Quantitative Finance Methods for ESG Stocks in Asia

Introduction to Quantitative Finance

Tom J. Espel

University of Hong Kong

Thursday 19th September, 2024

Objectives

Figure 1: In Q4'2023, AI was mentioned in the earnings calls of 179 of the S&P 500's companies.
Bob Rich's *The Matrix* for Hedgeye on May 29th, 2024.

Outline

Quantitative Finance

Definition (Quantitative Finance)

Quantitative Finance is the branch of Mathematics dedicated to studying financial markets.

Quantitative Finance is usually associated with systematic investing in the secondary markets. Its practitioners are called **quantitative analysts** (a.k.a quants or strats).

Quantitative Finance

- Quantitative Finance originated in Europe in the early 20th century, alongside the study of stochastic processes. One of its early founders is **Louis Bachelier**.
- It grew significantly in the USA in the 1970s with the growth of derivatives contracts and some landmark contributions in option pricing by **Black, Scholes, and Merton**.
- In 1982, **Jim Simons** established Renaissance Technologies, the first and most successful quantitative hedge fund.

Quantitative Finance

- The **Masters in Computational Finance** at Carnegie Mellon University was created in 1994 by Shreeve and is regarded as the first modern-era program in Quantitative Finance.
- Quants became increasingly important after the 2001 dot com bubble, then after the 2008 GFC for risk modeling.
- As more data became available and high-frequency venues grew, some moved to the front office and portfolio management roles.

Quantitative Finance

- The interest of Quantitative Finance is to create models that explain and predict asset prices in financial markets. **Economics** and **Game Theory** were its main contributors..
- It is an active and growing field of research with both public and private institutions. Recently, developments in **Computer Science** have been increasingly influential.
- Its historical research and practitioner hubs are **NyLonKong**, alongside Chicago, Paris, Singapore, and Tokyo.

Models

Asset prices are usually studied in the form of their returns[1].

Definition (Simple Return)

Let P_t be the price of an asset at time t, $t > 1$. The return R_t is defined as

$$R_t = \frac{P_t}{P_{t-1}} - 1$$

Models

The properties of financial returns have been extensively studied[1, 2]. Some of the most important empirical findings are called **stylized facts**.

- Historical daily returns tend to follow a normal distribution over large samples.
- Very short-horizon returns tend to follow a geometric distribution.
- The absolute or squared returns are highly auto-correlated (volatility clustering).

Models

Volatility is a fundamental concept in Finance. It expresses the uncertainty of a security's price over a set horizon. The most common mathematical definition relies on returns.

Definition (Volatility)

Volatility $\sigma(T;t)$ at time t over horizon T is the standard deviation of the returns observed in the interval $[t-T;t]$.

- We usually just mark volatility as $\sigma(t)$.
- Because of its definition as a standard deviation, there are implications on calculus (product, sum).
- By convention, it is annualized and expressed in percent.

Theories

The **Modern Portfolio Theory** (MPT) is a dominant framework in Finance.
Markowitz introduced it in 1952[3] and established the mean-variance optimization framework.

- Investors are risk averse (the variance) and want to maximize returns (the mean)
- The most common formulation is when an investor picks a combination of a risk-free asset and a market portfolio

Theories

We can represent the possible combinations on a graph with variance and expected returns[4].

