Financial Data Analysis

Forecasting Asset Returns and Market Efficiency (ME)

Roman Liesenfeld
(University of Kiel)

October 7, 2008
Contents

2 Forecasting Asset Returns and Market Efficiency (ME) 3
  2.1 Martingale .................................................. 3
  2.2 Random Walk (RW) Hypotheses .......................... 5
  2.3 Tests for Random-walk I .................................. 9
  2.4 Test for Random-walk III ................................. 20
  2.5 Tests for Long-Range Dependence ...................... 28
2 Forecasting Asset Returns and Market Efficiency (ME)

2.1 Martingale

- The possibility to forecast asset returns is closely related to the concept of “market efficiency”.

- The “classical” definition of market efficiency was introduced by Fama (1970):

  
  "A market in which prices always fully reflect available information is called [informationally] efficient."

- Implications:

  1) Prices react instantaneously to new information and the best prediction of tomorrow’s price, is today’s price.

  2) It is not possible to make profits by trading on the available information.

- Using different information sets, which should be reflected in prices, leads to different forms of ME:

  1) **Weak-form of ME**: Information set includes only past prices.

  2) **Semistrong form of ME**: Information set includes publicly available information.

  3) **Strong form of ME**: Information set includes all information, including private information.

- A stochastic specification which is often associated with the hypotheses of efficient market (in the weak form) is so called **martingale model**:

  A martingale is a stochastic process \( \{ p_t \} \) which satisfies the following conditions:

  \[
  \mathbb{E} [ p_{t+1} \mid p_t, p_{t-1}, \ldots ] = p_t
  \]

  or equivalently

  \[
  \mathbb{E} [ p_{t+1} - p_t \mid p_t, p_{t-1}, \ldots ] = 0.
  \]

  This implies:

  - The best forecast (in the sense of minimizing the MSE) of tomorrow’s price is today’s price.
- In other words, all information contained in past prices is instantly and fully reflected in asset’s current price.

- Formally, this implies the absence of any linear relationship between \( p_{t+1} - p_t \) and past prices: if \( g(\cdot) \) denotes an arbitrary function and \( I_t = \{ p_t, p_{t-1}, \ldots \} \), then we have

\[
\text{cov} \left( (p_{t+1} - p_t), g(I_t) \right) = \mathbb{E} \left[ (p_{t+1} - p_t) \cdot g(I_t) \right] = \mathbb{E} \left[ \mathbb{E} (p_{t+1} - p_t \mid I_t) \cdot g(I_t) \right] = 0.
\]

- A problem of associating a martingale with market efficiency is the fact, that the martingale ignores the trade-off between expected return and risk, which is supposed in modern financial economics:

According to the CAPM, e.g., we have

\[
\mathbb{E} [p_{t-1} - p_t \mid \cdot] = \text{linear function of the risk premium} > 0,
\]

where the risk premium is necessary to attract investors to bear the risk.

In particular, the expected return in a CAPM-model is:

\[
\mathbb{E} [r_{t+1} \mid I_t] = r_{f,t+1} + \beta (\mathbb{E} [r_{m,t+1} \mid I_t] - r_{f,t+1}),
\]

with:
- \( r_{f,t} \) = risk free rate
- \( r_{m,t} \) = return of the market portfolio
- \( \beta \) = beta factor
- \( (\cdot) \) = risk premium for bearing the systematic risk associated with the asset
- \( I_t \) = information set

Hence, the asset returns should be properly adjusted for risk before applying the martingale specification.

- Another statistical model associated with the hypothesis of efficient market is the random-walk (RW) model:

- closely related to the martingale;

- often used to test the implication of the EM-hypothesis that prices changes should not be predictable;

- will be discussed in more detail in the next section: 1) Three versions of the RW-hypothesis will be introduced; 2) tests will be developed for each of these versions.
2.2 Random Walk (RW) Hypotheses

**RW I: iid increments**

- The simplest version of a RW for a price process is given by

\[
\begin{align*}
    p_t &= \mu + p_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{iid } (0, \sigma^2) \\
    p_t - p_{t-1} &= \mu + \varepsilon_t,
\end{align*}
\]

where \( p_t \) is the log-price, \( \mu \) is the expected price change (drift), and \( \varepsilon_t \) is a White Noise process.

- The implications of this RW-model are:
  - The average increment in prices is given by \( \mu \), representing the deterministic time trend of the price process.
  - Increments are serially independent and identically distributed.
  - For \( \mu = 0 \), the price process is a martingale, since
    \[
    \mathbb{E} [p_t - p_{t-1} | I_{t-1}] = \mu.
    \]
    However, note that under the RW I increments/returns are not only uncorrelated with past prices (or functions of them) but also independent.

- To interpret the properties of a RW I, consider the first two moments of \( p_t \mid p_0 \):
  \[
  \begin{align*}
  \mathbb{E} [p_t \mid p_0] &= \mu t + p_0, \quad p_0 = \text{“starting value”} \\
  \text{var} [p_t \mid p_0] &= \sigma^2 t.
  \end{align*}
  \]
  Hence, mean and variance are increasing without bounds, reflecting the non-stationarity of the a RW. (This also holds for the two other forms of the RW below).

