F500: Empirical Finance
Lecture 2: Efficient Markets Hypothesis and Predictability of Asset Returns

Oliver Linton
obl20@cam.ac.uk

January 29, 2020

Outline

1. Variance Ratio Tests
2. Empirical Evidence
3. Fads Model Alternative to EMH and Long Run implications
4. Trading strategy based evidence
5. Nonlinear/Nonparametric Tests
    1. Cowles and Jones test
    2. Runs test
6. Regression based tests of Semi-Strong EMH
7. Discussion of evidence

Reading: Linton (2019), Chapter 4

Variance Ratio Tests

Variance ratio tests are a close relative to autocorrelation and Box-Pierce tests.

- Suppose that returns are stationary and in particular $Er_t = \mu$ and $var(r_t) = \sigma^2$. Consider the 2-period return

$$r_{t:t+2} = p_{t+2} - p_t = p_{t+2} - p_{t+1} + p_{t+1} - p_t = r_{t+2} + r_{t+1}.$$

- We have

$$\mathrm{var}(r_{t:t+2}) = \mathrm{var}(r_{t+2}) + \mathrm{var}(r_{t+1}) + 2\mathrm{cov}\left(r_{t+2}, r_{t+1}\right).$$

- Under the assumption that returns are uncorrelated (RW3), we further have $\mathrm{cov}(r_{t+2}, r_{t+1}) = 0$ and so

$$\mathrm{var}(r_{t:t+2}) = \mathrm{var}(r_{t+2}) + \mathrm{var}(r_{t+1}) = 2\mathrm{var}(r_t).$$

  Therefore

$$VR(2) = \frac{\mathrm{var}(r_{t:t+2})}{2\mathrm{var}(r_t)} = 1.$$

- Define the q period return

$$r_{t:t+q} = r_{t+q} + r_{t+q-1} + \ldots + r_{t+1}.$$

- Under H_0 (RW3), we have

$$VR(q) = \frac{\mathrm{var}(r_{t:t+q})}{q\mathrm{var}\left[r_t\right]} = 1.$$

Under the alternative $VR(q) \neq 1$

- A useful result. For a general stationary process with ACF $\{\rho(j),$ $j = 1, 2, \ldots\}$

$$VR(q) = 1 + 2\sum_{j=1}^{q-1}\left(1 - \frac{j}{q}\right)\rho(j).$$

It is just a "linear functional" of the correlogram. The direction of the ratio (compared with 1) depends on all the first $q-1$ autocorrelations and their relative magnitudes. Under EMH, $\rho(j) = 0$ and $VR(q) = 1$

Example

Suppose that $q = 5$ and $\rho(1) = 0.3$, $\rho(2) = 0.3$, $\rho(3) = 0.2$, $\rho(4) = 0.1$:

$$VR(5) = 1 + 2\left(\frac{4}{5} \times 0.3 + \frac{3}{5} \times 0.3 + \frac{2}{5} \times 0.2 + \frac{1}{5} \times 0.1\right) = 2.04$$

$$\sum_{j=1}^{4} \rho(j)^2 = 0.23$$

Example

Suppose that $\rho(1) = 0.3$, $\rho(2) = -0.3$, $\rho(3) = -0.2$, $\rho(4) = 0.1$:

$$VR(5) = 1 + 2\left(\frac{4}{5} \times 0.3 + \frac{3}{5} \times -0.3 + \frac{2}{5} \times -0.2 + \frac{1}{5} \times 0.1\right) = 1$$

$$\sum_{j=1}^{4} \rho(j)^2 = 0.23$$

Testing based on Sample Versions

- $nq + 1$ (log) price observations $p_0, \ldots, p_{nq}$
- High Frequency returns (H)

$$r_t = p_t - p_{t-1}, \; t = 1, \ldots, nq,$$

- Low Frequency returns (L)

$$r_{t-q:t} = p_t - p_{t-q},$$

  - Non-overlapping returns (Monday to Monday, Monday to Monday etc), we let $t = q, 2q \ldots, nq,$ a total of n returns
  - Overlapping returns (Monday to Monday plus Tuesday to Tuesday etc), we let $t = q+1, q+2, \ldots, nq,$ a total of $nq - q$ such returns.

$$\widehat{\mu} = \overline{r} = \frac{1}{nq}\sum_{t=1}^{nq} r_t = \frac{1}{nq}\sum_{t=1}^{nq}(p_t - p_{t-1}) = \frac{1}{nq}\left(p_{nq+1} - p_1\right)$$

$$\widehat{\sigma}_H^2 \equiv \frac{1}{nq}\sum_{t=1}^{nq}\left(r_t - \widehat{\mu}\right)^2$$

$$\overbrace{\widehat{\sigma}_{LN}^2(q)}^{\text{"Friday to Friday"}} \equiv \frac{1}{n}\sum_{k=1:t=qk+1}^{n}\left(r_{t-q:t} - q\widehat{\mu}\right)^2$$

$$\overbrace{\widehat{\sigma}_{LO}^2(q)}^{\text{"Monday to Monday,...,Friday to Friday"}} \equiv \frac{1}{nq-q}\sum_{t=q+1}^{nq}\left(r_{t-q:t} - q\widehat{\mu}\right)^2$$

Define overlapping and no-overlapping statistics

$$\widehat{VR}_j(q) = \frac{\widehat{\sigma}^2_{Lj}(q)}{q\widehat{\sigma}^2_H}, \quad j = O, N$$

