F500: Empirical Finance, Lecture 1
Efficient Markets Hypothesis and Predictability of Asset Returns I

Oliver Linton
obl20@cam.ac.uk

June 25, 2019

Outline

Reading: Linton (2019), Chapter 3.

Prices and Returns

Definition

The capital gain (Return) associated with a price process $\{P_t\}$, over the holding period $[t, t+s]$ is

$$R_{t:t+s} = \frac{P_{t+s} - P_t}{P_t} = \mathcal{R}_{t:t+s} - 1$$

Definition

Continuously compounded returns or Logarithmic return

$$r_{t:t+s} \equiv \log\left(1 + R_{t:t+s}\right) = \log \frac{P_{t+s}}{P_t} = p_{t+s} - p_t$$

Usually take $s = 1$ and denote $R_t = R_{t-1:t}$ and $r_t = r_{t-1:t}$.

Definition

Stock Index values. For some weights w_{jt}

$$I_t = \sum_{j=1}^{J} w_{jt} P_{jt}$$

Equal weighted, price weighted (Dow Jones), value weighted (S&P500)

- Dividends should be added to capital gain to make total return. For indexes, this is usually done through reinvestment. For individual stocks dividends may be payed once or twice a year and so quite dificult to work with.
- Taxes, inflation, and exchange rates may also be relevant to investors when calculatng their return.

Calendar Time or Trading Time

- We will use time series analysis, which requires equally spaced data. In many applications this requires an additional justification.
- For daily frequency, usually take closing price to closing price, in which case there are "gaps".
- There are two approaches to this

Definition

Calendar time - returns are generated in calendar time, observe $P_1, P_2, P_3, P_4, P_5, P_8, P_9, \ldots$ and so Friday to Monday is a three day return. Have to deal with the gaps.

Definition

Trading time - returns are only generated when exchange is open so Friday to Monday is a one day return. Dont' have to deal with the gaps.

Remarks on Prices vs Log Prices

- Nice feature of log returns is that they can take any value, whereas actual returns are limited from below by limited liability, i.e., you cant lose more than your stake means that

$$R_t \geq -1,$$

whereas

$$r_t \in \mathbb{R}.$$

- Therefore, r_t is logically consistent with a normal distribution, whereas R_t is not.

- Logarithmic returns are **time additive**

$$\begin{aligned} r_{t:t+H} &= \log P_{t+H} - \log P_t \\ &= \log P_{t+H} - \log P_{t+H-1} + \ldots + \log P_{t+1} - \log P_t \\ &= r_{t+H} + r_{t+H-1} + \ldots + r_{t+1} \end{aligned}$$

e.g., weekly returns are the sum of the five daily returns

- Not true for actual returns. In that case

$$\begin{aligned} 1 + R_{t:t+H} &= \mathcal{R}_{t,H} = \frac{P_{t+H}}{P_t} = \frac{P_{t+H}}{P_{t+H-1}} \times \frac{P_{t+H-1}}{P_{t+H-2}} \times \cdots \times \frac{P_{t+1}}{P_t} \\ &= \mathcal{R}_{t+H-1,1} \times \cdots \times \mathcal{R}_{t,1} \\ &= [1 + R_{t+H-1,1}] \times \cdots \times [1 + R_{t,1}] \end{aligned}$$

Can take geometric mean to give a per period figure $(1 + R_{t:t+H})^{1/H}$.

- However, returns are **portfolio additive**

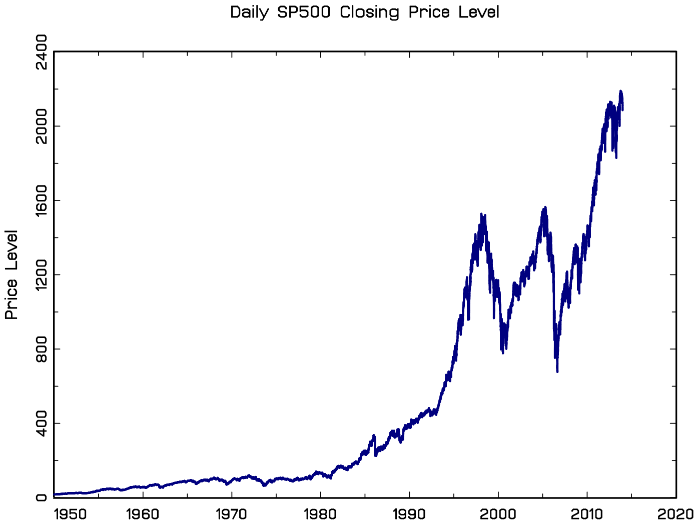
$$R_t\left(w\right) = w_1 \; R_{1t} + w_2 \; R_{2t} + \ldots + w_N \; R_{Nt}$$

- But log returns are not

$$r_t\left(w\right) = \log\left(\frac{w_1 \; P_{1t} + \ldots + w_N P_{Nt}}{w_1 \; P_{1,t-1} + \ldots + w_N P_{N,t-1}}\right) \neq w_1 \; r_{1t} + \ldots + w_N \; r_{Nt}$$

- For small returns, such as high-frequency returns, $r \approx R$, i.e., we have by Taylor theorem
$$r = \log(1 + R) \approx R.$$
This means that in such cases returns are similar. Over long horizon though, returns and log returns are quite different.