- The problem of the RW I applied to model prices over a longer time span: The assumption of iid increments is typically not plausible:
  - Structural breaks due to changes in the institutional and/or economic environment.
  - Volatility clustering. (see, Figure 1)
Hence, we would expect changes in the distributional behavior of prices and returns. This leads to the following weaker forms of the RW–hypotheses:
Figure 1: Time plots of monthly log returns for IBM stock and S&P 500 index from January 1926 to December 1999 (T02, p.115).
**RW II: independent increments**

- The RW II has the form

\[ p_t = \mu + p_{t-1} + \varepsilon_t; \quad \varepsilon_t = \text{serially independent process with zero mean} \]

- This form allows for different distributional forms for \( \varepsilon_t \) including heteroscedasticity.
- An even less restrictive and weaker form of the RW-hypothesis is given by:

**RW III: uncorrelated increments**

- The RW III has the form

\[ p_t = \mu + p_{t-1} + \varepsilon_t; \quad \mathbb{E} [\varepsilon_t] = 0, \quad \mathbb{E} [\varepsilon_t \varepsilon_s] = \begin{cases} \sigma_{\varepsilon}^2, & \text{if } s = t \\ 0, & \text{else.} \end{cases} \]

- The RW III represents the weakest form of the RW-hypothesis allowing for nonlinear serial dependencies such as serial correlation in \(|\varepsilon_t|\) or \(\varepsilon_t^2, \varepsilon_t^3\).
2.3 Tests for Random–walk I

- Based on the assumption of iid increments one has the choice between parametric and non-parametric procedures to test the RW.

Non-parametric tests are in general more robust than parametric test, but often more difficult to construct.

Here, I’ll discuss two non-parametric tests, the Cowles–Jones-test and the runs-test.

### Cowles–Jones (CJ)–Test

- The basic idea is to compare the observed number of sequences and reversals in the time series of prices with the expected number of reversals and sequences under a RW I.

- A **sequence** is defined as two consecutive returns with the same sign, i.e.:
  \[ \{r_t, r_{t+1}\} > 0 \text{ or } \{r_t, r_{t+1}\} < 0. \]

- A **reversal** is defined as two consecutive returns with opposite signs
  \[ \{r_t > 0, r_{t+1} < 0\} \text{ or } \{r_t < 0, r_{t+1} > 0\}. \]

- Construction of the CJ–test statistic:
  - The model to be tested, is the RW I:
    \[
    r_t = p_t - p_{t-1} = \mu + \varepsilon_t; \quad \varepsilon_t \sim \text{iid } (0, \sigma^2),
    \]
    with a symmetric pdf for \( \varepsilon_t \).

  - Define the indicator (representing the sign of the return):
    \[
    I_t = \begin{cases} 
    1, & \text{if } r_t \geq 0 \\ 
    0, & \text{else} 
    \end{cases}
    \]
    and note that under the RW I, \( I_t \) is a iid Bernoulli random variable with
    \[
    \pi \equiv P(I_t = 1) = P(r_t \geq 0) \\
    1 - \pi = P(I_t = 0) = P(r_t < 0).
    \]

    Furthermore, note that under the RW I, the probability \( \pi \) is given by:
    \[
    \pi \begin{cases} 
    > 0.5, & \text{if } \mu > 0 \\
    = 0.5, & \text{if } \mu = 0 \\
    < 0.5, & \text{if } \mu < 0
    \end{cases}
    \]

  - Now, the number of sequences \( N_s \) in a time series with \( n + 1 \) return observations \( \{r_t\}_{t=1}^{n+1} \) can be obtained as
    \[
    N_s = \sum_{t=1}^{n} Y_t \text{ with } Y_t = I_t I_{t+1} + (1 - I_t) (1 - I_{t+1}),
    \]

9
where
\[ Y_t = \begin{cases} 
1, & \text{if } I_t = I_{t+1} = 1 \text{ or } I_t = I_{t+1} = 0 \quad \text{(sequence)} \\
0, & \text{if } I_t = 1, I_{t+1} = 0 \text{ or } I_t = 0, I_{t+1} = 1 \quad \text{(reversal)}
\end{cases} \]

- Note, that the indicator \( Y_t \) (indicating a sequence) is a Bernoulli random variable with
\[ \pi_s \equiv P (Y_t = 1) = P (I_{t+1} = 1, I_t = 1) + P (I_{t+1} = 0, I_t = 0). \]

Hence, under the RW I which implies serially independent \( r_t \)'s and \( I_t \)'s, this probability for a sequence has the form:
\[ \pi_s = P (I_{t+1} = 1) \cdot P (I_t = 1) + P (I_{t+1} = 0) \cdot P (I_t = 0) \]
\[ = \pi^2 + (1 - \pi)^2; \]

and for a reversal:
\[ 1 - \pi_s = 2\pi(1 - \pi). \]

- Furthermore observe that
\[ \pi_s \begin{cases} 
= 0.5, & \text{if } \mu = 0 \\
> 0.5, & \text{if } \mu \neq 0.
\end{cases} \]

This reflects the fact that under a RW without a drift, sequences and reversals are equally likely, and under a RW with a drift, sequences are more likely than reversals (due to the deterministic trend).

- The CJ-Statistic is given by
\[ \hat{CJ} = \frac{\# \text{ sequences}}{\# \text{ reversals}} = \frac{N_s}{n - N_s} = \frac{1}{n} \sum_{i=1}^{n} Y_i - \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} Y_i Y_j. \]

In order to use \( \hat{CJ} \) for a test of the RW I, it is necessary to know the properties of \( \hat{CJ} \) under the RW I:

- According to the weak law of large numbers (WLLN) the probability limit (for \( n \to \infty \)) of \( \frac{1}{n} \sum_{t=1}^{n} Y_t \) is given by
\[ \text{plim} \frac{1}{n} \sum_{t=1}^{n} Y_t = E [Y_t] = 1 \cdot \pi_s + 0 \cdot (1 - \pi_s) = \pi_s = \pi^2 + (1 - \pi)^2. \]

Hence, the CJ-statistic has the following probability limit:
\[ \text{plim} \hat{CJ} = \frac{\pi_s}{1 - \pi_s} = \frac{\pi^2 + (1 - \pi)^2}{2\pi (1 - \pi)}. \]

This implies that under a RW without drift, one would expect
\[ \hat{CJ} \approx 1, \quad \text{since} \quad \pi_s = 0.5. \]
and under a RW with drift
\[ \widehat{CJ} > 1, \text{ since } \pi_s > 0.5. \]

- In order to construct a formal test for the RW I-hypothesis based upon the CJ-statistic the sampling distribution of \( \widehat{CJ} \) must be known.