- Under RW1 we have

$$\sqrt{nq}\left(\widehat{VR}_j(q) - 1\right) \Longrightarrow N(0, V_j(q))$$

$$V_N(q) = 2(q-1) \quad ; \quad V_O(q) = \frac{4q-2}{6q} 2(q-1)$$

- Overlapping method is more efficient (30%-50% smaller variance)

$$V_O(q) \leq V_N(q)$$

- Formal test of null hypothesis of no predictability. Let

$$Z_j(q) = \frac{\sqrt{nq}\left(\widehat{VR}_j\,(q) - 1\right)}{\sqrt{V_j(q)}}$$

- Reject null hypothesis if $|Z_j(q)| > z_{\alpha/2}$ where z_α is normal critical value.
    - If $Z_j(q) > 0$, this means positive autocorrelation (momentum) and
    - if $Z_j(q) < 0$, this means negative autocorrelation (contrarian)

Empirical Evidence

Campbell, Lo and Mackinlay (1997) CRSP data 1962-1978 and 1978-1994; $q =$2,4,8, and 16 weeks. Heteroskedasticity consistent standard errors (RW2)

- Table 2.5 Variance ratios for weekly stock indexes (value weighted and equal weighted). Equal weighted strongly significant, greater than one, and increasing with horizon. Less so in second half 1978-1994. Value weighted similar but not significant. 78-94, VR less than one
- Table 2.6 Three Size sorted portfolios. Smallest and medium strongly significant, greater than one and increasing with horizon. Largest not significant in most recent period
- Table 2.7. Average of VR for 411 individual stocks. No standard errors. Less than one and declining with horizon.

Robust Standard Errors for the Variance Ratio Statistic

Define $\widetilde{X}_t = r_t - \bar{r}$ and for $j, k = 1, \ldots, q$

$$\widehat{\lambda}_{jk} = \frac{1}{T}\sum_{t=1}^{T} \widetilde{X}_{t-j}\widetilde{X}_{t-k}\widetilde{X}_t^2 \quad ; \quad \widehat{\gamma}_0 = \frac{1}{T}\sum_{t=1}^{T} \widetilde{X}_t^2$$

RW1. iid

$$\widehat{V}_{O1}(q) = \frac{4}{T}\sum_{j=1}^{q-1}\left(1 - \frac{j}{q}\right)^2 = \frac{4\left(2q^2 - 3q + 1\right)}{6qT}$$

RW2. Heteroskedasticity robust (Univariate White's)

$$\widehat{V}_{O2}(q) = \frac{4}{T}\sum_{j=1}^{q-1}\left(1 - \frac{j}{q}\right)^2 \frac{\widehat{\lambda}_{jj}}{\widehat{\gamma}_0^2}$$

RW2.5.Heteroskedasticity and Leverage robust (System White's)

$$\widehat{V}_{O3}(q) = \frac{4}{T}\sum_{j=1}^{q-1}\sum_{k=1}^{q-1}\left(1 - \frac{j}{q}\right)\left(1 - \frac{k}{q}\right)\frac{\widehat{\lambda}_{jk}}{\widehat{\gamma}_0^2}$$

Updated Evidence

- We add data from 1994-2013 and used robust standard errors
- We first test for the absence of serial correlation in each of three weekly size-sorted equal-weighted portfolio returns (smallest quantile, central quantile, and largest quantile).
- We compare with the results reported in Campbell, Lo and Mackinlay (1997, P71, Table 2.6).
- We divide the whole sample to three subsamples: 62:07:06-78:09:29 (848 weeks), 78:10:06-94:12:23 (847 weeks) and 94:12:30-13:12:27 (992 weeks).

Table 1-A: Variance ratios for weekly small-size portfolio returns

Sample period	Lags $q=2$	$q=4$	$q=8$	$q=16$
62:07:06—78:09:29	**1.43**	**1.93**	**2.46**	**2.77**
T=848	$(8.82)^*$	$(8.49)^*$	$(7.00)^*$	$(5.59)^*$
	$(8.82)^*$	$(10.81)^*$	$(11.00)^*$	$(9.33)^*$
	$(12.46)^*$	$(14.47)^*$	$(14.39)^*$	$(11.70)^*$
78:10:06—94:12:23	**1.43**	**1.98**	**2.65**	**3.19**
T=847	$(6.20)^*$	$(7.07)^*$	$(7.37)^*$	$(6.48)^*$
	$(6.20)^*$	$(8.62)^*$	$(10.69)^*$	$(10.70)^*$
	$(12.52)^*$	$(15.25)^*$	$(16.26)^*$	$(14.45)^*$
94:12:30—13:12:27	**1.21**	**1.47**	**1.7**	**1.82**
T=992	$(3.30)^*$	$(3.58)^*$	$(3.35)^*$	$(2.50)^*$
	$(3.30)^*$	$(4.13)^*$	$(4.15)^*$	$(3.44)^*$
	$(6.59)^*$	$(7.91)^*$	$(7.43)^*$	$(5.82)^*$