S&P500 index was 17.03 on 10/1/1950 and was 1842.37 on 10/01/2014. Gross return is 108.183. Annual return of 7.6%.



The Random Walk Hypothesis

The traditional model for stock prices, says that prices evolve randomly.

Definition

The random walk

$$X_t = \mu + X_{t-1} + \varepsilon_t,$$

where $X_t = p_t$ or $X_t = P_t$. Three general assumptions:
(1) RW1: $\varepsilon_t \sim IID$; $E\varepsilon_t = 0$;
(2) RW2: ε_t *independent over time*; $E\varepsilon_t = 0$;
(3) RW3: For all k : $\operatorname{cov}(\varepsilon_t, \varepsilon_{t-k}) = 0$

Historically, μ was often assumed to be zero and ε_t normally distributed, even stronger than (1).
We also consider the more natural assumption of MDS

Definition

A martingale is a time-series process X_t obeying

$$E\left[X_{t+1} \mid X_t,\ X_{t-1},\ \ldots\right] = X_t$$

or equivalently, call $\varepsilon_{t+1} = X_{t+1} - X_t$ a martingale difference sequence (MDS) if

$$E\left[\varepsilon_t \mid X_{t-1},\ X_{t-2},\ \ldots\right] = 0.$$

This corresponds with the notion of a fair game: If you toss a coin against opponent and bet successively at fair odds with initial capital X_0, current capital X_t is a martingale.

This is the case that $\mu = 0$; More generally, we might assume for stock returns that X_t is a martingale plus drift, i.e.,

$$\varepsilon_{t+1} = X_{t+1} - X_t - \mu$$

is a martingale difference (MDS). Increments are essentially unpredictable given past information.

Martingale property implies that

$$\operatorname{cov}(\varepsilon_t, g(X_{t-1},\ X_{t-2},\ \ldots)) = 0$$

for any (measurable) function g (trading strategy).
In particular returns are uncorrelated (RW3) but also

$$\operatorname{cov}(r_t, g(r_{t-1}, \ldots, r_{t-p})) = 0$$

So stronger condition than RW3. Call it RW2.5.

Theorem

Provided $E\varepsilon_t^2 \leq C < \infty$,

$$RW1 \Longrightarrow RW2 \Longrightarrow RW2.5 \Longrightarrow RW3 : \operatorname{cov}\left(\varepsilon_t, \varepsilon_{t-k}\right) = 0$$

Efficient Markets Hypothesis

Definition

Fama (1970, JoF): A market in which prices always fully reflect available information is called efficient (EMH)

- If prices are predictable $\Rightarrow$ opportunities for superior returns (free lunch) $\Rightarrow$ will be competed away immediately by a lot of hungry traders $\Rightarrow$ unpredictable random walk
  - If a security believed to be underpriced, buying pressure $\Rightarrow$ jump up to a level where no longer thought a bargain
  - If a security believed to be overpriced, (short-)selling pressure $\Rightarrow$ jump down to a level where no longer thought too expensive
- As a result, market forces respond to news quckly and make prices the best available estimates of fundamental values, i.e. values justified by likely future cash flows and preferences of investors/consumers

We distinguish among three forms of market efficiency depending on the information set with respect to which efficiency is defined

1. **Weak form.** (1) Information from historical prices are fully reflected in the current price; (2) One can't earn **abnormal profits** from trading strategies based on past prices alone.
2. **Semi strong form.** (1) All public information (past prices, annual reports, quality of management, earnings forecasts, macroeconomic news, etc.) is fully reflected in current prices; (2) One can't earn abnormal profits from trading strategies based on public information.
3. **Strong form.** (1) All private and public information is fully reflected in current prices; (2) One can't earn abnormal profits from trading strategies based on all information including public and private.

$$\textit{Strong} \Longrightarrow \textit{Semi strong} \Longrightarrow \textit{Weak}$$

"Normal" Return

- Suppose that μ_t is the required or normal return over the interval $[t, t+1]$, that arrives from an asset pricing model, and R_{t+1} is the realized random return. Then under the null hypothesis ($\mathcal{F}_t$) it holds that the return on any risky asset over the same interval satisfies

$$E(R_{t+1}|\mathcal{F}_t) = \mu_t.$$

  You can't make more money on average than μ_t.

- We can write $\mu_t = R_{ft} + \pi_t,$ where R_{ft} is the risk free rate and π_t is the risk premium (known at time t). In conclusion, we should allow

$$R_{t+1} = R_{ft} + \pi_t + \varepsilon_{t+1} = \mu_t + \varepsilon_{t+1},$$

  where ε_t is an MDS with respect to some information set $\mathcal{H}_t,$ (prices, i.e., all past ε) $\mathcal{G}_t$ (public), $\mathcal{F}_t$ (all).

- The risk premium π_t we may suppose is non-negative and is determined by some economic model.