It can be shown that the asymptotic distribution for \( \widehat{CJ} \) is given by:
\[ \widehat{CJ} \sim N \left( \frac{\pi_s}{1 - \pi_s}, \text{asy.var} \left( \widehat{CJ} \right) \right) \]

with
\[ \text{asy.var} \left( \widehat{CJ} \right) = \frac{\pi_s (1 - \pi_s) + 2 \left[ \pi^3 + (1 - \pi)^3 - \pi_s^2 \right]}{n (1 - \pi_s)^4}. \]

The unknown parameters can be estimated consistently by
\[ \hat{\pi} = \frac{1}{n} \sum_{t=1}^{n} I_t \quad \text{and} \quad \hat{\pi}_s = \hat{\pi}^2 + (1 - \hat{\pi})^2. \]

- This asymptotic distribution of the CJ-statistic is obtained as follows:
  
  - First, observe that under the RW I the sequence of indicators \( \{Y_t\} \) with
    \[ Y_t = I_t I_{t-1} + (1 - I_t) (1 - I_{t-1}) \]
  
  constitute a homogeneous Markov Chain with states \( Y_t = 1 \) and \( Y_t = 0 \).

  - In particular, under the RW I we have
    \[ Y_t, Y_{t-1} : \text{are stochastically dependent} \]
    \[ Y_t, Y_{t-k} : k > 1 \text{ are stochastically independent} \]

  with the following homogeneous transition probabilities:
  \[ P_t \left( Y_t = y_t \mid Y_{t-1} = y_{t-1}, Y_{t-2} = y_{t-2}, \cdots \right) = P \left( Y_t = y_t \mid Y_{t-1} = y_{t-1} \right). \]

  Note, that the \( Y_t \)'s are neither iid nor serially uncorrelated.

  - To establish an asymptotic distribution for \( \frac{1}{n} \sum_t Y_t \) (and hence for \( \widehat{CJ} \)), one can use the Central Limit Theorem for homogeneous Markov Chains provided by Fish (1976)\footnote{see, Fish, M. Wahrscheinlichkeitsrechnung und mathematische Statistik, VEB Deutscher Verlag der Wissenschaften, Berlin, p. 314}.
Theorem: Let \( \{ x_n \} \) be a sequence of random variables constituting a homogeneous Markov Chain with a finite number of states and \( \mathbb{E}[x_j] = \mu < \infty \forall i \). Let the states be irreducible with transition probabilities being strictly positive, and

\[
\lim_{n \to \infty} \frac{1}{n} \text{var} \left[ \sum_{i=1}^{n} x_i \right] = \sigma^2 > 0.
\]

Then

\[
\sqrt{n} \left( \frac{1}{n} \sum_{i=1}^{n} x_i - \mu \right) \xrightarrow{d} N \left( 0, \sigma^2 \right).
\]

- In order to apply this result to \( \frac{1}{n} \sum_t Y_t \), we need to calculate \( \sigma_y = \lim_{n \to \infty} \frac{1}{n} \text{var} \left[ \sum_t Y_t \right] \):

First, note that

\[
\text{var} \left[ \sum_t Y_t \right] = \sum_t \text{var} [Y_t] + 2 \sum_{t>s} \text{cov} (Y_t, Y_s),
\]

with

\[
\text{var} (Y_t) = 1^2 P (Y_t = 1) + 0^2 P (Y_t = 0) - [1P (y_t = 1) - 0P (Y_t = 0)]^2
\]

\[= \pi_s - \pi_s^2
\]

\[= (1 - \pi_s) \pi_s,
\]

\[
\text{cov} (Y_t, Y_{t-j}) = 0 \quad \forall \ |j| > 1,
\]

\[
\text{cov} (Y_t, Y_{t-1}) = \sum_{y_t, y_{t-1}} Y_t Y_{t-1} P (Y_t = y_t, Y_{t-1} = y_{t-1}) - \mathbb{E} [Y_t] \mathbb{E} [Y_{t-1}]
\]

\[= 1 \cdot 1 \cdot P (Y_t = 1; Y_{t-1} = 1) - \pi_s^2
\]

\[= \pi^3 + (1 - \pi)^3 - \pi_s^2.
\]

Hence, we obtain:

\[
\text{var} \left[ \sum_t Y_t \right] = n \cdot (1 - \pi_s) \pi_s + 2 (n - 1) \left[ \pi^3 + (1 - \pi)^3 - \pi_s^2 \right]
\]

and

\[
\sigma_y^2 = \lim_{n \to \infty} \frac{1}{n} \text{var} \left[ \sum_t Y_t \right] = (1 - \pi_s) \pi_s + 2 \left[ \pi^3 - (1 - \pi)^3 - \pi_s^2 \right].
\]

This yields according to the Theorem given above the following limiting distribution for stabilizing transformation of \( \frac{1}{n} \sum_t Y_t \):

\[
\sqrt{n} \left( \frac{1}{n} \sum_t Y_t - \pi_s \right) \xrightarrow{d} N \left( 0; \sigma_y^2 \right).
\]
• Now, note that the CJ-statistic is a continuous function in $Z = \frac{1}{n} \sum_i Y_i$:

$$\bar{CJ} = f(Z) = \frac{Z}{1 - Z}, \quad \text{with} \quad \frac{df(\cdot)}{dZ} = \frac{1}{(1 - Z)^2}.$$  