Table 1-B: Variance ratios for weekly medium-size portfolio returns

Sample period	Lags $q=2$	$q=4$	$q=8$	$q=16$
62:07:06—78:09:29	**1.25**	**1.54**	**1.79**	**1.91**
T=848	$(5.41)^*$	$(5.55)^*$	$(4.35)^*$	$(3.22)^*$
	$(5.41)^*$	$(6.41)^*$	$(5.93)^*$	$(4.69)^*$
	$(7.37)^*$	$(8.42)^*$	$(7.78)^*$	$(6.05)^*$
78:10:06—94:12:23	**1.20**	**1.37**	**1.54**	**1.56**
T=847	$(3.29)^*$	$(3.35)^*$	$(3.18)^*$	$(2.14)^*$
	$(3.29)^*$	$(3.72)^*$	$(3.90)^*$	$(2.93)^*$
	$(5.73)^*$	$(5.80)^*$	$(5.36)^*$	$(3.74)^*$
94:12:30—13:12:27	**0.99**	**1.05**	**1.02**	**0.89**
T=992	(-0.02)	(0.38)	(0.10)	(-0.38)
	(-0.02)	(0.43)	(0.11)	(-0.48)
	(-0.04)	(0.78)	(0.20)	(-0.78)

Table 1-C: Variance ratios for weekly large-size portfolio returns

Sample period	Lags $q=2$	$q=4$	$q=8$	$q=16$
62:07:06—78:09:29	**1.05**	**1.15**	**1.21**	**1.19**
T=848	(1.05)	(1.64)	(1.23)	(0.68)
	(1.05)	(1.54)	(1.32)	(0.84)
	(1.59)	(2.33)*	(2.06)*	(1.29)
78:10:06—94:12:23	**1.03**	**1.06**	**1.08**	**1.01**
T=847	(0.63)	(0.61)	(0.54)	(0.03)
	(0.63)	(0.65)	(0.59)	(0.04)
	(0.95)	(0.91)	(0.75)	(0.04)
94:12:30—13:12:27	**0.93**	**0.94**	**0.89**	**0.81**
T=992	(−0.99)	(−0.46)	(−0.53)	(−0.62)
	(−0.99)	(−0.52)	(−0.61)	(−0.77)
	(−2.05)*	(−1.01)	(−1.14)	(−1.35)

- The results for the earlier sample periods are broadly similar to those in Campbell, Lo and Mackinlay (1997, P71, Table 2.6).
  - The variance ratios are greater than one and deviate further from one as the horizon lengthens.
  - The departure from the random walk model is strongly statistically significant for the small and medium sized firms, but not so for the larger firms.
- When we turn to the later period 1994-2013 we see that the variance ratios all reduce.
  - For the smallest stocks the statistics are still significantly greater than one and increase with horizon. However, they are much closer to one at all horizons and the statistical significance of the departures is substantially reduced.
  - For medium sized firms, the variance ratios are reduced. They are in some cases below one and also no longer increasing with horizon. They are insignificantly different from one.
  - For the largest firms, the ratios are all below one but are statistically inseparable from this value.

- The test statistics change quite a lot depending on which variance estimator $\widehat{V}_{O1}(q)$, $\widehat{V}_{O2}(q)$ or $\widehat{V}_{O3}(q)$ one uses, and in some cases this could affect ones conclusions, for instance, for large-size portfolio, test statistics based on $\widehat{V}_{O1}(q)$ in some periods are statistically significant.
- One interpretation of these results is that the stock market (at the level of these portfolios) has become closer to the efficient benchmark. This is consistent with the evidence presented in Castura, Litzenberger, Gorelick, and Dwivedi (2010) for high frequency stock returns. The biggest improvements seem to come in the most recent period, especially for the small stocks.

Alternative Hypothesis: The Fads Model

Definition

Suppose log prices have a permanent/transitory decomposition:

$$p_t = p_t^* + u_t$$

$$p_t^* = \mu + p_{t-1}^* + \varepsilon_t, \varepsilon_t \sim IID(0, \sigma^2)$$

where u_t is a stationary process, while p_t^* is a random walk plus drift. p_t^* represents "fundamental value" and u_t represents "fads" or pricing errors

This says that markets are short term inefficient but in the long run the efficient price dominates andy pricing errors.

How could we detect this? It follows that

$$\Delta p_t = r_t = \overbrace{\mu + \varepsilon_t}^{\text{iid fundamental return}} + \overbrace{u_t - u_{t-1}}^{\text{mean zero stationary fad}}$$

Consider the variance ratio for horizon q of the observed return series r_t.

$$r_{t-q:t} = q\mu + \sum_{k=1}^{q} \varepsilon_{t-k} + u_t - u_{t-q}$$

$$\mathrm{var}(r_{t-q:t}) = q\mathrm{var}(\varepsilon_t) + \mathrm{var}(u_t - u_{t-q}).$$

If u_t is i.i.d., then

$$VR(q) = \frac{q\mathrm{var}(\varepsilon_t) + \mathrm{var}(u_t - u_{t-q})}{q\mathrm{var}(\varepsilon_t) + q\mathrm{var}(u_t - u_{t-1})} = \frac{q\mathrm{var}(\varepsilon_t) + 2\mathrm{var}(u_t)}{q\mathrm{var}(\varepsilon_t) + 2q\mathrm{var}(u_t)} < 1$$

