095	1.092873	0.001452	0.052022	0.475269	0.297969	0.298064
8	2/10/2017	MU	0.360894	0.000272	1.906027	0.000521	0.678468	0.466950	0.906337	0.906609
9	2/10/2017	ILMN	0.102850	0.000358	0.984149	0.001360	0.345376	0.437460	0.241631	0.241989

- High beta
- Low r-squared
- Portfolio not top performing

	VRTX	NVDA	ATVI	TSLA	JD	ISRG	LRCX	PYPL	MU	ILMN
181	-0.004331	-0.027716	-0.013787	-0.020071	-0.033405	-0.015560	-0.004286	-0.001701	-0.001669	-0.014133
182	-0.005275	-0.009784	-0.003281	-0.042901	-0.014294	0.001266	0.025502	0.006939	0.004446	0.005826
183	-0.014383	-0.045722	-0.027929	-0.017527	-0.034577	-0.013837	-0.037825	-0.024577	-0.033835	-0.000149
184	-0.004302	0.005598	-0.006134	0.000753	-0.015016	-0.000690	-0.004299	-0.008858	-0.019986	-0.014707
185	0.004100	0.021687	0.021306	-0.012474	-0.015245	0.007295	0.057063	0.006177	0.081707	-0.003433
186	0.003683	-0.000285	0.007107	-0.004026	-0.027056	0.001756	0.013455	0.009430	0.023186	0.006251
187	0.016045	0.017436	0.015151	0.004407	-0.011711	0.019306	0.017555	0.001563	0.035455	0.001005

Results

	Dates	Title	r2	specific_risk	beta	alpha	alpha_significance	absolute_returns	systematic_risk	total_risk
0	2/10/2017	VRTX	0.153709	0.000553	1.540778	0.002283	0.203283	0.712586	0.592259	0.592812
1	2/10/2017	ISRG	0.200560	0.000096	0.755998	0.001893	0.012045	0.488818	0.142585	0.142681
2	2/10/2017	ATVI	0.242589	0.000272	1.434711	0.001624	0.196800	0.573247	0.513524	0.513796
3	2/10/2017	TSLA	0.145605	0.000441	1.331995	0.001574	0.325138	0.544164	0.442626	0.443067
4	2/10/2017	PYPL	0.346526	0.000095	1.092873	0.001452	0.052022	0.475269	0.297969	0.298064
5	2/10/2017	ILMN	0.102850	0.000358	0.984149	0.001360	0.345376	0.437460	0.241631	0.241989
6	2/10/2017	JD	0.333218	0.000223	1.621815	0.001252	0.271072	0.542991	0.656197	0.656420
7	2/10/2017	NVDA	0.241062	0.000547	2.026196	0.001128	0.526473	0.599815	1.024223	1.024770
8	2/10/2017	EA	0.259101	0.000159	1.144536	0.001081	0.260080	0.418724	0.326807	0.326965
9	2/10/2017	CERN	0.153864	0.000137	0.766338	0.001079	0.226305	0.344250	0.146512	0.146648

Positive alpha portfolio

Minimum total risk portfolio

- Not so many tech companies

	Dates	Title	r2	specific_risk	beta	alpha	alpha_significance	absolute_returns	systematic_risk	total_risk
0	2/10/2017	TSCO	0.001484	0.000277	0.098691	-0.001357	0.285068	-0.224957	0.002430	0.002707
1	2/10/2017	ORLY	0.001172	0.000458	-0.112696	-0.001673	0.305018	-0.323228	0.003168	0.003627
2	2/10/2017	COST	0.020284	0.000143	0.264130	-0.000180	0.843156	0.019345	0.017405	0.017547
3	2/10/2017	ROST	0.015743	0.000188	0.266445	-0.000702	0.501338	-0.074167	0.017711	0.017899
4	2/10/2017	DLTR	0.018114	0.000205	0.299133	0.000087	0.936316	0.074256	0.022323	0.022529
5	2/10/2017	VOD	0.043517	0.000090	0.311569	0.000468	0.518167	0.145182	0.024218	0.024308
6	2/10/2017	MNST	0.025836	0.000188	0.343631	0.000770	0.461535	0.205867	0.029459	0.029647
7	2/10/2017	PAYX	0.072053	0.000083	0.390339	-0.000662	0.340774	-0.042659	0.038012	0.038095
8	2/10/2017	ESRX	0.035739	0.000182	0.399220	-0.001077	0.295046	-0.115660	0.039761	0.039943
9	2/10/2017	DISCA	0.031266	0.000227	0.415813	-0.001974	0.086250	-0.273948	0.043135	0.043362