Applying a first-order Taylor expansion to $\bar{CJ}$ and using the limit distribution for $\sqrt{n} (Z - \pi_s)$ then yields:

$$\sqrt{n} \left( \bar{CJ} - f(\pi_s) \right) \xrightarrow{d} N \left(0, \left(\frac{df(\pi_s)}{dZ}\right)^2 \sigma_Y^2 \right)$$

and hence

$$\sqrt{n} \left( \bar{CJ} - \frac{\pi_s}{1 - \pi_s} \right) \xrightarrow{d} N \left(0, \frac{1}{(1 - \pi_s)^2} \left[1 - \pi_s + 2\left(\pi^3 + 1 - \pi - \pi_s^2\right)\right]\right),$$

- Based on this asymptotic result, one can construct the following $t$-statistic:

$$t_{\bar{CJ}} = \frac{\bar{CJ} - \hat{\pi}_s / (1 - \hat{\pi}_s)}{\left(\text{asy.var } \bar{CJ}\right)^{1/2}},$$

which is under the RW I approximately $N(0,1)$-distributed. Based on this statistic the RW I is rejected at significance level $\alpha$, if

$$|t_{\bar{CJ}}| > Z_{1-\alpha/2} \quad \text{(two-sided test)},$$

where $Z_{1-\alpha/2}$ is the $(1 - \alpha/2)$-quantile of a $N(0,1)$-distribution.

- Figure 2 shows the daily closing prices of the Daimler-Benz stock and the german stock index DAX, and the daily spot exchange rates for the US-Dollar/Deutsche Mark together with the corresponding CJ-statistic.

- As to the quality of this test, one can show, that it can have a rather low power w.r.t. certain alternatives.

In particular, CLM97 construct a two-state Markov Chain for $I_t$ implying serially correlated $I_t$’s and $r_t$’s, (and hence violating the RW I hypothesis) such that $\bar{CJ}$ converges to the same probability limit as under a RW I. Hence, the test based on $\bar{CJ}$ can have the tendency to accept the RW I model, even if it is false, too often.

- Also, note that the test is based upon the assumption of symmetric return distributions, which might be not appropriate.

\[\text{see, e.g., M96, Theorem 5.39, p. 287}\]
Figure 2. The Cointegration statistic for the return series of the Daimler-Benz stock, the German stock index DAX, and the spot exchange rate for the US-Dollar Deutsche Mark.

daily closing prices: Daimler-Benz AG

CJ = 1.12
t_cj = 1.94
$p_i_s/(1-p_i_s) = 1.002$

daily closing prices DAX

CJ = 0.99
t_cj = -0.16
$p_i_s/(1-p_i_s) = 1.006$

daily exchange rates: $/DM

CJ = 1.02
t_cj = 0.27
$p_i_s/(1-p_i_s) = 1.003$
- Another test for RW I is the (Wald–Wolfowitz) runs test. This test compares the number of runs which is expected under the RW I with the observed number of runs.

- A runs is defined as a sequence of consecutive realized values with the same sign.

- Consider, for example, the following sequence of returns \( \{r_t\} \):

\[
\begin{array}{c|cccccccc}
 t & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
\hline
 r_t & 2 & -1 & -2 & 1 & 2 & 3 & -2 & 3 & 2 \\
 sign & + & - & - & + & + & + & - & + & + \\
 run no. & 1 & 2 & 3 & 4 & 5 & & & & \\
\end{array}
\]

Hence, this particular series contains 5 runs.

- The runs test is a test which can be used to test the iid assumption. Note that under the RW I

\[
r_t = p_t - p_{t-1} = \mu + \varepsilon_t, \quad \varepsilon_t \sim \text{iid } (0, \sigma^2) \]

we have

\[
r_t \sim \text{iid } (\mu, \sigma^2).
\]

Hence, the runs test testing the iid property of the returns can be used as a test for RW I.

- The basic idea of the runs test of the iid property is that if the random variables under consideration are truly iid, then neither too few nor too many runs should be observed in the sample:
  
  - Too few runs could indicate clustering or trending.
  - Too many runs could indicate a systematic alternating pattern.

- To perform the runs test, we require the sampling distribution of the total number of runs under the iid property of returns:

Let

\[
W = \text{number of runs observed in the sequence } \{r_t\}_{t=1}^n, \\
n_1 = \text{number of positive returns} \\
n_0 = (n - n_1) = \text{number of negative returns}.
\]

Then, it can be shown that, for given \( n_0, n_1 \), the sampling distribution of \( W \) under

\footnote{The essential sampling distributions, which are necessary for the construction of runs test were derived by Wald and Wolfowitz (1940) (Annals of Mathematical Statistics, vol. 11.)}
iid returns \( r_t \) is given by\(^4\):

\[
P(W = w \mid n_1, n_0) = \begin{cases} 
2 \left( \frac{n_1-1}{w/2-1} \right) \left( \frac{n_0-1}{w/2-1} \right), & \text{for } w \text{ even} \\
\binom{n}{n_1} \left( \frac{n_0-1}{(w-1)/2} \right) + \binom{n}{n_1} \left( \frac{n_0-1}{(w-3)/2} \right), & \text{for } w \text{ odd}
\end{cases}
\]

with mean and variance:

\[
E[W] = \frac{2n_1n_0}{n} + 1
\]

\[
\text{var}[W] = \frac{2n_0n_1(2n_0n_1 - n)}{n^2(n-1)}.
\]

This sampling distribution is obtained using combinatorial arguments.

- The computations involved in calculating \( P(W = w \mid n_1, n_0) \) can become quite tedious. Fortunately, the distribution of runs \( W \) is approximately Normal for large samples\(^5\). In particular, for \( n_0 \to \infty \) and \( n_1 \to \infty \), it follows that under \( H_0 : r_t \sim \text{iid} \)

\[
Z = \frac{W - E[W]}{\sqrt{\text{var}[W]}} = \frac{W - \frac{2n_1n_0}{n} - 1}{\sqrt{\frac{2n_0n_1(2n_0n_1 - n)}{n^2(n-1)}}} \xrightarrow{d} N(0,1).
\]

- Hence, based on the runs test statistic \( Z \), the RW I hypothesis is rejected at a significance level \( \alpha \), if

\[
|Z| \geq Z_{1-\alpha/2} \quad \text{(two-sided test)},
\]

where \( Z_{1-\alpha/2} \) is the \((1 - \alpha/2)\)-quantile of a \( N(0,1) \)-distribution.