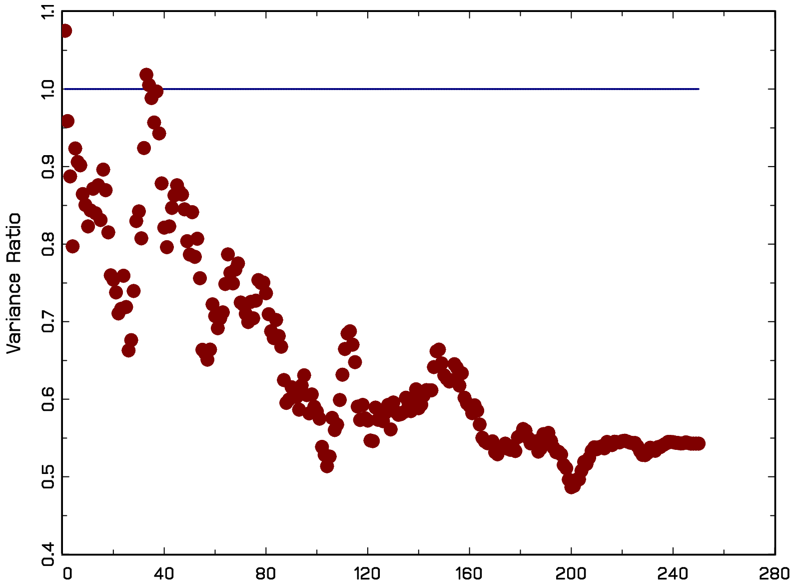
for any $q \geq 1$. However, if u_t is not iid, which is what we may anticipate, then $VR(q)$ could be larger or smaller than one.

Suppose only that the fad component is covariance stationary. Then, for a large enough q the variance ratio of observed returns r is less than one. In fact, as $q \to \infty$

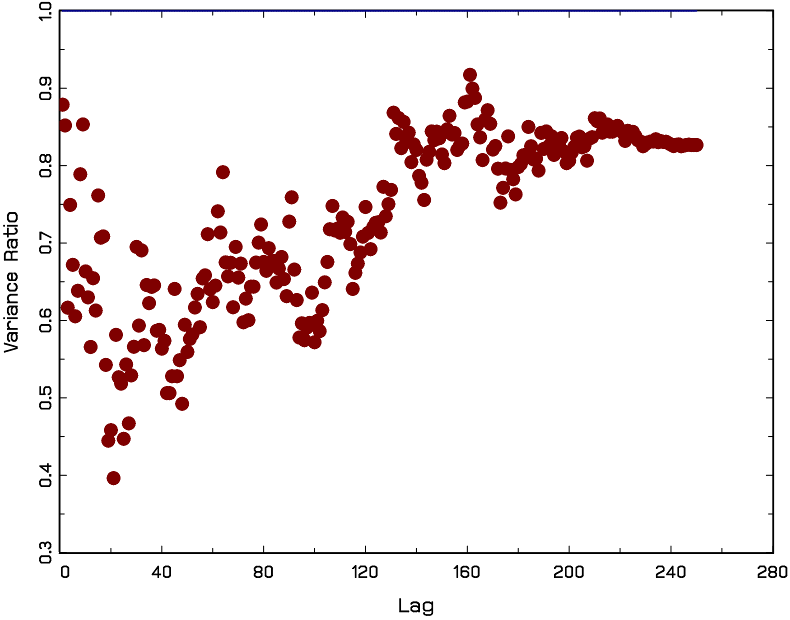
$$
\begin{aligned}
VR(q) &= \frac{q\mathrm{var}(\varepsilon_t) + \mathrm{var}(u_t - u_{t-q})}{q\mathrm{var}(\varepsilon_t) + q\mathrm{var}(u_t - u_{t-1})} \to \frac{\mathrm{var}(\varepsilon_t)}{\mathrm{var}(\varepsilon_t) + \mathrm{var}(u_t - u_{t-1})} \\
&= \frac{\mathrm{var}(\Delta p) - \mathrm{var}(\Delta u)}{\mathrm{var}(\Delta p)} = 1 - \frac{\mathrm{var}(\Delta u)}{\mathrm{var}(\Delta p)}
\end{aligned}
$$

This says that if the fads model is true we should find VR less than one for long lags.

Shows $VR(lag)$ for $lag = 2, \ldots, 250$. Variance ratio of SP500 Daily Returns, 1950-2017



Variance ratio of FTSE100 Daily Returns, 1984-2017



- The (weak) fads model has clear implications only for long horizons.
- Econometric testing of this model is problematic. CLM say that long horizon return tests often have low power and unreliable asymptotics (they explain based on $q/T \to (0, \infty)$).
- If $q/T \to 0$ fast enough then maybe ok. Newey-West standard errors. Large standard errors

Trading Strategy Based Evidence

We have so far emphasized statistical evidence and statistical criteria to judge the presence or absence of predictability. We now consider whether such predictability can yield a profit, and how large a profit it might yield.

- Lo and MacKinlay (1990). Consider a set of assets with returns $\{r_{it},\ i = 1, \ldots, n\}$. Define the equally weighted portfolio with return in period t given by $\overline{r}_t = \sum_{i=1}^n r_{it}/n$. Consider the following portfolio weights (operating from period $t-1$ to t) for any $j \geq 1$

$$w_{it}(j) = \frac{1}{n}\left(r_{i,t-j} - \overline{r}_{t-j}\right).$$

- This can be considered a momentum strategy: puts positive weight (buys) on winners and negative weight (sells) on losers as defined at time $t-j$. By construction the weights satisfy $\sum_{i=1}^n w_{it}(j) = 0$ so this is a zero net investment.

