FIN3080 Report 123090559

Data collection

As required, the data was downloaded ended up with 6 cvx files including these categories: Monthly Closing Price(1), Market Value of Tradable Shares(1), Monthly Return without Cash Dividend Reinvested(1), Return on Assets-B(2), Return on Equity - B(2), R&D expenses(3), Total Assets(4), Total Liabilities(4), Establishment Date(5), Market Type(5), Earnings per Share-TTM1(6), Net Assets per Share(6) during that time and A shares options were downloaded

Specifically, Return on assets/Equity - B were chosen because the average denominator(e.g., average total assets or equity over a period) to mitigate distortions from short-term volatility. This approach smooths out abrupt changes in balance sheet items, offering a more stable reflection of a company's long-term profitability.

Noted that there are many types of Earning per share. Choose Earnings per Share-Trailing Twelve Months (EPS-TTM) leveraging 12-month rolling data to reduce the impact of quarterly anomalies. TTM metrics dilute extreme values and provide a clearer view of sustainable earnings performance.

Establishment date and market type were downloaded all due to the requests in the pdf.

Problem 1

(a)

Coding to manipulate data process

1. Data cleaning and preparing:

Drop all the B shares columns (not necessary actually but just to make sure in case forgetting to click the A share option when downloading), drop all the unnecessary data.

Standardize columns. To better code in the following steps, set the titles into readable names like ”stock”, ”date”, “price” etc.

Data alignment and merging. Price Dates: Convert to monthly periods. EPS/book value dates are adjusted to the next month's start for accurate alignment. Merges on date columns. Datasets are merged based on these dates. Missing values are filled by carrying forward the last known value.

P/E and P/B ratios are calculated using three columns (closing price, earnings per share, and book value per share) to assess stock valuation.

R&D/Asset ratio Standardizes date formats in R&D expense and asset/liability datasets to quarterly intervals . Integrates the datasets using common identifiers. Computes the R&D-to-Asset ratio by dividing quarterly R&D expenses by corresponding total asset values for each merged record. This metric quantifies innovation investment intensity relative to a firm’s asset base.

Firm ages. Quarter-end dates are converted, use function to change quarter indicators to the last day of the specific quarter, and gets the firm age by comparing this date to the establishment date. The age is expressed in years.

Results shown in the file.

(b)

Summary: shown in the table as rounded to 2 decimals and firm age in year.

Table 1

Main board

	Monthly return	P/E	P/B	ROA	ROE	R&D/asset	Firm age
count	561774	546803	555717	318351	313611	103566	339255
mean	0.01	39.04	2.31	0.02	0.02	0.01	17.47
std	0.16	6218.11	677.81	0.24	1.06	0.02	7.32
min	-0.89	-1288000	-282500	-64.82	-276.27	0	-0.6
25%	-0.07	12.71	1.6	0	0.01	0	12.16
50%	0	28.27	2.57	0.01	0.02	0.01	17.23
75%	0.08	56.93	4.21	0.03	0.05	0.02	22.29
max	22.05	1279000	28171.14	64.75	36.49	0.91	73.05

Table 2

GEM *board*

	Monthly return	P/E	P/B	ROA	ROE	R&D/asset	Firm age
count	113574.0	107601.0	113631	69807	69637	46841	71612.0
mean	0.01	95.25	4.75	0.02	0.0	0.02	16.67
std	0.18	10627.14	18.23	0.05	6.66	0.03	5.45
min	-0.86	-692857.14	-733.51	-1.88	-1756.04	0.0	0.69
25%	-0.08	24.4	2.38	0	0.01	0.01	12.78
50%	0	41.96	3.53	0.01	0.02	0.02	16.44
75%	0.08	72.03	5.55	0.03	0.05	0.03	20.14
max	6.4	2087500.0	3484.12	0.97	1.32	2.7	42.57

Overview:

2. The Main Board has a substantially larger observation compared to the GEM Board, indicating a broader spectrum of listed entities.

3. **P/E Ratio**: The GEM Board tends to display higher. It is likely to be more companies that are undervalued in the Main Board. **P/B Ratio**: Analogously, the GEM Board also shows higher than P/B ratio statistics. However, its maximum and standard deviation are less than the Main board.