- A slight adjustment to \( Z \) is often made to account for the fact that while a Normal random variable is continuous, \( W \) is a discrete random variable:

\[
Z = \frac{W - E[W] + \frac{1}{2}}{\sqrt{\text{var}[W]}} \quad \text{(continuity correction)}.
\]

- Figure 3 shows the daily closing prices of the Daimler-Benz stock and the German stock index DAX, and the daily spot exchange rates for the US-Dollar/Deutsche Mark together with the corresponding run statistic.

- In order to construct runs, we have transformed the returns into a dichotomous variable using two types of returns: positive and negative,

\[
I_t = \begin{cases} 
0 & \text{if } r_t \leq 0 \\
1 & \text{if } r_t > 0
\end{cases}.
\]

\(^4\)See, e.g., M96, p. 663–666

This approach can be generalized by using more than 2 classes to transform the variable \( r_t \), e.g.,

\[
I_t^* = \begin{cases} 
1 & \text{if } c_0 < r_t \leq c_1 \\
2 & \text{if } c_1 < r_t \leq c_2 \\
& \vdots \\
1 & \text{if } c_{q-1} < r_t \leq c_q 
\end{cases}
\]

with the number of observations of type \( i \)

\[
n_i = \sum I_{(c_{i-1},c_i]}(r_t), \quad i : 1 \rightarrow q
\]

and

\[
w_i = \text{number of runs of type } i \\
w = \sum_{i=1}^{q} w_i = \text{total number of runs}
\]

For further details, see CLM97.
Figure 3: The run statistic for the return series of the Daimler-Benz stock, the German stock index DAX, and the spot exchange rate for the US-Dollar/Deutsche Mark.
- As to the quality of this test:
  - The runs test is independent of distributional assumptions. In particular, it does not require symmetric distributions for \( r_t \), which is a requirement for the CJ-test.
  - A disadvantage of the runs test is that it exhibits low power against higher serial dependence (e.g. serial correlation in \( r_t^2 \) or \(|r_t|\)).

**Tests for Random-walk II**

The assumption of the RW I of identical return distribution is clearly implausible, especially, when applied to data that span a long horizon (see, e.g., the volatility clustering in Figure 9).

Hence, a RW I with a serially independent error process and a varying distributional forms for \( \varepsilon_t \) might be more adequate.

However, testing for independence without assuming identical distributions is quite difficult in a time series context; and if we place no restriction on how the distribution of \( \varepsilon_t \) varies over time, it is completely impossible to conduct a test for independence.

Hence, we immediately turn to tests for the RW III.
2.4 Test for Random–walk III

**Autocorrelation tests**

The most direct test of the RW–hypothesis is to check for serial correlation in the return series.

- Under the RW III:

\[ r_t = p_t - p_{t-1} = \mu + \varepsilon_t; \quad \mathbb{E} [\varepsilon_t] = 0, \quad \mathbb{E} [\varepsilon_t \varepsilon_s] = \begin{cases} \sigma^2, & \text{if } s = t \\ 0, & \text{else} \end{cases} \]

(and under RW I and RW II) the returns \( r_t \) should be serially uncorrelated at all lags. Hence, the RW III can be tested by testing

\[ H_0 : \text{“} r_t \text{’s are serially uncorrelated”, i.e., } H_0 : \rho_\ell = 0, \forall \ell, \]

where \( \rho_\ell \) is the lag-\( \ell \) autocorrelation for \( r_t \).

- In order to use the sample autocorrelation

\[ \hat{\rho}_\ell = \frac{\hat{\gamma}_\ell}{\hat{\gamma}_0}, \quad \hat{\gamma}_\ell = \text{lag-\( \ell \) sample autocovariance for } \{r_t\} \]

\to test \( H_0 : \rho_\ell = 0 \), the sampling distribution of \( \hat{\rho}_\ell \) must be known:

- It can be shown that if \( \{r_t\}_{t=1}^T \) is a sequence of iid variables with finite variance \( \sigma^2 \) and \( \mathbb{E} [r_0^2] \propto \sigma^6 \), then \(^6\):

\[ \mathbb{E} [\hat{\rho}_\ell] = -\frac{T - \ell}{T(T - 1)} + O \left( \frac{T^{-2}}{T} \right). \]

(The expression \( O(T^{-2}) \) represents a term which when divided by \( T^{-2} \) converges to a constant as \( T \to \infty \).) Hence, under the RW I, the sample autocorrelation is a biased estimate for \( \rho_\ell = 0 \).

However, this bias vanishes for \( T \to \infty \).

- Furthermore, it can be shown that under the RW I with uniformly bounded sixth moment:

\[ \sqrt{T} \hat{\rho}_\ell \overset{\alpha}{\sim} N (0, 1). \]

• This last result can be used to perform a test of \( H_0 : \rho_\ell = 0 \) and hence of the RW I and RW III hypotheses. \( H_0 \) is rejected at a significance level \( \alpha \) if

\[
\sqrt{T} |\hat{\rho}_\ell| > Z_{1-\alpha/2} \quad \text{(two-sided test),}
\]

where \( Z_{1-\alpha/2} \) is the \((1 - \alpha /2)\) quantile of a \( N(0, 1)\)-distribution.

