The Efficient Market Hypothesis and Related Tests

The cc return $r_t = \underbrace{\log P_t}_{p_t} - \underbrace{\log P_{t-1}}_{p_{t-1}}$.

Recall that

$$r_t = \underbrace{E\left[r_t \mid \mathcal{F}_{t-1}\right]}_{\mu_t} + \underbrace{r_t - E\left[r_t \mid \mathcal{F}_{t-1}\right]}_{\epsilon_t}$$

It can be shown that ϵ_t is a MDS.

The EMH implies that

1. $E\left[p_t \mid \mathcal{F}_{t-1}\right] = p_{t-1}$, so $\mu_t = 0$, so $r_t = \epsilon_t$. **CC return r_t is a MDS.**

2. $p_t - p_{t-1} = r_t$ is a MDS, so **log price p_t is a random walk process**.

If we further assume that $E[\epsilon_t^2] = \sigma^2$, then 1. and 2. above further imply:

a. **CC return r_t is $WN(0, \sigma^2)$.**

b. **log price p_t is a random walk process.**

This means we can test EMH by testing a and b.

Testing for White Noise

Let ρ_j denote the auto correlation coefficient of r_t. We can test the following null hypothesis:

$$H_0 : \sum_{j=1}^{m} \rho_j^2 = 0$$

We use the Ljung-Box statistic as test statistics:

$$Q_m = (T+2) \sum_{j=1}^{m} (1 - \frac{j}{T}) \hat{\rho}_j^2$$

where $\hat{\rho}_j$ is the sample auto correlation coefficient $\hat{\rho}_j = \frac{\hat{\gamma}_j}{\hat{\gamma}_0}$ for $|j| < T$. Q_m is approximately $\chi^2(m)$ under H_0. This implies we can use $\chi^2_{1-\alpha}(m)$ as the critical value.

The Python code to perform Ljung-Box test is as follows:

In [102…

```
# pip install pandas_datareader
import pandas_datareader as web
```

```python
import numpy as np
import pandas as pd

df = web.get_data_yahoo("AMZN", start = "2001-01-01", end = "2021-12-1", interva
df.reset_index(inplace = True)
df['Date'] = pd.to_datetime(df['Date']) # convert the strings to dates

df['CC'] = np.log(df['Adj Close'] / df['Adj Close'].shift(1))
CC = df['CC'][1:]

from statsmodels.stats.diagnostic import acorr_ljungbox
print("m = 2\n========================\n", acorr_ljungbox(CC, lags=[2]))
```

```
m = 2
========================
     lb_stat  lb_pvalue
2  2.781885   0.248841
```

The test with $m = 2$ has p-value larger than 10%, so we cannot reject the null hypothesis that cc return of GS is a WN process.

However, it's kind of important to include larger values of m in the test.

In [103…

```python
print("m = 5\n========================\n", acorr_ljungbox(CC, lags=[5]))

print("m = 10\n========================\n", acorr_ljungbox(CC, lags=[10]))

print("m = 20\n========================\n", acorr_ljungbox(CC, lags=[20]))
```

```
m = 5
========================
    lb_stat  lb_pvalue
5  7.37175   0.194428
m = 10
========================
       lb_stat  lb_pvalue
10  13.887108   0.178204
m = 20
========================
       lb_stat  lb_pvalue
20  24.623522   0.216215
```

Testing for Random Walk

The EMH implies that $E[r_t|\mathcal{F}_{t-1}] = 0$, or $r_t = p_t - p_{t-1} = \epsilon_t$. We can impose an AR(1) model for the log prices.

$$p_t = \alpha + \beta p_{t-1} + \epsilon_t.$$

To test for random walk is equivalent to testing $H_0 : \beta = 1$. We can consider the Augmented Dickey-Fuller test. Under H_0,

$$ADF = \frac{\hat{\beta} - 1}{s(\hat{\beta})} \to_d \frac{\int_0^1 U(r)dW(r)}{\left(\int_0^1 U(r)^2 dr\right)^{1/2}}$$