For simplicity suppose that r_{it} is stationary with mean zero and variance one. The expected profit of this strategy is

$$\begin{aligned}
\pi_+(j) &= \sum_{i=1}^{n} E[w_{it}(j) r_{it}] \\
&= \frac{1}{n}\sum_{i=1}^{n} E\left[\left(r_{it-j} - \overline{r}_{t-j}\right) r_{it}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} \left[\left(E(r_{it-j} r_{it}) - \frac{1}{n}\sum_{l=1}^{n} E(r_{l,t-j} r_{it})\right)\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} \rho_{ii}(j) - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{l=1}^{n} \rho_{il}(j),
\end{aligned}$$

where $\rho_{il}(j)$ is the cross-autocorrelation between asset i and l with lag j. Under the (multivariate) EMH hypothesis, $\pi_+(j) = 0$ for all j.

- There are famous several papers that use essentially this methodology
  - Jegadeesh and Titman (1993) who found short-term momentum, i.e., good and bad recent performance (3-12 months) continues over time (which is consistent with positive autocorrelation and not zero as with a random walk); they considered NYSE and AMEX over the period 1965 to 1989. They considered trading strategies that selected stocks based on their returns over the past J months and then evaluated their performance over a K month holding period.
  - De Bondt and Thaler (1985, 1987) suggest on the other hand that stock prices overreact to information over the longer term, suggesting that contrarian strategies (buying past losers and selling past winners) achieve abnormal returns. They consider monthly return data for New York Stock Exchange (NYSE) common stocks for the period between January 1926 and December 1982. They show that over 3- to 5-year holding periods stocks that performed poorly over the previous 3 to 5 years achieve higher returns than stocks that performed well over the same period.

Nonparametric Tests of RW1

- The ACF and Variance ratio tests generally require that

$$Er_t^4 \leq C < \infty,$$

  which may be problematic for daily data (e.g., 1987 crash).
- Specifically, sample correlations could be dominated by one or two observations.

- A second issue is that one cannot reject the null hypothesis of the absence of linear predictability implicit in the above methodology even when there does exist predictability.

Example

Suppose that $X = \cos\theta$ and $Y = \sin\theta$, where θ is uniform on $[0, 2\pi]$. Then $Y^2 = 1 - X^2$ so X and Y are functionally related and not independent. However,

$$\mathrm{cov}(X, Y) = \int_0^{2\pi} \cos(\theta)\sin(\theta)d\theta - \int_0^{2\pi} \cos(\theta)d\theta \times \int_0^{2\pi} \sin(\theta)d\theta = 0$$

so that the two random variables are uncorrelated. The point is that even if there is no linear predictability there may be nonlinear predictability.

- In this example $\mathrm{corr}(X^2, Y^2) = 1$, so this suggests looking at $\mathrm{cov}(r_t, g(r_{t-j}))$ for different functions g.

- We consider sign, runs, and Cowles and Jones (1937) tests of the efficient market hypothesis. These can detect nonlinear predictability but will typically require the stronger assumption of RW1 regarding independence but weaker assumptions on moments.
- Consider

$$\mathrm{sign}(X_t) = \begin{cases} 1 & \text{if } X_t > 0 \\ -1 & \text{if } X_t < 0 \end{cases}$$

  and zero else. Can think of this as "red" and "black" outcomes on roulette wheel.
- If returns are not predictable (and there is no drift or mean), we expect red and black to be equally likely and to not occur in long sequences of reds or blacks.

- Specifically, suppose that X_t is independent with **median** zero (RW2). Then

$$Y_t = \mathrm{sign}(X_t)$$

is i.i.d. Bernoulli with 50/50 chance of being ± 1 so $EY_t = 0$ and $\mathrm{var} Y_t = 1.$ Furthermore, $EY_t Y_{t-j} = 0$ for all $j.$
- Can apply statistical tests of iid null hypothesis based on the autocorrelation of signs of returns.
- This is what Cowles and Jones did essentially.

- Now suppose that X_t is independent and identically distributed but does not have **median** zero. Then

$$\begin{aligned} EY_t &= \Pr(X_t > 0) - \Pr(X_t < 0) \\ &= 1 - 2F_{X_t}(0) \neq 0. \end{aligned}$$

  Furthermore, $\mathrm{var}(Y_t) = 1 - E^2(Y_t) < 1.$

- Nevertheless, $\mathrm{cov}(Y_t, Y_{t-j}) = 0$ and we can test this implication. But we need to correct for the non-zero mean.

Cowles and Jones test

Define continuations and reversals

$$\begin{aligned} \textit{continuations}, \quad r_t \times r_{t+1} &> 0 \\ \textit{reversals}, \quad r_t \times r_{t+1} &< 0 \end{aligned}$$

we consider tests based on counting the relative frequency of continuations and reversals.