Technical analysts

- Chartists try to identify regularity of some patterns in stock prices, hoping to exploit them and profit. They believe patterns repeated in prices. e.g. Head and Shoulders

This one predicts that future prices decline.

- Incompatible with weak form hypothesis
- Lo and Hasanhodzic (2010) connected the analysis of chartists to nonlinear time series analysis. They show how to convert observed price history into a numerical score that identifies say "head and shoulderness". They show that there is some basis to their work, but provide the tools to replace them by automated systems.

Fundamental analysts

- They estimate future cash flows from securities and their riskiness, based on analysis of company-relevant data such as balance sheets as well as the economic environment in which it operates, to determine the proper price of securities. Graham and Dodd class book on investing espoused by Warren Buffett.
- For example, buy stocks with low P/E and sell high P/E stocks
- Warren Buffet: Ratio of Stock market valuation to GDP
- Incompatible with semi-strong form hypothesis

Two Theoretical critiques of EMH

- Grossman and Stiglitz (1980, AER) point out that if information collection and analysis are costly, there must be compensation for such activity in terms of extra risk-adjusted returns, otherwise rational investors would not incur such expenses. Therefore, **Markets cannot be fully informationally efficient**, rather an 'equilibrium degree of disequilibrium'. Weak form may hold but semistrong harder to justify.
- Shleifer and Vishny (1997, JF). Textbook arbitrage is a costless, riskless and profitable trading opportunity; in practice it is usually costly and risky. Also is conducted by a small number of highly specialized professionals using other people's capital (agency relationship). If the misspricing temporarily worsens, investors/clients may judge the manager as incompetent and refuse to provide additional capital (margin call) and make withdrawals, thus forcing him to liquidate positions at the worst time. He loses performance fees, and perhaps a career ender). Therefore, a rational specialized arbitrageur stays away.

An Econometric Critique: Joint Hypothesis Problem

- Any test of weak form EMH must assume an equilibrium asset pricing model that defines ‘normal’ security returns against which investor returns are measured.
- If we reject the hypothesis that investors can’t achieve superior risk-adjusted returns, we don’t know if markets are inefficient or if the underlying model is misspecified.
- Therefore, we can never reject EMH.

- In the next sections we will mostly be assuming that $\mu_t = \mu$ is constant or its variation is small. This can be justified if the frequency is high and or risk aversion is small.
- We test the implication of the weak form EMH that demeaned returns are uncorrelated
- We assume that RW1 holds to make life easy; ideally should assume only RW2.5, we will get there later.

Testing of EMH under RW1

The population autocovariance and autocorrelation functions of a stationary series Y_t

$$\gamma_s = \mathrm{cov}(Y_t, Y_{t-s}) = E\left[(Y_t - EY_t)(Y_{t-s} - EY_{t-s})\right]$$

$$\rho_s = \frac{\gamma_s}{\gamma_0}$$

for $s = 0, \pm 1, \pm 2, \ldots$ Take $Y_t = r_t$ or R_t. The efficient markets hypothesis (RW3) says that $\gamma_s, \rho_s = 0 \quad$ for all $s \neq 0$.
Can estimate these quantities by the sample equivalents

$$\widehat{\gamma}_s = \frac{1}{T}\sum_{t=s+1}^{T}(Y_t - \overline{Y})(Y_{t-s} - \overline{Y})$$

$$\widehat{\rho}_s = \frac{\widehat{\gamma}_s}{\widehat{\gamma}_0}.$$

- Assume further that Y_t is i.i.d. It can be shown that for any k,

$$\sqrt{T}\widehat{\rho}_k \Longrightarrow N(0,1)$$

under the null hypothesis of no correlation.

- Therefore, you can test the null hypothesis by comparing $\widehat{\rho}_k$ with the so-called 'Bartlett intervals'

$$[-z_{\alpha/2}/\sqrt{T}, z_{\alpha/2}/\sqrt{T}],$$

where z_α are normal critical values. Values of $\widehat{\rho}_k$ lying outside this interval are inconsistent with the null hypothesis. Literally, this is testing the hypothesis that $\rho_k = 0$ versus $\rho_k \neq 0$ for a given k.

- Under the alternative hypothesis

$$\sqrt{T}\widehat{\rho}_k \xrightarrow{P} \infty$$

for at least one k.

- In fact, under this assumption we have

$$\sqrt{T}(\widehat{\rho}_1, \ldots, \widehat{\rho}_P)^\intercal \Longrightarrow N(0, I_P).$$