In [104…

```python
from statsmodels.tsa.stattools import adfuller

# augmented Dickey-Fuller test
adfuller(np.log(df['Adj Close']), regression = 'c')
# The test statistic, approximate p-value, the number of lags used.,
# The number of observations used for the ADF regression
# Critical values for the test statistic at the 1 %, 5 %, and 10 % levels
# The maximized information criterion if autolag is not None
```

Out[104…

```
(-0.20540301411430542,
 0.9378745206897964,
 0,
 251,
 {'1%': -3.4566744514553016,
  '5%': -2.8731248767783426,
  '10%': -2.5729436702592023},
 -392.7678447135629)
```

Besides, we can conduct the KPSS test with $H_0 : |\beta| < 1$.

In [105…

```python
from statsmodels.tsa.stattools import kpss
import warnings
warnings.filterwarnings("ignore") # suppress the warnings

kpss(np.log(df['Adj Close'])) # The KPSS test statistic, the p-value of the test
# The critical values at 10%, 5%, 2.5% and 1%
```

Out[105…

```
(2.343413698824865,
 0.01,
 10,
 {'10%': 0.347, '5%': 0.463, '2.5%': 0.574, '1%': 0.739})
```

Both tests imply that the log prices are not stationary.

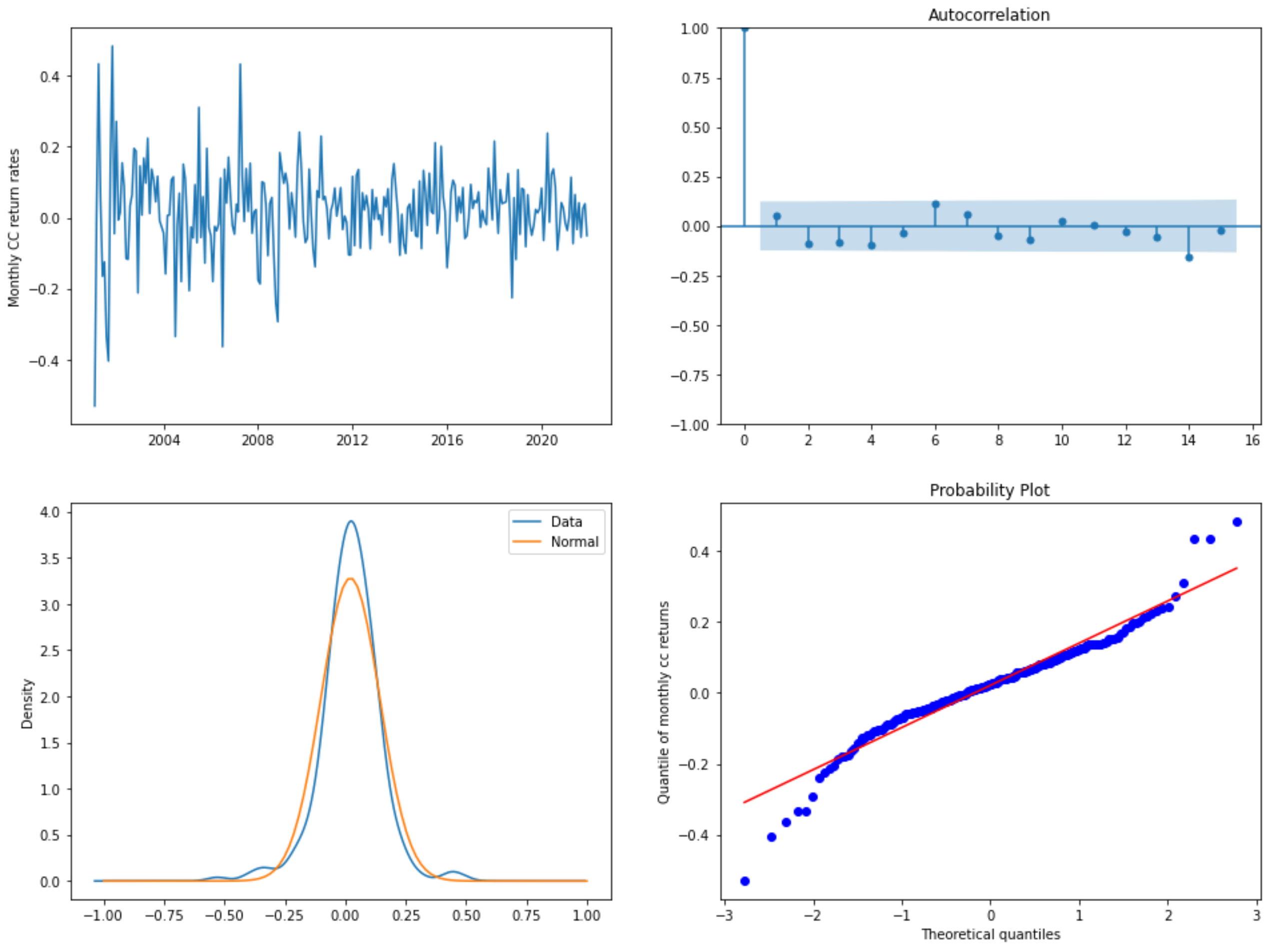
Descriptive statistics of cc returns

In what follows, we will create a four panel plot containing the time series, sample acf, density plot, and normal Q-Q plot.

In [108…

```python
import matplotlib.pyplot as plt
from scipy.stats import norm, probplot
from statsmodels.graphics.tsaplots import plot_acf