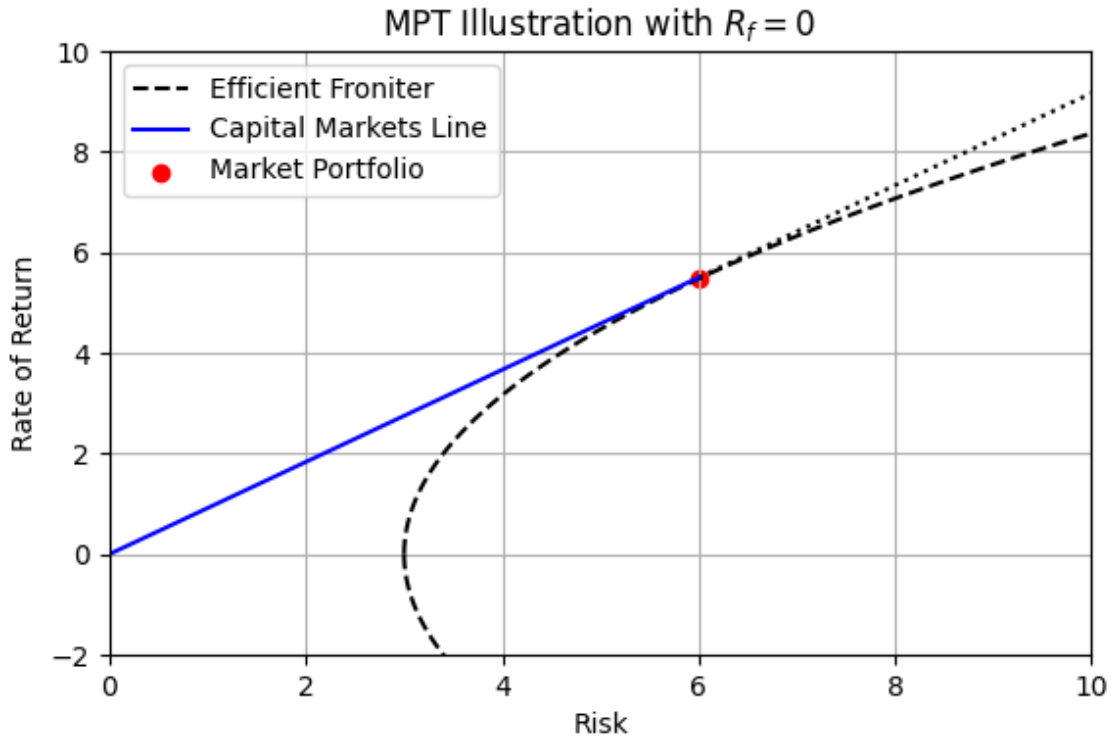

Figure 2: Illustration of MPT key concepts.

Risk and Quantitative Finance

The balance between risk and return can be expressed Mathematically. The most popular of such representations is the Sharpe Ratio, developed by Sharpe in 1966[5, 6].

Definition (Sharpe Ratio)

Given the returns R of an asset and the returns R_f of the risk-free asset, we define the Sharpe Ratio.

$$SR = \frac{\mathbf{E}(R - R_f)}{\sqrt{\mathbf{VAR}(R - R_f)}}$$

There are other popular ratios, such as Treynor, Sortino, and Sterling.

Data

Quantitative Finance relies on various sources of data.

- Market prices, in particular historical data.
- Public information (quantified): economic data releases, earnings, etc.
- Public information (non-quantified): transcripts, news headlines, etc.
- Increasingly, synthetic data is also used (so-called market generators[7]).

Data

Data can be inaccurate for multiple reasons.

- Down-sampling for market data: loss of resolution.
- Data and reporting errors and omissions.
- Earnings management and other fraudulent practices.

Even when data is accurate, the market is unaware of non-public material information (MNPI).

High Frequency Trading

- High-frequency trading is one form of systematic, quantitative trading that involves exchanging securities and other financial assets quickly.
- It is often computerized and limited to liquid, electronic markets: equities and foreign exchange.
- Both market-makers and market-takers are involved in high-frequency trading.

High Frequency Trading

The case **for**.

+ It makes markets efficient at medium and low frequencies by resolving arbitrages at high-frequency.
+ High-frequency trading has also helped provide more liquidity and, in the long run, has pushed transaction costs down.

The case **against**.

- It is speculative and can lead to over-reactions, including flash crashes (May 6th, 2010, flash crash of the S&P500).
- High barriers to entry and *technological arms race* make it difficult for new participants to compete.

Capital Asset Pricing Model

The intuition behind the **Capital Asset Pricing Model** (CAPM) is that one can decompose the price of a security as follows[2].