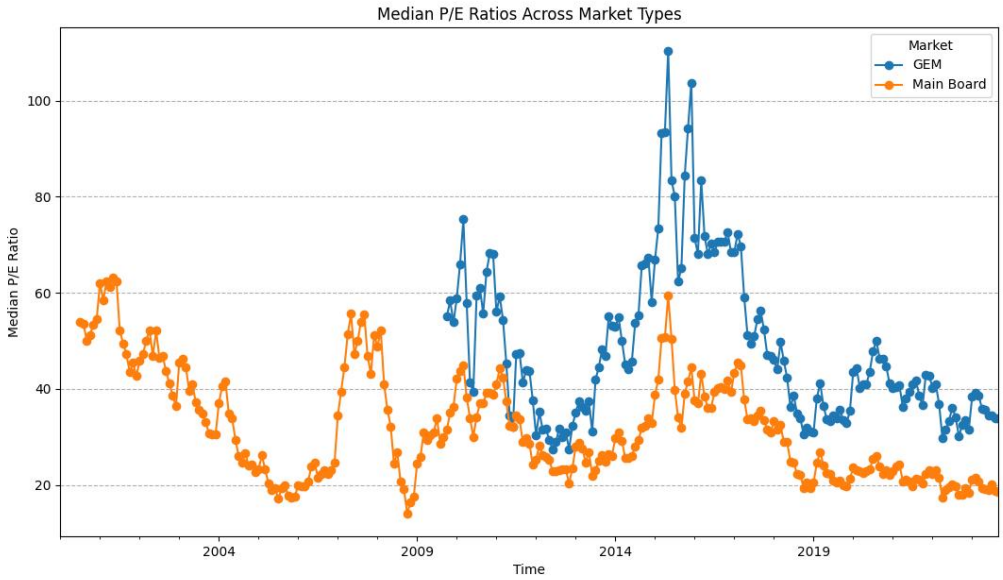
4. Regarding ROA and ROE columns,apart from outlier values,both boards share comparable statistics. The GEM Board demonstrates a considerably higher ratio of R&D expenses to total assets, highlighting a greater focus on innovation.

5. Firm Age: In terms of quarterly firm age, the Main Board is larger than the GEM Board across all statistical measures. Summarizing, this comparison indicates that firms on the main board are generally more stable, whereas those on the GEM board present higher risks and returns.

Problem 2

Figure 1

Two time-series for median P/E ratio by market type Average



(i) Yes. At sep.2023, it is advisable to invest newly into both markets. As looking back into the trend, it is indicated that sep.2023 is in a low point, signifying a possible potential undervaluation of listed equities. Furthermore, historical data reveals cyclical market patterns, implying a likelihood of future capital appreciation. So it is advisable.

(ii) As the cyclical PE ratios, experiencing highs and lows, the possible strategy is: **Buying or going long on index ETFs during periods of low P/E ratios periods and selling or shorting them during high P/E periods.** The characteristics of two boards are: main board has lower ratios, while more stable than GEM board. So it is advisable to allocate portfolios according to risk reference.

Problem 3

Table 3

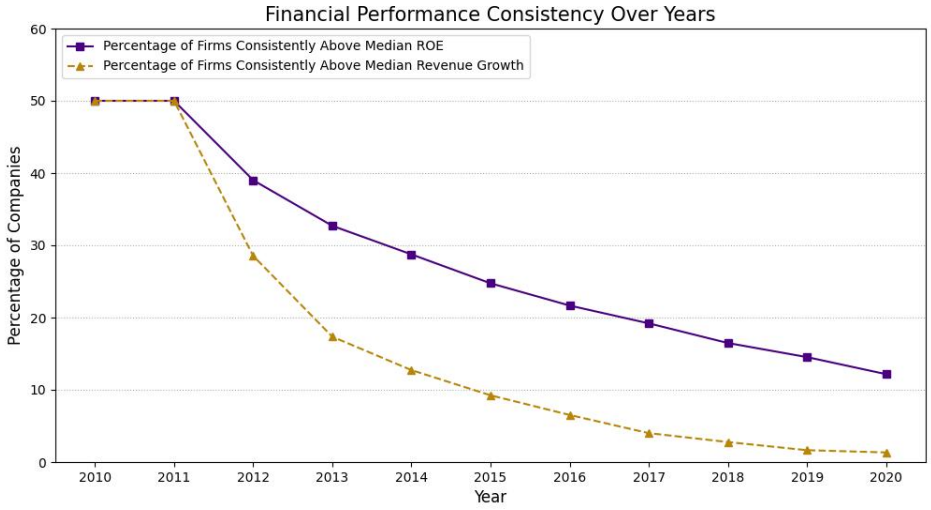
The Annual Median for ROE and Total Revenue Growth Rate(2 decimals in the table)

year	ROE median	Revenue growth median
2011	0.0912	0.1613
2012	0.0731	0.0595
2013	0.0684	0.1057

2014	0.0701	0.0753
2015	0.0689	0.0313
2016	0.0753	0.0991
2017	0.0819	0.16
2018	0.0701	0.1079
2019	0.0718	0.0664
2020	0.0781	0.039

Figure 2

Time-series of the percentages of companies that consistently maintain above-median ROE and total revenue growth rate



Overview:

Even though the benchmark was just the "median level" (not very high), very few companies stayed above the median consistently. Less than 10% of them outperformed the market over the decade. This proves how hard it is to beat the market for a long time, even with a simple starting rule.