fig, ax = plt.subplots(nrows=2, ncols=2, figsize=(16, 12)) # a 2-row, 2-column f
ax[0,0].plot(df['Date'][1:], CC)
ax[0,0].set_ylabel('Monthly CC return rates')

plot_acf(CC, lags = 15, ax = ax[0, 1])

CC.plot(kind='density', label = 'Data', ax = ax[1, 0])
xseq = np.linspace(-1, 1, 100)
ax[1, 0].plot(xseq, norm.pdf(xseq, np.mean(CC), np.std(CC)), label = 'Normal')
ax[1, 0].legend()

probplot(CC, dist="norm", plot=ax[1, 1])
```

```python
ax[1, 1].set_ylabel('Quantile of monthly cc returns')

fig.show()
```



In [109…

```python
from scipy.stats import jarque_bera

jarque_bera(CC) # reject H_0: CC follows a normal distribution
```

Out[109…

```
Jarque_beraResult(statistic=132.48036803060393, pvalue=0.0)
```

Stylized Facts on stock return

1. Nearly mds: $E[r_t|\mathcal{F}_{t-1}] \approx 0$
2. Volatility clustering; periods of high and low volatility
   - Persistent volatility modeling
3. No (or weak) autocorrelation
4. Heavier tail than normal distribution

ARCH and GARCH

The ARCH and GARCH models are employed to model the time-varying volatility of stock returns.

ARCH(1):

$$
\begin{aligned}
r_t &= \sigma_t e_t,\ e_t \sim \text{ iid } N(0,1) \\
\sigma_t^2 &= \omega + \alpha_1 r_{t-1}^2,\ 0 < \alpha_1 < 1, \omega > 0
\end{aligned}
$$

GARCH(1, 1):

$$
\begin{aligned}
r_t &= \sigma_t e_t,\ e_t \sim \text{ iid } N(0,1) \\
\sigma_t^2 &= \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2
\end{aligned}
$$

In what follows, I will show that the ARCH(1) model are consistent with the stylized facts.