• Note that the RW III implies that all \( \rho_\ell \)’s are zero.
A test for the joint hypothesis

\[
H_0 : \rho_1 = \rho_2 = \cdots = \rho_m = 0
\]

is the Box–Pierce test which is based upon the previous results. The test statistic is given by

\[
Q_m = T \sum_{\ell=1}^{m} \hat{\rho}_\ell^2 = \sum_{\ell=1}^{m} \left( \sqrt{T} \hat{\rho}_\ell \right)^2 .
\]

Under the RW I-hypothesis

\[
Q_m \overset{a}{\sim} \chi^2(m),
\]

which results from the fact that \( Q_m \) a sum of squared random variables which are asymptotically \( N(0, 1)\)-distributed.

• In practice, the selection of \( m \) may affect the performance of the test:

- If \( m \) is too small: the presence of higher order autocorrelation may be missed;
- If \( m \) is too large: the power may be low due to many insignificant higher-order autocorrelation.

• To increase the power of the test in finite samples, \( Q_m \) can be modified as follows (Ljung–Box statistic):

\[
Q'_m = T(T + 2) \sum_{\ell=1}^{m} \frac{\hat{\rho}_\ell^2}{T - \ell} .
\]

Note, that this modification has no impact on the asymptotic properties as \((T + 2)/(T - \ell) \rightarrow 1\) for \( T \rightarrow \infty\).

- Figure 4 shows the daily closing prices of the Daimler-Benz stock and the german stock index DAX, and the daily spot exchange rates for the US-Dollar/Deutsche Mark together with the corresponding Ljung-Box statistic.

- One issue which requires a careful interpretation of test results is:

the asymptotic distribution of \( \hat{\rho}_\ell, Q_m, Q'_m \) are obtained under the assumption that the \( \nu_i \)'s are iid

However, suppose that \( t_i \) is uncorrelated but not serially independent (due to, e.g., volatility clustering effects) the tests of \( H_0 : \nu_i \) is serially uncorrelated” discussed above might lead to biased results.
Figure 4: The ACF and Ljung-Box statistic for the return series of the Daimler-Benz stock, the German stock index DAX, and the spot exchange rate for the US-Dollar/Deutsche Mark.
**Variance–ratio test**

- The variance–ratio (VR) test exploits the fact that under the RW I-III the variance of $k$-period log returns must be linear in the number of periods $k$.

- In particular, for $k = 2$ we obtain:

$$\text{var} \left( r_t [2] \right) = \text{var} \left( r_t + r_{t-1} \right) = \text{var} \left( r_t \right) + \text{var} \left( r_{t-1} \right) + 2 \text{cov} \left( r_t, r_{t-1} \right).$$

Hence, under a RW for log-prices, we have

$$\text{var} \left( r_t [2] \right) = 2 \cdot \text{var} \left[ r_t \right].$$

- Now, in order to check the RW-hypothesis, one can use the variance–ratio:

$$VR = \frac{\text{var} \left( r_t [2] \right)}{2 \cdot \text{var} \left[ r_t \right]},$$

which, under stationarity, can be expressed as:

$$VR = \frac{2 \text{var} \left[ r_t \right] + 2 \text{cov} \left[ r_t, r_{t-1} \right]}{2 \text{var} \left[ r_t \right]} = 1 + \rho_1.$$

Note in particular, that under a RW with $\rho_k = 0$, $k \geq 1$, the variance-ratio is:

$$VR = 1.$$

- In order to construct a statistical test of the RW-hypothesis based on VR, we need the corresponding sampling distribution of VR, which will be derived in the following:

- Let’s start with the Null under which the distribution of VR is derived:\footnote{The normality assumption is imposed just for expositional convenience.}

$$H_0 : r_t = \mu + \varepsilon_t, \quad \varepsilon_t \sim \text{iidN} \left( 0, \sigma^2 \right),$$

implying that $\text{var} \left[ r_t \right] = \sigma^2$, $\text{var} \left( r_t [2] \right) = 2\sigma^2$.

- Furthermore, let the data be

$$\{ p_k \}_{k=0}^{2n} : \text{ time series with } 2n + 1 \text{ log-prices.}$$
- These data can be used to estimate

\[ \text{E} [r_l] = \mu \quad \text{and} \quad \text{var} [r_l] = \sigma^2 \]

as follows:

\[ \hat{\mu} = \frac{1}{2n} \sum_{k=1}^{2n} (p_k - p_{k-1}) \]

\[ \hat{\sigma}_a^2 = \text{var} [r_l] = \frac{1}{2n} \sum_{k=1}^{2n} (p_k - p_{k-1} - \hat{\mu})^2 \]

\[ \hat{\sigma}_b^2 = \frac{1}{2} \text{var} (r_l [2]) = \frac{1}{n} \sum_{k=1}^{n} (p_{2k} - p_{2k-2} - 2\hat{\mu})^2 \]

Note that

\[ \hat{\mu}, \hat{\sigma}_a^2: \quad \text{conventional ML estimators} \]

\[ \hat{\sigma}_b^2: \quad \text{estimator, which makes use of the fact that under the Null} \quad \frac{1}{2} \text{var} (r_l [2]) = \sigma^2. \]

- According to standard asymptotic theory, these estimators exhibit under \( H_0 \) the following properties\(^8\):

\[ \text{plim} \hat{\sigma}_a^2 = \sigma^2; \quad \text{plim} \hat{\sigma}_b^2 = \sigma^2 \quad \text{(consistency)} \]

and:

\[ \sqrt{2n} (\hat{\sigma}_a^2 - \sigma^2) \overset{d}{\sim} \mathcal{N} (0, 2\sigma^4) \]

\[ \sqrt{2n} (\hat{\sigma}_b^2 - \sigma^2) \overset{d}{\sim} \mathcal{N} (0, 4\sigma^4) \quad \text{(asymptotic normality)} \]

Furthermore, note that \( \hat{\sigma}_a^2 \), which is the ML-estimator based on the complete data set, is asymptotically efficient (i.e., it exhibits an asymptotic normal distribution with the smallest possible variance, see, e.g., M96.).