- **systematic** component: market, cannot be hedged
- **idiosyncratic** component: specifics of the individual asset

Both components are associated with returns and risks. The CAPM's main contribution is its risk decomposition.

Capital Asset Pricing Model

The CAPM was highly impactful because it relates a portfolio's risk to non-diversifiable risk (market excess returns). It is also prevalent because of its simplicity.

Remarks

- Sharpe was awarded the Nobel Prize in 1990 alongside Markowitz and Miller.
- Sharpe initially submitted the CAPM in 1962 but was rejected as irrelevant. It was only and accepted only in 1964.

Capital Asset Pricing Model

Proposition (Capital Asset Pricing Model)

Let R the returns of an asset, R_M the returns of the market portfolio and R_f the returns of the risk-free asset. There exists a β that verifies the below.

$$\mathbf{E}R(t) - R_f = \beta\left(\mathbf{E}R_M(t) - R_f\right)$$

Single Factor Models

- The β in the CAPM is the expression of the sensitivity of the excess returns of the portfolio, so the excess returns of the underlying market.
- The β in the CAPM is the sole factor computed for the unique variable, so we refer to the CAPM as a single-factor model.

The CAPM can be expressed as an ex-post model by calibrating the observed returns.

$$\mathbf{E}R(t) - R_f = \beta\left(\mathbf{E}R_M(t) - R_f\right) + \epsilon(t)$$

We must introduce a term, $\epsilon(.)$, which captures the model errors. Under CAPM, we expect the values of $\epsilon(.)$ to be independent and identically (usually normally) distributed.

Regression

- Ex-post single-factor models are often calibrated using regressions.
- Regressions allow us to find the parameters between a dependent variable and a set of independent variables.
- The most common form is the **linear regression**, where the mathematical dependency between the independent and dependent variables is assumed to be linear (sum).

Regression

Definition (Linear Regression)

Let **X** be the matrix of p observations of n independent variables, and **y** be the column vector of the p observations of the dependent variable. We can write

$$\mathbf{y} = \mathbf{X}\beta + \epsilon$$

with β a column vector of dimension (n+1), and ϵ the column vector of the p residuals.

Regression

The parameter β of the regression can be estimated using different techniques.

- The choice of technique is determined in part by the assumptions one can make on β.
- Some constraints on $\mathbf{X}$ can also limit the choice of techniques available.
- We almost always assume some form of stationarity of the distributions.

The most common and popular estimator is called **Ordinary Least Squares**.

Regression : Ordinary Least Squares

Ordinary Least Squares (OLS) requires several assumptions. The main ones are contact variance and independence of errors. We also must allow $\mathbf{X}^T\mathbf{X}$ to have an inverse.
The main idea of OLS is to choose the $\hat{\beta}$ that reduces the sum of the squared distances between the dependent variable and the linear function of independent variables. It is called the **Sum of Squared Residuals** (SSR).

$$\hat{\beta} = \arg\min_{\beta} ||\mathbf{y} - \mathbf{X}\beta||^2$$

Regression : Ordinary Least Squares

We can show that $\hat{\beta}$ has a closed-form solution.

Proposition (OLS Estimation of β)

$$\hat{\beta} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$$

We can compute the estimated $\hat{\mathbf{y}}$ using $\hat{\beta}$.

Regression : Ordinary Least Squares

The ex-post CAPM is a **simple linear regression** because it only has one explanatory variable. We can, therefore, write the beta in closed form.

Proposition (Simple Linear Regression β)

$$\hat{\beta} = \frac{\mathbf{COV}(x, y)}{\mathbf{VAR}x} = \rho_{x,y} \frac{\sigma_y}{\sigma_x}$$

with $\rho_{x,y}$ the correlation coefficient between x and y and σ the standard deviation.