The Box–Pierce Q statistic

$$Q = T \sum_{j=1}^{P} \widehat{\rho}_j^2$$

can be used to test the joint hypothesis that $\rho_1 = 0, \ldots, \rho_P = 0$ versus the general alternative. We have

$$Q \Longrightarrow \chi_P^2$$

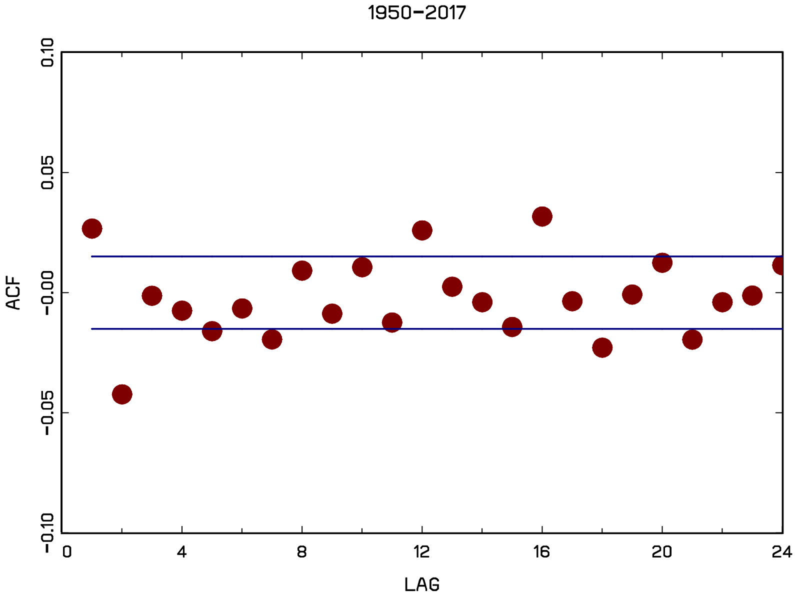
under the null hypothesis, so reject when $Q > \chi_P^2(\alpha)$ for an α-level test.

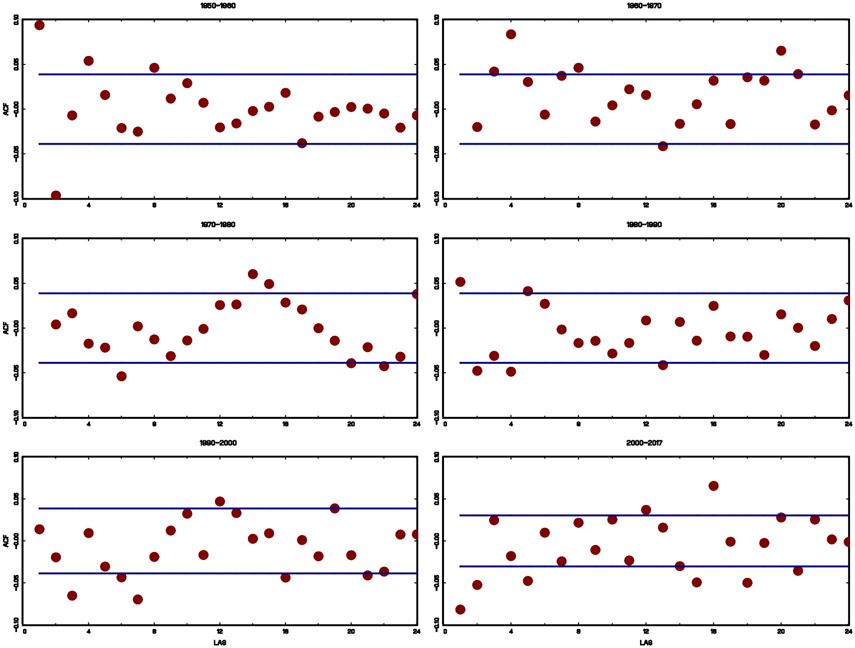
- Box-Ljung version is known to have better finite sample performance (smaller bias)

$$Q = T(T+2) \sum_{j=1}^{p} \frac{\widehat{\rho}_j^2}{T - j}$$

- CLM results. 19620703-19941230, Daily, weekly, monthly. CRSP value weighted and equal weighted indexes. A sample of 411 individual securities from the CRSP database.
- Table 2.4
  - Positive (first lag) autocorrelation for daily indexes (0.1-0.43), which are significant using the iid standard errors $1/\sqrt{T}$. Statistically significant Q_5 and Q_{10}.
  - Weaker at weekly and monthly horizon. Weaker for value weighted versus equal weighted.
  - Results are not stable across subperiods
- Small negative autocorrelation for individual stocks at daily horizon. How to explain the different results between individual stocks and index? Read notes.
- Lead lag relations between large and small stocks (Explained by cross-correlation)
- Violation of weak form efficiency implies violation of semi-strong and strong. Question is whether the violation is large in an economic sense, stable over time, robust to different assumptions.