Example

Contingency Table (S&P500 data)

today/tomorrow	up	down
up	0.297	0.227
down	0.228	0.232

Continuations $0.297 + 0.232 = 0.53$ and Reversals $0.228 + 0.227 = 0.46$

Under the null hypothesis of independence and constant probability $\pi = \Pr \ (r_t > 0)$ we have

$$\begin{aligned}
\pi_c &= \Pr[\ r_t \times \ r_{t+1} > 0] \\
&= \Pr[\ r_t > 0, r_{t+1} > 0] + \Pr[\ r_t < 0,\ r_{t+1} < 0] \\
&= \Pr[\ r_t \ > 0] \Pr[r_{t+1} > 0] + \Pr[\ r_t < 0] \Pr[\ r_{t+1} < 0] \\
&= \pi^2 + (1-\pi)^2,
\end{aligned}$$

and

$$\frac{\Pr[\mathit{continuation}]}{\Pr[\mathit{reversal}]} = \frac{\Pr[\ r_t \times \ r_{t+1} > 0]}{1 - \Pr[\ r_t \ \times r_{t+1} > 0]} = \frac{\pi_c}{1-\pi_c}.$$

If $\pi = 1/2$ as CJ assumed (zero drift), $\pi_c/(1-\pi_c) = 1.$

Definition

Compare empirical continuations with reversals

$$T_c = \#\ of\ \ continuations, \quad r_t \ \times r_{t+1} > 0$$

$$T_r = \#\ of\ \ reversals, \quad r_t \ \times r_{t+1} \leq 0$$

$$\widehat{CJ} = \frac{T_c}{T_r} = \frac{T_c}{T - T_c} = \frac{\left(\frac{T_c}{T}\right)}{1 - \left(\frac{T_c}{T}\right)}$$

- The CJ statistic is a rational function of the sample mean of a binomial random variable. For large T, the sample mean of a binomial is approximately normal, using central limit theory.

Theorem

Provided $\pi = 1/2$, *we have*

$$\sqrt{T}\left(\widehat{CJ} - 1\right) \Longrightarrow N\left(0, 4\right)$$

$$\tau_{CJ} = \sqrt{T}\left(\frac{\widehat{CJ} - 1}{2}\right) \Longrightarrow N\left(0, 1\right)$$

Reject null hypothesis if $\left|\tau_{CJ}\right| > z_{\alpha/2}$.

Theorem

When $\pi \neq /12$, *we have*

$$\sqrt{T}\left(\widehat{CJ} - m(\pi)\right) \Longrightarrow N\left(0, V(\pi)\right)$$

$$m(\pi) = \frac{\pi_c}{1-\pi_c} \quad ; \quad V(\pi) = \frac{\pi_c\left(1-\pi_c\right) + 2\left(\pi^3 + (1-\pi)^3 - \pi_c^2\right)}{\left(1-\pi_c\right)^4}$$

where $\pi_c = \pi^2 + (1-\pi)^2$. *Let* $\widehat{\pi}$ *be the fraction of positive returns in sample and let*

$$\tau_{CJ} = \sqrt{T}\left(\frac{\widehat{CJ} - m(\widehat{\pi})}{\sqrt{V(\widehat{\pi})}}\right) \Longrightarrow N\left(0, 1\right)$$

Reject null hypothesis if $\left|\tau_{CJ}\right| > z_{\alpha/2}$.

- One can directly estimate π by counting the number of positive returns in the sample. Then let
- CJ statistic is non-parametric (no moments) but test does require the i.i.d assumption. CJ detects some departures from RW1 but has limited power. Similar to Kendalls Tau, which was introduced around the same time

An alternative approach (for the case when $\pi \neq 1/2$) is to redefine continuations and reversals

Definition

Let $\widehat{\mu}$ is the sample median of returns.

$$T_c = \#\ of\ continuations, \quad (r_t - \widehat{\mu}) \times (r_{t+1} - \widehat{\mu}) > 0$$

$$T_r = \#\ of\ reversals, \quad (r_t - \widehat{\mu}) \times (r_{t+1} - \widehat{\mu}) \leq 0$$

$$\widehat{CJ} = \frac{\widehat{T}_c}{\widehat{T}_r} = \frac{\left(\frac{\widehat{T}_c}{T}\right)}{1 - \left(\frac{\widehat{T}_c}{T}\right)}$$

- Under the null hypothesis

$$\sqrt{T}\left(\frac{\widehat{CJ} - 1}{2}\right) \Longrightarrow N(0,1)$$

CJ test for Dow stocks

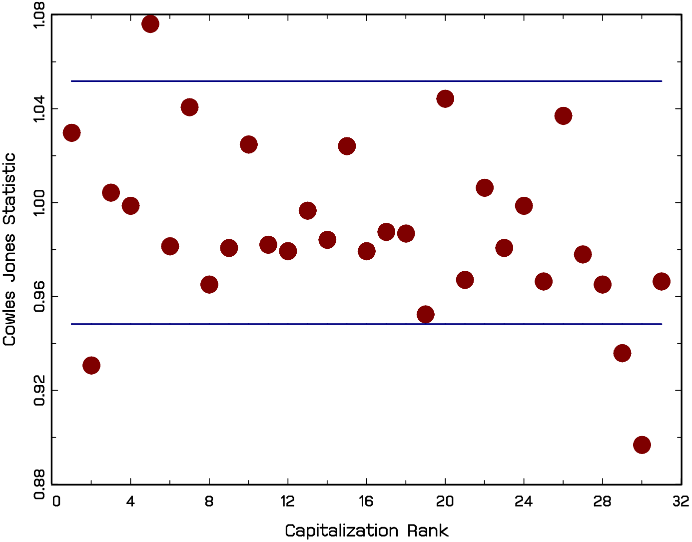


Figure: Cowles Jones statistic with standard errors

Runs Tests

- A "run" is a series of ups or downs. For example, $++--+++$ has three runs. CLM discusses tests based on the number of runs in sample. There is a formal (nonparametric) test of RW1 using this.
- Instead we ask: How many days in a row should we see up markets or down markets?
- This is similar to the question of how many reds or blacks in a row should we see at the roulette wheel. In August 18, 1913, the colour black came up 26 times in a row at the casino in Monte Carlo. In MC roulette, there are 18 reds, 18 blacks, and a zero, so the probability of making 26 blacks in a row ex ante is