Regression : Ordinary Least Squares

Proposition (Coefficient of Determination)

The coefficient of determination r^2 is the ratio of the variance derived from the estimated $\hat{\mathbf{y}}$ over the measured variance of $\mathbf{y}$.

$$r^2 = \frac{\frac{1}{n}\sum_{k=1}^{n}(\hat{y}_k - \bar{y})^2}{\frac{1}{n}\sum_{k=1}^{n}(y_k - \bar{y})^2}$$

r^2 is often used to assess **goodness of fit**.

Regression

Two other popular estimations introduce a penalty coefficient λ, providing **regularization**.

- Ridge Regression uses $\mathcal{L}_2$ regularization
- LASSO (Least Absolute Shrinkage and Selection Operator) Regression uses $\mathcal{L}_1$ regularization

The only difference between Ridge and LASSO regression is the form of the regularization.

Recall the ex-post CAPM.

$$\mathbf{E}R(t) - R_f = \beta\left(\mathbf{E}R_M(t) - R_f\right) + \epsilon(t)$$

When we compute a regression, we can obtain a non-zero intercept.

$$\mathbf{E}R(t) - R_f = \left(\alpha + \epsilon(t)\right) + \beta\left(\mathbf{E}R_M(t) - R_f\right)$$

Jensen's Alpha

When we assume $\mathbf{E}\epsilon = 0$, we can formulate a CAPM-style model.

Definition (Jensen's Alpha)

Let R be the return of a portfolio, R_M be the market portfolio, and R_f be the return of the risk-free asset. There exists a systematic risk β such that.

$$R(t) - R_f = \alpha + \beta\left(R_M(t) - R_f\right)$$

and α is called the **Jensen alpha**.

It was formalized in 1967 by Jensen[8] and is used to evaluate a portfolio manager's **skill**.

What is ESG?

Definition (ESG)

ESG stands for

- E nvironmental
- S ocial
- G overnance

and is an investing principle, with its current formulation attributed to the UN. It relates to impact investing.

Application

Figure 3: A taste of what's to come: we will study the Meiji corporation, the makers of the chocolate brand.

Application

Can we describe the price returns of Asian stocks with ESG indices under the CAPM framework?

What is ESG?

- Analysts can assign companies **ESG scores**. The leading scoring companies are MSCI and Morningstar.
- Assigning a score is complex, and such quantifications have been challenged. It remains a very active field of research.
- Based on those scores, some stocks are included in **ESG indices**. This creates another form of securities classification, which is cross-sectional compared with sectors.

Indices - Construction

- Indices are weighted price averages, with weights set by private providers.
- The weights are periodically rebalanced (monthly, quarterly), and each index has a different methodology.
- Some indices are regarded as broad enough to constitute a reference: S&P500 (USA), HSI (Hong Kong), Nikkei225 (Japan).

When indices are broad enough, they are used as benchmarks for investment performance. Products that replicate the indices (such as index ETFs) are said **pure** β.

Indices - Construction

Indices can have different focus areas.

- Geography: where the company is listed.
- Product: stocks, currencies (DXY), commodities.
- Sector: industry or type of asset as underlying.
- Other factors: such as ESG.

Because it's a different analysis from the classic industry-sector segmentation, we will examine ESG indices and their stocks.

ESG Indices

We aim to illustrate the concepts of α and β in the context of market fragmentation.

Definition (Market Fragmentation)

Market fragmentation is the state of a market when it is broken into smaller markets due to asset or structural characteristics.

Examples of fragmentation include geography (multi-venue assets), standardization (contract structure), and regulation.

ESG Indices

ESG indices are often cross-sectional, using a similar investment universe as other well-known conventional indices. It also highlights how the choice of a benchmark can impact performance analysis.

Scope	**Benchmark**	**ESG index**
Hong Kong	HSI	HSI ESG
Japan	Nikkei 225	FTSE Blossom Japan Index

Table 1: Selection of market indices in Hong Kong and Japan.

HSI ESG index

We review the characteristics of the HSI ESG index, which we will refer to as *HK ESG*.