T>15,000




1950-1960
1960-1970
1970-1980
1980-1990
1990-2000
2000-2017
ACF
LAG
0.10
0.05
0.00
-0.05
-0.10
0
4
8
12
16
20
24

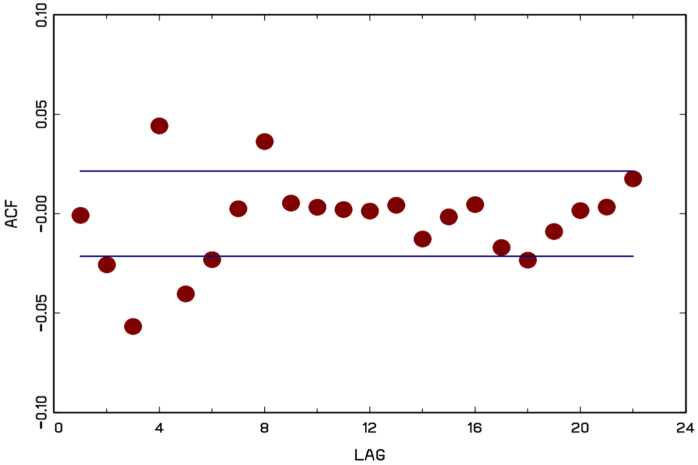


Figure: Correlogram of FTSE100 daily returns from 1984-2017

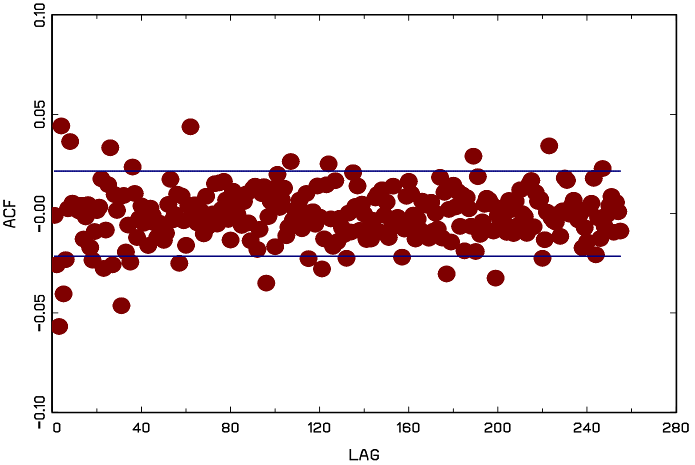


Figure: Correlogram of FTSE100 daily returns long horizon

Dow Jones Industrial Average, as of January, 2013

Name	Cap $b	Name	Cap $b
Alcoa Inc.	9.88	JP Morgan	172.43
AmEx	66.71	Coke	168.91
Boeing	58.58	McD	90.21
Bank of America	130.52	MMM	65.99
Caterpillar	62.07	Merck	127.59
Cisco Systems	108.74	MSFT	225.06
Chevron	216.27	Pfizer	191.03
du Pont	42.64	Proctor & Gamble	188.91
Walt Disney	92.49	AT&T	200.11
General Electric	222.31	Travelers	28.25
Home Depot	94.47	United Health	53.21
HP	29.49	United Tech	77.89
IBM	219.20	Verizon	126.43
Intel	105.20	Wall Mart	231.02
Johnson[2]	198.28	Exxon Mobil	405.60

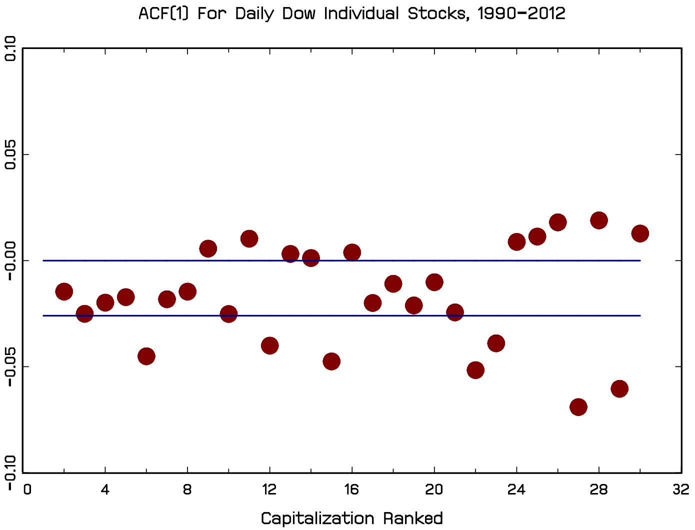


Figure: ACF(1) of daily Dow Stock returns against market capitalization

CLM Table 2.7 - averages of autocorrelations across firms. Suppose that we compute $\rho_i(.)$ for a cross section of stocks ($i = 1, \ldots, n$) and report the average estimated value

$$\widehat{\overline{\rho}}(k) = \frac{1}{n}\sum_{i=1}^{n}\widehat{\rho}_i(k).$$

The EMH implies that $\overline{\rho}(k) = \frac{1}{n}\sum_{i=1}^{n}\rho_i(k) = 0$ for all k. What is the sampling distribution of this estimate? Not given in CLM.

Theorem

Suppose that Y_{it} are i.i.d. across t with finite variance, and let

$$v = \frac{1}{n^2}\left(n + \sum\sum_{i \neq j}\omega_{ij}^2\right),$$

where $\omega_{ij} = \mathrm{corr}(Y_{it}, Y_{jt})$. Then for any p, n as $T \to \infty$

$$\sqrt{T}(\widehat{\overline{\rho}}(1), \ldots, \widehat{\overline{\rho}}(p))^{\intercal} \Longrightarrow N(0, vI_p).$$