$$\left(\frac{18}{37}\right)^{26} = 7.3087 \times 10^{-9}$$

which is low, but not impossible. Since this is the maximum observed run length out of many millions presumably of rolls (since 1796) should correct for this "selection bias" in the probability calculation.

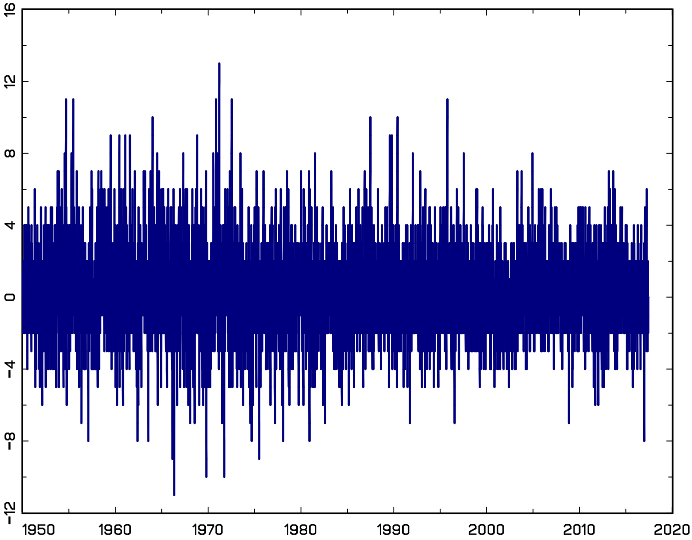


Figure: Length of daily runs on the S&P500

- As you can see from the plot, there are both positive and negative runs through the data.
  - The longest run is of duration 13 days on both the positive and negative side.
  - The mean of Z is 0.139 so that on balance we have more positive days than negative ones.
- In fact the distribution of the maximum run length is known exactly (under iid null hypothesis), and the number 10 is close to the centre of its theoretical distribution for the 5000 or so observations considered in this sample. No evidence against EMH from this.

Regression Based Tests (Semi Strong)

If we assume that stock returns are unforecastable (or not much) given past prices this does not preclude them being forecastable given additional information.

- Calendar effects. Suppose that

$$R_t = \mu + \beta^{\intercal} X_t + \varepsilon_t,$$

  where X_t is observed (public information) at time t deterministic seasonal dummy variables. The EMH (along with constant mean or risk premium) says that $\beta = 0$. Standard regression F test for the inclusion of X_t.

- Examples: Day of the week effect, month of the year effect, etc

Sample period	Mon	Tue	Wed	Thur	Fri
1950-1960	−0.00128	0.00002	0.00109	0.00093	0.00188
	(3.969)	(0.075)	(3.463)	(2.926)	(5.859)
1960-1970	−0.00158	0.00013	0.00099	0.00053	0.00089
	(5.496)	(0.465)	(3.464)	(1.862)	((3.121)
1970-1980	−0.00120	−0.00013	0.00064	0.00039	0.00068
	(3.095)	(0.356)	(1.694)	(1.011)	(1.787)
1980-1990	−0.00107	0.00113	0.00141	0.00027	0.00088
	(-2.194)	(2.384)	(2.991)	(0.556)	(1.823)
1990-2000	0.00116	0.00062	0.00088	−0.00029	0.00062
	(2.882)	(1.601)	(2.247)	(-0.748)	(1.554)
2000-2016	−0.00013	0.00056	0.00010	0.00053	−0.00028
	(0.296)	(1.331)	(0.231)	(1.228)	(0.655)

Sample period	Mon	Tue	Wed	Thur	μ
1950-1960	-0.00315	-0.00185	-0.00078	-0.00094	0.00188
	(6.946)	(4.098)	(1.743)	(2.089)	(5.859)
1960-1970	-0.00245	-0.00074	0.00013	-0.00034	0.00087
	(6.056)	(1.836)	(0.312)	(0.849)	(3.053)
1970-1980	-0.00188	-0.00082	$-3.24E-6$	-0.00029	0.00068
	(3.465)	(1.521)	(0.0061)	(0.529)	(1.788)
1980-1990	-0.00195	0.00022	0.00055	-0.00061	0.00088
	(2.844)	(0.325)	(0.808)	(0.900)	(1.823)
1990-2000	0.000571	$9.38E-6$	0.00026	-0.00091	0.00062
	(1.010)	(0.019)	(0.473)	(1.630)	(1.555)
2000-2010	0.00034	0.00089	0.00038	0.0010	-0.00054
	(0.384)	(1.011)	(0.430)	(1.129)	(0.859)
2010-2018	$-6.79E-5$	0.00062	0.00027	0.00024	0.00026
	(0.102)	(0.946)	(0.414)	(0.370)	(0.563)

- Consider the (predictive) regression

$$R_{t+j} = \mu + \beta^{\intercal} X_t + \varepsilon_{t+j},$$

  where X_t is observed (public information) at time t or deterministic like seasonal dummy variables. The EMH (along with constant mean or risk premium) says that $\beta = 0$. Standard regression F test for the inclusion of X_t.
- Examples: price/earnings ratio effects, dividend rate, and so on. Lots of evidence on this. Shiller website. Some econometric issues when X is very persistent process.