- It uses the same underlying as the HSI index.
- The weights are adjusted with tilt factors based on the sustainability scores the Hong Kong Quality Assurance Agency provided.
- No component can have a contribution greater than 8%.
- Rebalancing is quarterly.

FTSE Blossom Japan Index

We will refer to the FTSE Blossom Japan Index as *Japan ESG*.

- Based on the FTSE ESG rating and comprising 310 Japanese businesses with overlap with the Nikkei225.
- The FTSE Japan All Cap index weights are adjusted with the FTSE ESG score. Only the companies that meet a minimum ESG score are included.
- Rebalancing is semi-annual.

Study

We chose dates between the rebalancing periods to simplify the analysis and reduce bias. Our analysis will be between June 21st and August 29th, 2024.

Date	# FTFBLSMJPN	# HSIESG	# ^HSI	# ^N225
	Missing: 2 (4%) Distinct: 48 (96%) Min 181.29 Max 240.75	Missing: 1 (2%) Distinct: 49 (98%) Min 2059.26 Max 2242.58	Missing: 2 (4%) Distinct: 48 (96%) Min 16647.33984... Max 18293.380859...	Missing: 3 (6%) Distinct: 47 (94%) Min 31458.419921... Max 42224.01953...
2024-06-21T00:00:00.000	222.49	2204.32	18028.51953125	38596.46875
2024-06-24T00:00:00.000	223.79	2203.14	18027.7109375	38804.6484375
2024-06-25T00:00:00.000	228.18	2203.1	18072.900390625	39173.1484375
2024-06-26T00:00:00.000	229.45	2207.04	18089.9296875	39667.0703125
2024-06-27T00:00:00.000	228.66	2166.35	17716.470703125	39341.5390625
2024-06-28T00:00:00.000	230.24	2169.76	17718.609375	39583.078125
2024-07-01T00:00:00.000	231.47	*Missing value*	*Missing value*	39631.05859375
2024-07-02T00:00:00.000	234.46	2169.7	17769.140625	40074.69140625
2024-07-03T00:00:00.000	235.76	2193.28	17978.5703125	40580.76171875
2024-07-04T00:00:00.000	238.35	2197.11	18028.279296875	40913.6484375
2024-07-05T00:00:00.000	237.38	2169.05	17799.609375	40912.37109375
2024-07-08T00:00:00.000	235.84	2138.34	17524.060546875	40780.69921875
2024-07-09T00:00:00.000	238.13	2142.6	17523.23046875	41580.171875
2024-07-10T00:00:00.000	239.31	2138.73	17471.669921875	41831.98828125

Figure 4: The data has gaps and needs a bit of processing.

Study

It's always a good practice to do some quick statistics to get to know our dataset.
We want to check some assumptions on distribution and potential trends.

Table 2: Statistical analysis of returns.

	HSI	**HK ESG**	**Nikkei225**	**Japan ESG**
Median	0.00	0.00022	0.0020	0.0044
Mean	-0.00021	-0.000035	0.00027	0.00018
Kurtosis	-0.000212	-0.000035	0.000272	0.000179
KS p-Val	1.87e-11	1.78e-11	3.13e-10	2.64e-10