Reject the null hypothesis if

$$\overline{\rho}(k) \notin \left[-z_{\alpha/2}\sqrt{\frac{1}{T}\widehat{v}}, z_{\alpha/2}\sqrt{\frac{1}{T}\widehat{v}} \right]$$

$$\widehat{v} = \frac{1}{n^2}\left[n + \sum_{i \neq j}\sum \widehat{\omega}_{ij}^2 \right], \qquad \widehat{\omega}_{ij} = \frac{1}{T}\sum_{t=1}^{T}(Y_{it} - \overline{Y}_i)(Y_{jt} - \overline{Y}_j)$$

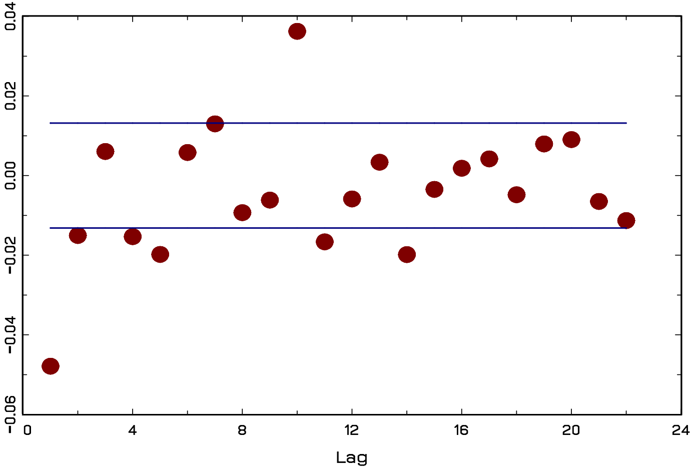


Figure: Average ACF of Dow stocks daily returns

Also care about the cross -autocorrelation

$$\mathrm{cov}(Y_t, X_{t-j}) = \gamma_{XY}(j), \quad \mathrm{corr}(Y_t, X_{t-j}) = \rho_{XY}(j) = \frac{\gamma_{XY}(j)}{\sqrt{\gamma_{XX}(0)\gamma_{YY}(0)}}$$

Under the EMH, $\gamma_{XY}(j) = \rho_{XY}(j) = 0$ for all $j = \pm 1, \pm 2, \ldots$

A portfolio is a weighted average of stock returns. In some case we find individual stocks have negative autocorrelation but stock indexes have positive autocorrelation. How can that be?

For

$$\begin{aligned} & \mathrm{cov}(w_Y Y_t + w_X X_t, w_Y Y_{t-j} + w_X X_{t-j}) \\ = \; & w_Y^2 \gamma_{YY}(j) + w_X^2 \gamma_{XX}(j) + w_Y w_X \gamma_{YX}(j) + w_Y w_X \gamma_{XY}(j). \end{aligned}$$

Suppose that $\gamma_{YY}(j), \gamma_{XX}(j) < 0$. If $w_Y, w_X > 0$ and $\gamma_{YX}(j), \gamma_{XY}(j) > 0$, then we may have

$$\mathrm{cov}(w_Y Y_t + w_X X_t, w_Y Y_{t-j} + w_X X_{t-j}) > 0.$$

Standard Errors under RW2.5

- Normality is not needed for the above simple distribution theory, but we do require at least $EY_t^2 < \infty$
- The RW1 theory is too restrictive. Should allow for
  - Heteroskedasticity
  - Dependence (in higher moments)
  - Nonstationarity or non-identically distributed
- In which case, the distribution of sample autocorrelations and Box-Pearce statistics is more complicated

Suppose that only RW2.5 (MDS) holds, i.e., nonlinear dependence is allowed in higher moments. Special case with known $EY_t = 0$. Then, for stationary processes we may show that

$$\frac{1}{T}\sum_t Y_t^2 \xrightarrow{P} EY_t^2 \quad < \quad \infty$$

$$\frac{1}{\sqrt{T}}\sum_t Y_t Y_{t-j} \Longrightarrow N(0, \mathrm{E}(Y_t^2 Y_{t-j}^2))$$

This is because by the MDS assumption $E(Y_t Y_{t-j}) = 0$, so that

$$\begin{aligned}
\mathrm{var}\left(\frac{1}{\sqrt{T}}\sum_t Y_t Y_{t-j}\right) &= E\left[\left(\frac{1}{\sqrt{T}}\sum_t Y_t Y_{t-j}\right)^2\right] \\
&= \frac{1}{T}\sum_t E(Y_t^2 Y_{t-j}^2) + \frac{1}{T}\sum\sum_{t\neq s} E\left(Y_t Y_{t-j} Y_s Y_{s-j}\right) \\
&= \frac{1}{T}\sum_t E(Y_t^2 Y_{t-j}^2)
\end{aligned}$$

because $E\left(Y_t Y_{t-j} Y_s Y_{s-j}\right) = 0$ by MDS when $t > s$ or $s > t$.