- Can also include nonlinear functions of observed variables to try to enhance predictability. More generally can fit nonlinear regression models.
- The following web site has a long list of examples of violations of EMH and explanations thereof http://www.behaviouralfinance.net/

Empirical Evidence Semi-strong Form

Many studies have identified the so-called 'anomalies' that seem difficult to reconcile with the EMH.

- Dividend/price ratio or price/earnings ratio seem to predict returns (return predictability),
- Small firms seem to have higher risk-adjusted returns than large firms (small firm effect, most of it in January),
- Firms with high book-to-market ratios seem to have higher returns even after controlling for risk (book-to-market effect)
- Calendar effects: Monday has negative returns up until 1990s

Some of the above may be reinterpreted as rational rewards for risk if the asset pricing model that helps us adjust for risk is misspecified (remember the Joint Hypothesis Problem). Other examples like Monday effect harder to fit into rational asset pricing paradigm.

Some Case Studies

Example

Massively Confused Investors Making Conspicuously Ignorant Choices. Rashes (JF, 2001). He examines the comovement of stocks with similar ticker symbols MCI (large telecom, Nasdaq) and MCIC (a closed end mutual fund, NYSE). Finds a significant correlation between returns, volume, and volatility at short frequencies. New information about MCI affects prices of MCIC and vice versa. Deviations from "fundamental value" tend to be reversed within several days, although there is some evidence that the return comovement persists for longer horizons. Arbitrageurs appear to be limited in their ability to eliminate these deviations from fundamentals.

Example

CUBA, Thaler (2016). It had around 70 percent of its holdings in US stocks with the rest in foreign stocks, but absolutely no exposure to Cuban securities, since it has been illegal for any US company to do business in Cuba since 1960. For the first few months of 2014 the share price was trading in the normal 10–15 percent discount range of the Net Asset Value (the value of the shares it itself held). Then on December 18, 2014, President Obama announced his intention to relax the United States' diplomatic relations with Cuba. The price of CUBA shares jumped to a 70 percent premium Although the value of the assets in the fund remained stable, the substantial premium lasted for several months, finally disappearing about a year later.

Example

More recently, a number of firms with names overlapping with Bitcoin but with no direct connection have experienced substantial price appreciation.

Anomaly Characteristics

- They are ‘small’. Small $ (e.g., MCI Jr. vs. MCI)
- Not scalable, e.g., illiquid
- Statistically suspect. Standard errors often based on iid. Even worse Data mining issues (**White’s Reality Check**)
- Fleeting, dont last long. E.g., the small stock premium, January effect, Monday effect. Heisenberg Principle of Finance/Goodhart’s Law (about policy instruments). Observing an anomaly brings about its extinction.
- Not realizable profit opportunities. Transaction costs: commissions and Bid/Ask spreads. Information costs, e.g., complex mortgage instruments

The degree of efficiency might be the relevant point for discussion. Comparison of inefficiency across markets or stocks or time.

Strong form efficiency

We don't expect markets to be strongly efficient

Old market makers had monopoly access to the order book, meaning they knew who wanted to trade and the prices they were willing to pay. Earned consistent returns (except in 1987 crash!).

Trading on inside information is regulated and limited in many countries.

Improper disclosure and misuse of information are kinds of insider dealing

Studies of trades by insiders (managers, etc. who have to report such trades to the Securities and Exchange Commission [SEC]) show they are able to make abnormal profit through their information.

Market abuse. Cornering the market for commodities eg silver. Painting the tape. Spoofing/Layering. Wash trades. Recent cases in US and UK.

Appendix

Standard errors for averages of individual stocks follows from similar arguments given for averages of autocorrelations. We have

$$\frac{1}{n}\sum_{i=1}^{n}\widehat{VR}_{Ai}(q) = 1 + 2\sum_{j=1}^{q-1}\left(1-\frac{j}{q}\right)\frac{1}{n}\sum_{i=1}^{n}\widehat{\rho}_i(j).$$

Under RW1 we have

$$\begin{aligned}
\text{var}\left[\frac{1}{n}\sum_{i=1}^{n}\widehat{VR}_{Ai}(q)\right] &= 4\sum_{j=1}^{q-1}\left(1-\frac{j}{q}\right)^2 \text{var}\left[\frac{1}{n}\sum_{i=1}^{n}\widehat{\rho}_i(j)\right] \\
&= 4\sum_{j=1}^{q-1}\left(1-\frac{j}{q}\right)^2 \frac{1}{n^2 T}\left[n + \sum_{i\neq j}\sum \omega_{ij}^2\right] \\
&\simeq \frac{\omega^2}{T}\frac{4}{6q}\left(2q^2 - 3q + 1\right).
\end{aligned}$$

For example with $q = 2$ this is approximately $1/2T$, which gives a standard error of approximately 0.017 for this case. See Table 2.7 of CLM