Study

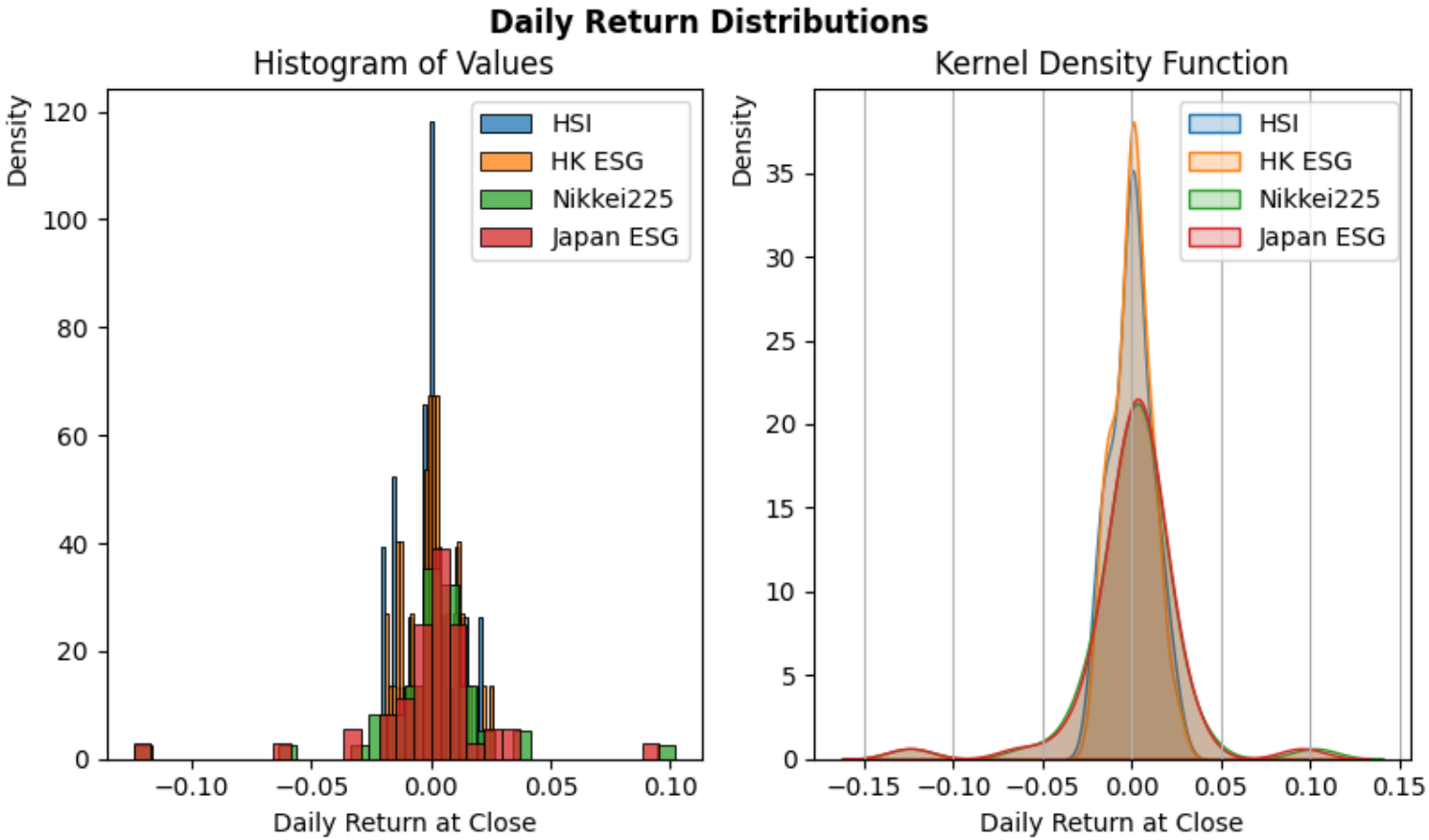

Figure 5: Plot of the daily close return distributions.

Study

Market	Company	Ticker
JP	Meiji	2269.T
	Toyota	7203.T
	Sony	6758.T
	MUFJ	8306.T
	NTT	9432.T
	Hitachi	6501.T
	SMBC	8316.T
	Nintendo	7974.T
	JAL	9201.T
	Kikkoman	2801.T

Table 3: List of Japanese companies and their ticker symbol.

Study

We selected a few Hong Kong companies, too.

Market	Company	Ticker
HK	CNOOC	0883.HK
	CLP	0002.HK
	SHK	0016.HK
	Swire	1972.HK
	SinoPec	0386.HK
	Tencent	0700.HK
	CMobile	0941.HK
	Sands	1928.HK
	AIA	1299.HK
	Alibaba	9988.HK

Table 4: List of HK Companies and their ticker symbol.