Accordingly:

\[ \text{asy var} [\hat{\sigma}_a^2] = 2\sigma^4 < \text{asy var} [\hat{\sigma}_b^2] = 4\sigma^4. \]

- Also note that under the Null of a RW the distance between the two variance estimates \( \hat{\sigma}_a^2 \) and \( \hat{\sigma}_b^2 \) should be small.

- In particular, since \( \hat{\sigma}_b^2 \) and \( \hat{\sigma}_a^2 \) are under the Null asymptotically normal, the same holds for the distance:

\[ \sqrt{2n} ([\hat{\sigma}_b^2 - \hat{\sigma}_a^2] - 0) \overset{d}{\sim} \mathcal{N} (0, 2\sigma^4), \]

where the asymptotic variance \( \text{asy var} [\hat{\sigma}_b^2 - \hat{\sigma}_a^2] = 2\sigma^4 \) is obtained as follows\(^9\).

\(^8\)See, e.g., M96.

\(^9\)See, CLM97, p. 51
• First note that
\[
\text{asy.var} (\hat{\sigma}_b^2) = \text{asy.var} (\hat{\sigma}_a^2 + \hat{\sigma}_b^2 - \hat{\sigma}_a^2) = \text{asy.var} (\hat{\sigma}_a^2) + \text{asy.var} (\hat{\sigma}_b^2 - \hat{\sigma}_a^2) + 2 \text{asy.cov} \left( [\hat{\sigma}_b^2 - \hat{\sigma}_a^2], \hat{\sigma}_a^2 \right).
\]

• Now, observe that the covariance between \(\hat{\sigma}_a^2\) and \(\hat{\sigma}_b^2 - \hat{\sigma}_a^2\) must be zero, since \(\hat{\sigma}_a^2\) is asymptotically efficient (otherwise it would be possible to construct a linear combination of \(\hat{\sigma}_a^2\) and \(\hat{\sigma}_b^2 - \hat{\sigma}_a^2\) with lower asymptotic variance than \(\hat{\sigma}_a^2\), which would be a contradiction).

Hence, we obtain
\[
\text{asy.var} (\hat{\sigma}_b^2 - \hat{\sigma}_a^2) = \text{asy.var} (\hat{\sigma}_b^2) - \text{asy.var} (\hat{\sigma}_a^2) = 4\sigma^4 - 2\sigma^4 = 2\sigma^4.
\]

- In order to construct a test statistic based on the distance of the variance estimators, we need a consistent estimation of the asymptotic variance \(2\sigma^4\).

For this purpose, e.g.,
\[
\hat{2}\sigma^4 = 2 (\hat{\sigma}_a^2)^2
\]
can be used. Then, we obtain under the Null
\[
d_1 = \frac{\sqrt{2n} (\hat{\sigma}_b^2 - \hat{\sigma}_a^2)}{\sqrt{2 (\hat{\sigma}_a^2)^2}} = \frac{\hat{\sigma}_b^2 - \hat{\sigma}_a^2}{\sqrt{n (\hat{\sigma}_a^2)^2}} \sim N (0, 1),
\]
which can be used to test the RW.

- In particular, the RW III and RW I is rejected at a significance level \(\alpha\), if
\[
|d_1| \geq Z_{1-\alpha/2} = (1 - \alpha/2)\text{-quantile of a N}(0,1)\text{-distribution}
\]

- Another possible test statistic is based on the estimated variance-ratio exploiting the fact that under \(H_0\) the variance is linear in the periods:
\[
\overline{VR} = \frac{\text{var} (r_t^2)}{2 \text{var} [r_t]} = \frac{1}{2} \text{var} (r_t [2]) = \frac{\hat{\sigma}_b^2}{\hat{\sigma}_a^2},
\]
which can be transformed to:
\[
\overline{VR} - 1 = \frac{\hat{\sigma}_b^2 - \hat{\sigma}_a^2}{\hat{\sigma}_a^2}
\]
- Under a RW I model:

\[ d_2 = \sqrt{n} \left( V R - 1 \right) \overset{\text{d}}{\sim} N \left( 0, 1 \right), \]

which results from the fact that \( d_1 = d_2 \):

\[ d_2 = \sqrt{n} \left[ \frac{\sigma_h^2 - \sigma_a^2}{\sigma_a^2} \right] = \frac{\hat{\sigma}_h^2 - \hat{\sigma}_a^2}{\sqrt{\frac{1}{n} (\hat{\sigma}_a^2)}} = d_1. \]

However, observe that this equality holds only if \( 2 (\hat{\sigma}_a^2) \) is used as an estimate of \( 2\sigma^4 \) in order to construct \( d_1 \). For other estimators \( d_1 \neq d_2 \).

- Figure 5 shows the daily closing prices of the Daimler-Benz stock and the german stock index DAX, and the daily spot exchange rates for the US-Dollar/Deutsche Mark together with the corresponding VR test statistic.

- Comments on the VR-test

  - The VR-test for 2-period returns can be easily generalized to multi-period returns. In particular, note that under the RW

\[ \text{var} \left( r_t [q] \right) = \text{var} \left[ r_t + r_{t-1} + \cdots + h_{t-q+1} \right] = q \cdot \text{var} \left[ r_t \right], \]

which can be exploited to construct a corresponding test statistic (for details, see, CLM97).

  - The discussion of the asymptotic distributions of the test statistics was based on the iid assumption of the returns.

However, this strong assumptions is typically violated since we observe volatility changes and volatility clustering over time.