Portfolios Comparison

	Portfolio Title	Annualized Returns	Annualized Volatility
0	max_r2	16.26%	30.05%
1	absolute_returns	9.57%	30.53%
2	min_r2_and_high_specific_risk	6.14%	29.2%
3	max_specific_risk	-0.79%	35.25%
4	high_systematic_risk	8.63%	38.46%
5	low_systematic_risk	13.18%	19.25%
6	positive_alpha	5.15%	28.91%
7	positive_alpha_and_high_beta	15.83%	36.2%
8	max_total_risk	8.5%	38.48%
9	min_total_risk	12.95%	19.23%
10	excess_return_over_total	10.21%	25.65%
11	Nasdaq-100	15.01%	25.22%

NDX VS MAX SPECIFIC RISK PORTFOLIO
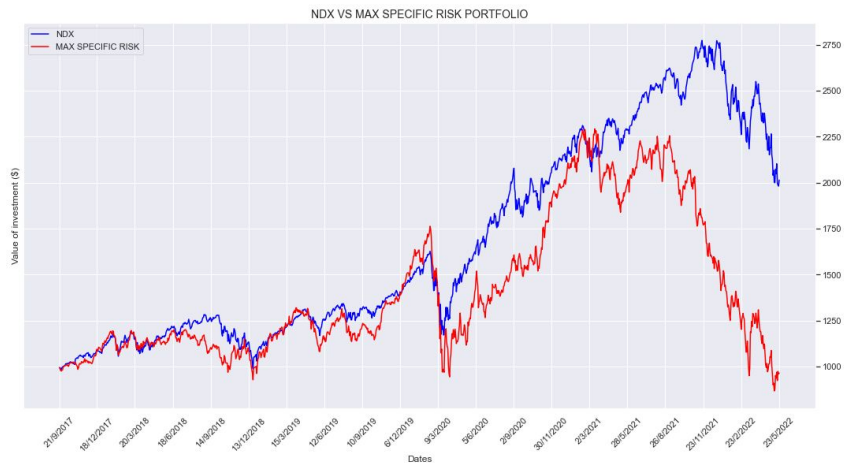
METRICS

Final Statistics

The metrics we take into account to compare the portfolios are:

- **Sharpe Ratio** to determine whether the risk you've taken on has paid off in your returns, compared to the returns you might have seen without taking on risk.

- **Maximum DrawDown** that expresses in percentage terms, and for a fixed period of time, the maximum loss of value that an investor can experience.

- **Calmar Ratio** to understand if the return of the investment is not at risk of significant drawdowns

We also calculate **Value at Risk**, **The Information Ratio** and **The Modigliani Ratio**.

The following table shows the values of the metric, calculated for each portfolio:

	Portfolio Title	Sharpe Ratio	MDD	CL	Var 90	Var 95	Var 99	IR	M2
0	max_r2	0.50644	-0.319009	0.509824	-0.0236161	-0.0304939	-0.0433954	0.147124	0.13814
1	absolute_returns	0.279216	-0.37224	0.257058	-0.0242679	-0.0312552	-0.0443621	-0.374972	0.0808451
2	min_r2_and_high_specific_risk	0.174381	-0.434591	0.141185	-0.023329	-0.0300114	-0.0425466	-0.425863	0.0544106
3	max_specific_risk	-0.0519942	-0.622202	-0.0126751	-0.0284863	-0.0365529	-0.0516845	-0.660405	-0.00267051
4	high_systematic_risk	0.197286	-0.433131	0.199285	-0.0307064	-0.0395083	-0.0560193	-0.325048	0.0601861
5	low_systematic_risk	0.630407	-0.245205	0.537489	-0.0150178	-0.0194234	-0.0276875	-0.0987685	0.169399
6	positive_alpha	0.142176	-0.378936	0.136007	-0.0231317	-0.0297472	-0.0421568	-0.735389	0.04629
7	positive_alpha_and_high_beta	0.408441	-0.398539	0.397201	-0.028597	-0.036882	-0.0524232	0.0519947	0.113429
8	max_total_risk	0.193723	-0.433131	0.196196	-0.0307253	-0.039531	-0.0560492	-0.331627	0.0592876
9	min_total_risk	0.619095	-0.245205	0.528119	-0.0150112	-0.0194124	-0.0276682	-0.111512	0.166546
10	excess_return_over_total	0.357313	-0.357034	0.285965	-0.0203041	-0.0261748	-0.0371875	-0.4256	0.100537

For example, we can see that:

- the portfolio that obtains the highest Sharpe Ratio value is the one obtained by taking into consideration the lowest Systematic Risk.