Study

- In the CAPM, we compare excess versus risk-free rate: to simplify our analysis, we assume $R_f = 0$.
- We'll start with OLS and assess goodness of fit with r^2.

Study - Meiji

Our OLS regression against the Japan ESG index gives us $\beta = 0.16$ and $r^2 = 0.10$.

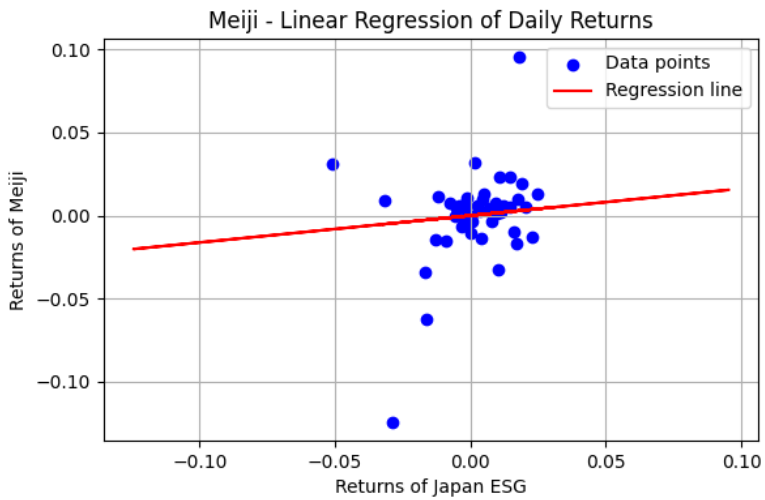

Figure 6: Regression graph for the daily returns of Meiji against the Japan ESG index.

Study - Meiji

Meiji has low β with both ESG and non-ESG indices.

Variable	β	r^2
Japan ESG	0.16	0.10
Nikkei225	0.15	0.09

Table 5: Regression statistics for Meiji with OLS.

Study

We can compare the regression beta results for both ESG and conventional indices.

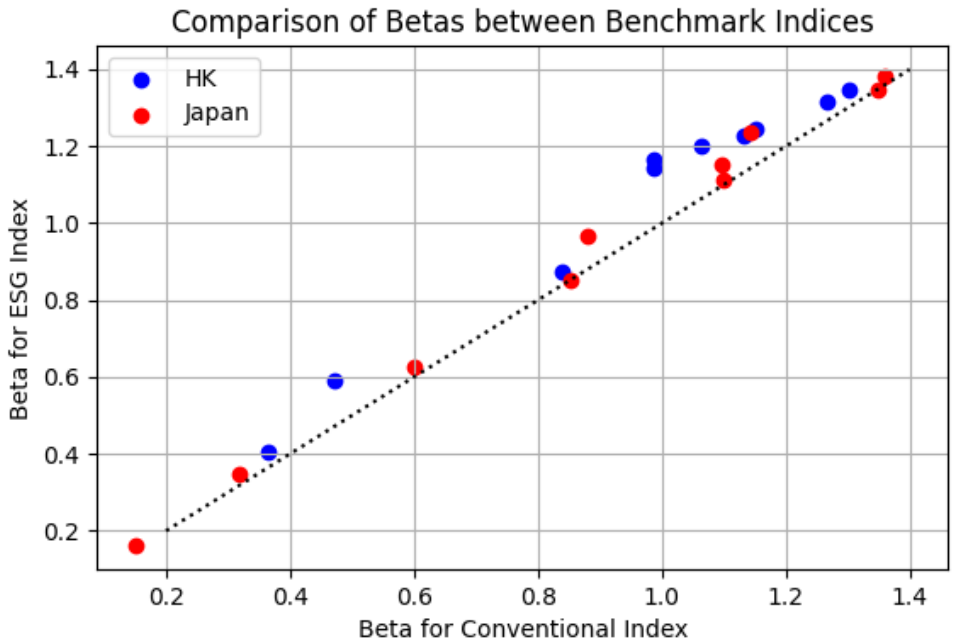

Figure 7: The range of r^2 in Hong Kong is 0.21-0.71 and for Japan 0.09-0.76 .