Therefore,

$$\sqrt{T}\widehat{\rho}(j) = \frac{\frac{1}{\sqrt{T}}\sum_t Y_t Y_{t-j}}{\frac{1}{T}\sum_t Y_t^2} \Longrightarrow N\left(0, \frac{\mathrm{E}(Y_t^2 Y_{t-j}^2)}{\mathrm{E}^2(Y_t^2)}\right).$$

But in general

$$\mathrm{E}(Y_t^2 Y_{t-j}^2) \neq \mathrm{E}(Y_t^2)\mathrm{E}(Y_{t-j}^2)$$

when dependent heteroskedasticity allowed for. In fact

$$\frac{\mathrm{E}(Y_t^2 Y_{t-j}^2)}{\mathrm{E}^2(Y_t^2)} = 1 + \frac{\mathrm{cov}(Y_t^2, Y_{t-j}^2)}{\mathrm{E}^2(Y_t^2)} = 1 + \frac{\mathrm{var}(Y_t^2)}{\mathrm{E}^2(Y_t^2)}\mathrm{corr}(Y_t^2, Y_{t-j}^2)$$

In fact

$$\frac{\mathrm{var}(Y_t^2)}{\mathrm{E}^2(Y_t^2)} = \frac{\mathrm{E}(Y_t^4) - \mathrm{E}^2(Y_t^2)}{\mathrm{E}^2(Y_t^2)} = (\kappa_4(Y_t) - 1)$$

so that the asymptotic variance is

$$\frac{\mathrm{E}(Y_t^2 Y_{t-j}^2)}{\mathrm{E}^2(Y_t^2)} = 1 + \overbrace{(\kappa_4(Y_t) - 1)}^{\textit{heavy}\ \text{tails}} \times \overbrace{\mathrm{corr}(Y_t^2, Y_{t-j}^2)}^{\textit{dependent}\ \text{heteroskedasticity}}$$

$$\overbrace{\leq}^{\mathrm{corr} \geq 0} 1 + (\kappa_4(Y_t) - 1)$$

where $\kappa_4(Y_t) \geq 1$ is the kurtosis of the series Y_t. The asymptotic variance of $\widehat{\rho}(j)$ can be arbitrarily large.

In principle, standard errors that allow for this dependence may be a lot wider than the Bartlett ones. In some cases they may be smaller.

	$\rho_{Y^2}(1)$	$\kappa_4(Y)$		$\rho_{Y^2}(1)$	$\kappa_4(Y)$
Alcoa	0.2844	10.5723	JP Morgan	0.1201	10.1305
AmEx	0.2172	9.9907	Coke	0.3217	12.8566
Boeing	0.3259	28.0133	McD	0.1801	7.4358
B of A	0.1724	9.5624	MMM	0.1137	7.4254
Caterpillar	0.1203	6.9534	Merck	0.0412	22.9570
Cisco	0.2117	8.4353	MSFT	0.1224	8.1895
Chevron	0.1340	8.3685	Pfizer	0.1497	6.2051
du Pont	0.2604	12.1940	P&G	0.0376	62.9128
Walt Disney	0.1797	6.9172	AT&T	0.1494	8.0896
GE	0.1262	10.1261	Travelers	0.3401	16.2173
Home Depot	0.2573	10.7188	United Health	0.0758	23.0302
HP	0.0656	17.1360	United Tech	0.0365	21.5274
IBM	0.0888	9.6413	Verizon	0.2203	7.9266
Intel	0.1116	9.9477	Wall Mart	0.1841	6.1845
Johnson[2]	0.1211	8.6509	Exxon Mobil	0.2947	11.7152
			S&P500	0.2101	11.4509

Correct Standard Errors for RW2.5

- In the general nonzero mean case case we have the same result with $\widetilde{Y}_t = Y_t - \overline{Y}_t$, specifically

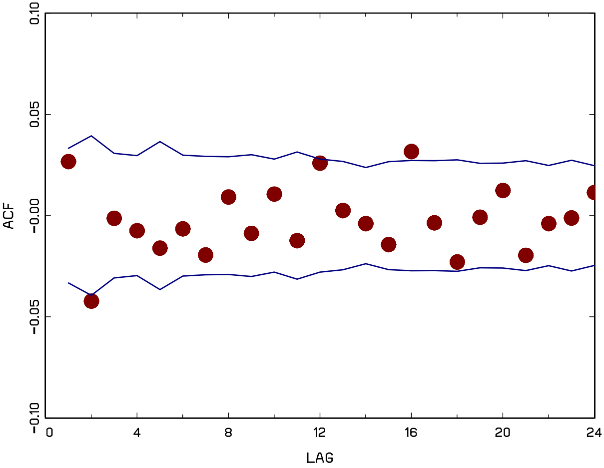
$$\sqrt{\frac{\left(\sum_t \widetilde{Y}_t^2\right)^2}{\sum_t \widetilde{Y}_t^2 \widetilde{Y}_{t-j}^2}}\widehat{\rho}(j) \Longrightarrow N(0,1).$$

- Therefore, instead of the Bartlett interval we should compute the interval

$$\left[-z_{\alpha/2}\sqrt{\frac{\sum_t \widetilde{Y}_t^2 \widetilde{Y}_{t-j}^2}{\left(\sum_t \widetilde{Y}_t^2\right)^2}}, z_{\alpha/2}\sqrt{\frac{\sum_t \widetilde{Y}_t^2 \widetilde{Y}_{t-j}^2}{\left(\sum_t \widetilde{Y}_t^2\right)^2}}\right]$$

- Robust tests can be constructed more generally. CLM try to do this under RW3 only, but their theory is not quite correct.