1. $E[r_t|\mathcal{F}_{t-1}] = 0$

$$
\begin{aligned}
E[r_t|\mathcal{F}_{t-1}] &= E[\sigma_t e_t|\mathcal{F}_{t-1}] \\
&= \sigma_t E[e_t|\mathcal{F}_{t-1}] \qquad [\text{by definition of conditional expectation}] \\
&= \sqrt{\omega + \alpha_1 r_{t-1}^2} E[e_t] \qquad [\text{since } e_t \sim iid\ N(0,1)] \\
&= 0.
\end{aligned}
$$

1. Volatility clustering

$$
\begin{aligned}
Var(r_t|\mathcal{F}_{t-1}) &= E[r_t^2|\mathcal{F}_{t-1}] \qquad [\text{by zero conditional mean}] \\
&= E[\sigma_t^2 e_t^2|\mathcal{F}_{t-1}] \\
&= \sigma_t^2 E[e_t^2] \\
&= \omega + \alpha_1 r_{t-1}^2.
\end{aligned}
$$

1. No autocorrelation. For $j > 0$,

$$
\begin{aligned}
Cov(r_t, r_{t-j}) &= E[r_t r_{t-j}] - E[r_t]E[r_{t-j}] \\
&= E[r_t r_{t-j}] \qquad [\text{zero conditional expectation implies zero mean}] \\
&= E[E[r_t r_{t-j}|\mathcal{F}_{t-1}]] \qquad [\text{law of iterated expectations}] \\
&= E[r_{t-j} E[r_t|\mathcal{F}_{t-1}]] \\
&= E[r_{t-j} 0] \\
&= 0.
\end{aligned}
$$

1. Heavier tail than normal distribution. For a normal distribution, its kurtosis is 3. See the picture Skewness and Kurtosis.jpeg in week 5 or 6.

$$
\begin{aligned}
kurt(r_t) &= E\left[\left(\frac{r_t - E[r_t]}{\sqrt{Var(r_t)}}\right)^4\right] \\
&= \frac{E[r_t^4]}{E[r_t^2]^2} \qquad [Var(r_t) = E[r_t^2] \text{ because of zero mean}] \\
&= \frac{E[E[r_t^4|\mathcal{F}_{t-1}]]}{E[r_t^2]^2} \qquad [\text{law of iterated expectations}] \\
&= \frac{E[\sigma_t^4 E[e_t^4|\mathcal{F}_{t-1}]]}{E[r_t^2]^2} \\
&= \frac{E[\sigma_t^4 E[e_t^4]]}{E[r_t^2]^2} \qquad [\text{since } e_t \sim iid\ N(0,1)] \\
&= 3\frac{E[\sigma_t^4]}{E[r_t^2]^2} \qquad [E[e_t^4] = 3 \text{ since } e_t \sim iid\ N(0,1)]
\end{aligned}
$$

Moreover, through iterations,

$$
\begin{aligned}
E[r_t^2] &= E[E[r_t^2|\mathcal{F}_{t-1}]] \\
&= E[\omega + \alpha_1 r_{t-1}^2] \\
&= \omega + \alpha_1 E[r_{t-1}^2] \\
&= \omega + \alpha_1(\omega + E[r_{t-2}^2]) \\
&= \omega + \omega\alpha_1 + \alpha_1 E[r_{t-2}^2] \\
&= \omega + \omega\alpha_1 + \omega\alpha_1^2 + \alpha_1^3 E[r_{t-3}] \\
&= \cdots \\
&= \frac{\omega}{1-\alpha_1}
\end{aligned}
$$

since $|\alpha_1| < 1$. Also,

$$
\begin{aligned}
E[\sigma_t^2] &= E[E[\sigma_t^2|\mathcal{F}_{t-1}]] \\
&= E[\omega + \alpha_1 r_{t-1}^2] \\
&= \omega + \alpha_1 E[r_{t-1}^2] \\
&= \omega + \alpha_1 \frac{\omega}{1-\alpha_1} \\
&= \frac{\omega}{1-\alpha_1}.
\end{aligned}
$$

By the Jensen's inequality, $E[f(X)] \geq f(E[X])$ if f is convex. Let $f(x) = x^2$ and $X = \sigma_t^2$. Then,

$$E[\sigma_t^4] = E[(\sigma_t^2)^2] \geq E[\sigma_t^2]^2.$$

As a result,

$$\frac{E[\sigma_t^4]}{E[r_t^2]^2} \geq \frac{E[\sigma_t^2]^2}{E[r_t^2]^2} = 1,$$

and we conclude $kurt(r_t) \geq 3$.