- the portfolio that has the highest value of MDD is the one built with the highest specific risk values and the maximum loss experienced by this portfolio was 62%.

MOMENTUM

Momentum Strategy

Momentum strategy is a system of buying stocks or other securities that have had high returns over the past months, and selling those that have had poor returns over the same period.

In our project, to build the portfolios we consider the titles with a positive momentum value.

To build portfolio taking into account also the momentum of each stock, we first select 1/3 of the stocks with the highest positive momentum value and then we proceed as before in the ranking of the stocks that are part of our portfolio based on selectors

For each portfolio we calculated the annualized return and annualized volatility.

```
+----+----------------------------+--------------------+----------------------+
|    | Portfolio Title            | Annualized Returns | Annualized Volatility |
+----+----------------------------+--------------------+----------------------+
|  0 | max_r2                     | 13.8%              | 30.67%               |
|  1 | absolute_returns           | 15.18%             | 30.19%               |
|  2 | min_r2_and_high_specific_risk | 7.15%           | 21.71%               |
|  3 | max_specific_risk          | 11.84%             | 29.63%               |
|  4 | high_systematic_risk       | 17.02%             | 34.98%               |
|  5 | low_systematic_risk        | 3.45%              | 19.62%               |
|  6 | positive_alpha             | 10.29%             | 28.43%               |
|  7 | positive_alpha_and_high_beta | 15.84%           | 34.31%               |
|  8 | max_total_risk             | 17.08%             | 34.99%               |
|  9 | min_total_risk             | 3.45%              | 19.62%               |
| 10 | excess_return_over_total   | 9.44%              | 27.57%               |
+----+----------------------------+--------------------+----------------------+
```

Depending on the type of selector, returns increased or decreased, more or less significantly, while variability remained very similar.

- the portfolio with the highest specific risk stocks went from negative to positive return.

- the portfolios with maximum total risk stocks and better absolute returns went from underperforming the index to overperforming it (the NDX annualized return was 15.01% and annualized volatility of 25.22%).

The following table shows the values of the metric, calculated for each portfolio:

	Portfolio Title	Sharpe Ratio	MDD	CL	Var 90	Var 95	Var 99	IR	M2
0	max_r2	0.415989	-0.324446	0.425407	-0.0242118	-0.0312308	-0.0443973	-0.112278	0.115333
1	absolute_returns	0.468196	-0.339194	0.447477	-0.0237689	-0.0306779	-0.0436378	0.0118272	0.128497
2	min_r2_and_high_specific_risk	0.28113	-0.260577	0.274246	-0.0172398	-0.0222074	-0.0315258	-0.519758	0.0813276
3	max_specific_risk	0.364244	-0.379188	0.312158	-0.0234508	-0.030232	-0.0429522	-0.221401	0.102285
4	high_systematic_risk	0.456785	-0.384213	0.4431	-0.0275676	-0.0355741	-0.0505931	0.133443	0.125619
5	low_systematic_risk	0.122868	-0.235603	0.146623	-0.015701	-0.0201908	-0.028613	-0.781894	0.0414215
6	positive_alpha	0.325353	-0.327148	0.314663	-0.022544	-0.0290507	-0.0412562	-0.353882	0.0924786
7	positive_alpha_and_high_beta	0.431187	-0.380265	0.416466	-0.0270677	-0.0349191	-0.0496471	0.0553922	0.119165
8	max_total_risk	0.458434	-0.384213	0.444641	-0.027568	-0.0355753	-0.0505957	0.137339	0.126035
9	min_total_risk	0.122868	-0.235603	0.146623	-0.015701	-0.0201908	-0.028613	-0.781894	0.0414215
10	excess_return_over_total	0.304666	-0.366548	0.257637	-0.0218825	-0.0281921	-0.0400279	-0.512999	0.0872623

For example, we can see that:

- the portfolios built with the lowest systematic risk has the lowest Sharpe Ratio value. Without the momentum calculation, this portfolio was the one with the highest value.

- considering the VaR with a confidence level of 90%, 95% and 99%, the portfolio at risk of having a greater loss than the others is the one built with the highest values of the total risk and it is the same obtained without the momentum calculation.

Thank you for your attention