Heteroskedasticity consistent standard errors



AutoRegression Tests

- Fit the autoregression

$$Y_t = \mu + \beta_1 Y_{t-1} + \ldots + \beta_P Y_{t-P} + \varepsilon_t$$

- Test the hypothesis (Standard F-test)

$$H_0 : \beta_1 = \cdots = \beta_P = 0$$

versus general alternative. When $P = 1$ this is equivalent to ACF test, but not for $P > 1$.

- Can do t-tests on the slope coefficients using OLS standard errors or Whites standard errors
- Under the iid assumption rw1, the asymptotic variance can be estimated by

$$\widehat{V}_{OLS} = \widehat{\sigma}_{\varepsilon}^2 (X^{\intercal}X)^{-1},$$

where X is the $(T - P - 1) \times P + 1$ matrix whose first column consists of ones, whose second column consists of the observations $Y_{P+1}, \ldots, Y_T$ etc. $\widehat{\sigma}_{\varepsilon}^2$ is the residual sample variance.

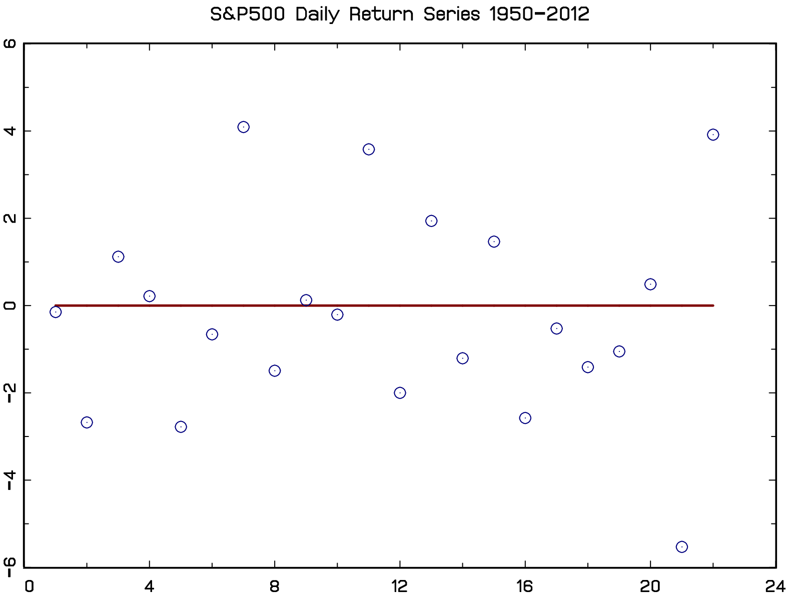
- The Whites standard errors are

$$\widehat{V}_W = (X^{\intercal}X)^{-1} X^{\intercal} D X (X^{\intercal}X/)^{-1},$$

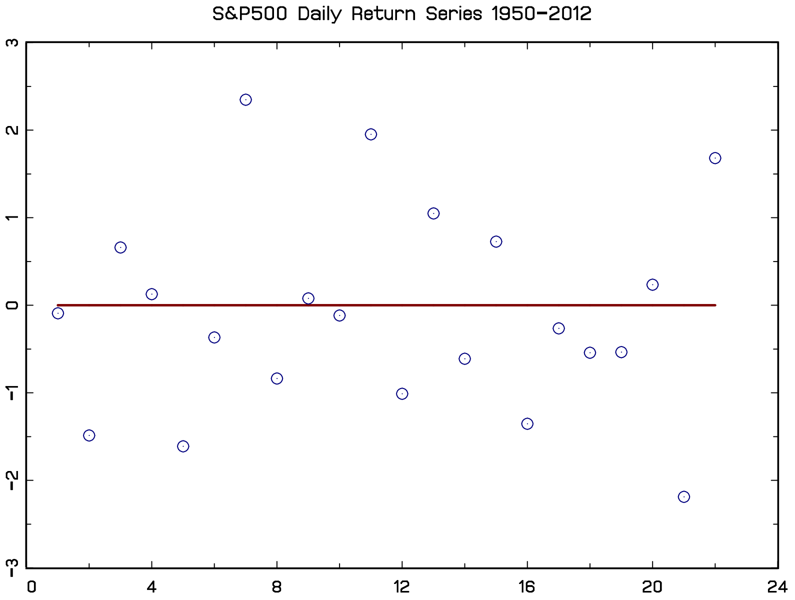
$$D = \mathrm{diag}\{\widehat{\varepsilon}_1^2, \ldots, \widehat{\varepsilon}_T^2\}$$

Figures shows t statistics for each coefficient for AR(22) with ols standard errors and then Whites standard errors.

OLS standard errors (t-stat)



White's standard errors (t-stat)



Advantages and disadvantages of regression tests

- Advantages
  - Designed more for the conditional moment hypothesis and prediction
- Disadvantages
  - If P is large, covariate matrix in OLS can be rank deficient, certainly when $P \longrightarrow \infty$.
  - Not graphical or directional