Discussion

- Meiji turned out to be an **outlier with low beta** compared with peers.
- Both ESG and conventional indices have **similar betas**, but for most companies selected, beta was higher on the ESG index.
- We observe low r^2, and that model would need to be refined.

Discussion

Can we conclude on the market fragmentation?

Multi-factor Models

- We can attempt models with more independent variables to improve fit.
- The explanatory variables can be other indices or other time series.

Based on our study, does creating a model with both conventional and ESG indices as variables make sense?

Regression: Beyond Ordinary Least Squares

In our examples, creating a model with both indices in the same region as variables would probably not make sense.

- No new information (financial interpretation): there is strong or full overlap between the underlying.
- Strong correlation (mathematical interpretation): OLS might perform poorly.

Table 6: Correlation matrix of the daily returns of indices.

	HSI	**HK ESG**	**Nikkei225**	**Japan ESG**
HSI	1.	-	-	-
HK ESG	0.99	1.	-	-
Nikkei225	0.30	0.27	1.	-
Japan ESG	0.33	0.30	0.98	1.

Regression: Beyond Ordinary Least Squares

When we are unsure about the overlap between our explanatory variables, we would need a tool that helps to pick one of the two indices for us.
This is where regularization can help, either Ridge or LASSO.

Characteristics

Another methodology for factor models, particularly in equities markets, uses characteristics.

Characteristics are other quantified parameters that characterize the stock, other than its returns. Some common characteristics used are:

- Intraday volume
- Market capitalisation
- Competitor's intraday volume

Characteristics

- Characteristics and other price factors are aggregated within one multi-factor model to predict stock returns.
- The most common form of regression practitioners use is Ridge or LASSO because its **regularization simplifies the model** and increases the explainability of the dependent variables.
- The backtesting and validation sample sizes and recalibration frequency are up to the trader's appreciation, and there is no firm rule[9].

Conclusion

- We have explored some important theories in Quantitative Finance, such as the **Capital Asset Pricing Model** (CAPM).
- We used **OLS regression** and computed the **CAPM betas** of some stocks against both ESG and conventional indices in Hong Kong and Japan.
- We have discussed potential **improvements** to our model, such as multi-factor models and other forms of regression.

Bibliography

R. S. Tsay, *Analysis of financial time series*.
Wiley series in probability and statistics, Cambridge, Mass: Wiley, 3rd ed ed., 2010.

C.-A. Lehalle and S. Laruelle, *Market microstructure in practice*.
Singapore: World Scientific, second edition ed., 2018.

H. Markowitz, "Portfolio Selection," *The Journal of Finance*, vol. 7, p. 77, Mar. 1952.

R. Grinold and R. N. Kahn, *Advances in active portfolio management: new developments in quantitative investing*.
New York: McGraw Hill, 2019.

W. F. Sharpe, "Mutual Fund Performance," *The Journal of Business*, vol. 39, p. 119, Jan. 1966.

W. F. Sharpe, "The Sharpe Ratio," 1994.

A. Kondratyev and C. Schwarz, "The Market Generator," *SSRN Electronic Journal*, 2019.

M. C. Jensen, "THE PERFORMANCE OF MUTUAL FUNDS IN THE PERIOD 1945–1964," *The Journal of Finance*, vol. 23, pp. 389–416, May 1968.

M. Isichenko, *Quantitative Portfolio Management: The art and science of statistical arbitrage*.
Wiley finance series, Hoboken, N.J: John Wiley & Sons, 1st ed., 2021.

AI Assistance Disclaimer

This presentation was partially created with the assistance of an AI chatbot named Perplexity (`www.perplexity.ai`). The AI was used for the following purposes:

- Reformatting data into LaTeX tables and csv files
- Generating Python code and LaTeX code based on prompts
- Searching for and providing details on stock ticker symbols

slides

handout

Thank you! :-)

research@tomespel.com