In order to account for this fact, CLM 97 derive a VR test which allows for special forms of serial dependence in the return process (ARCH–effects). This modified test then can be used to check the RW III hypothesis, of uncorrelated returns.
Figure 5: The ACF and Ljung-Box statistic for the return series of the Daimler-Benz stock, the German stock index DAX, and the spot exchange rate for the US-Dollar/Deutsche Mark.
2.5 Tests for Long–Range Dependence

- The test procedures for the RW-hypotheses discussed so far, test the Null:
  \[ r_t = p_t - p_{t-1} = \mu + \varepsilon_t, \quad \varepsilon_t \sim \text{iid or } \varepsilon_t \sim \text{WN} \]
  against the alternative
  \[ r_t \text{ is serially dependent or correlated.} \]

- In particular, these tests are designed to reveal short–range serial dependence in the return process, i.e.,
  
  correlation between \( r_t \) and \( r_{t-\tau} \), where \( |\tau| \) is comparably small.

- However, a departure from the RW which might be important and which cannot be detected by the tests discussed so far, is the long–range dependence, i.e.,
  
  correlation between \( r_t \) and \( r_{t-\tau} \), where \( |\tau| \) is comparably large.

- The phenomenon of long–range dependence, implying an unusually high degree of persistence and a very long memory in the time series behavior is an important issue discussed in hydrology and geophysics\(^\text{10}\).

- In the following, I provide a short review of the relevant instruments which are necessary to analyze long–range dependence.

**Fractionally integrated processes**

- Fractionally integrated processes are typical examples for processes with a long–range dependence.

Recall that the RW-model for the log prices

\[ r_t = p_t - p_{t-1} = \varepsilon_t, \quad \varepsilon_t \sim \text{WN} \]

can be written as

\[ (1 - L)^k p_t = \Delta^k p_t = \varepsilon_t \quad \text{with } k = 1, \]

where \( \Delta^k \equiv (1 - L)^k \) is the difference operator.

Hence, the \( k \)-th difference \( (k = 1) \) of the \( p_t \)-series is stationary. The typical assumption is that \( k \) is an integer, typically \( k = 1 \).

\(^{10}\text{See, e.g., CLM97.}\)
- Granger and Joyeux (1980)\textsuperscript{11} suggested that non-integer values for $k$ might also be useful, and introduced the fractionally integrated processes characterized by non-integer values for $k$.

- In order to express $(1 - L)^k$ is terms of a sum, which is helpful for interpretation, one can use the binomial theorem for non-integer powers $k$\textsuperscript{12}:

\[
(1 - L)^k = \sum_{j=0}^{\infty} (-1)^j \binom{k}{j} L^j,
\]

where

\[
\binom{k}{j} = \frac{k(k-1)(k-2)\cdots(k-j+1)}{j!} \quad \text{(binomial coefficient)}
\]

\[
(-1)^j \binom{k}{j} = \frac{\Gamma(j-k)}{\Gamma(-k)\Gamma(j+1)}
\]

\[
\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt, \quad \text{(Gamma function)},
\]

with $\Gamma(x+1) = x\Gamma(x)$.

- Applying this expansion to $p_t$ as defined above yields:

\[
(1 - L)^k p_t = \sum_{j=0}^{\infty} \frac{\Gamma(j-k)}{\Gamma(-k)\Gamma(j+1)} p_{t-j} = \varepsilon_t,
\]

which can be solved for $p_t$:

\[
\frac{\Gamma(-k)}{\Gamma(-k)\Gamma(1)} p_t + \sum_{j=1}^{\infty} \frac{\Gamma(j-k)}{\Gamma(-k)\Gamma(j+1)} p_{t-j} = \varepsilon_t
\]

\[
p_t = \sum_{j=1}^{\infty} \phi_j p_{t-j} + \varepsilon_t, \quad \text{with} \quad \phi_j = - \frac{\Gamma(j-k)}{\Gamma(-k)\Gamma(j+1)}.
\]

Hence, a fractionally integrated process can be expressed as an AR($\infty$) process with a particular sequence of coefficient $\{\phi_j\}$.

- $p_t$ may also be viewed as a MA($\infty$) process. In particular, assume that $(1 - L)^k$ is invertible (which is the case for $k < 1/2$). Then, we obtain

\[
p_t = (1 - L)^{-k} \varepsilon_t
\]

\[
= \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad \text{with} \quad \psi_j = (-1)^j \binom{-k}{j} = \frac{\Gamma(j+k)}{\Gamma(k)\Gamma(j+1)}.
\]


\textsuperscript{12}Another possibility is to apply a Taylor-series expansion to $f(L) = (1 - L)^k$ around $L = 0$ (see, H94, p. 448).
- It can be shown by a corresponding Taylor series expansion that:

\[ \psi_j = \frac{\Gamma(j + k)}{\Gamma(k)\Gamma(j + 1)} \simeq (j + 1)^{k-1}, \quad \text{as} \quad j \to \infty. \]

Hence, the impulse response function of the MA–representation of an invertible fractionally integrated process decays (for large \( j \)) hyperbolically with \( j \).

Remember, that for a stationary AR(1) \( p_t = (1 - \phi_1 L)^{-1} \varepsilon_t = \sum_j \phi_j^t \varepsilon_t \), e.g., the impulse response function is

\[ \psi_j = \phi_j^t, \]

which decays geometrically with \( j \), and hence much faster.

This slow decay of the impulse function of fractionally integrated processes reflects the “long-memory” and “strong persistence” of these processes.

- Closely related to this result is the fact that for \(-\frac{1}{2} < k < \frac{1}{2}\) the ACF of a fractionally integrated process can be approximated by

\[ \rho_\ell \simeq \frac{\Gamma(1 - k)}{\Gamma(k)} \ell^{2k-1}, \quad \text{as} \quad \ell \to \infty, \]

where:

\[ \rho_\ell > 0, \quad \text{if} \quad \begin{cases} k > 0, \quad \ell > 0, \\ k < 0, \quad \ell < 0. \end{cases} \]

Hence, for large \( \ell \) and \( k \in (-\frac{1}{2}, \frac{1}{2}) \) the ACF decays hyperbolically with lag \( \ell \). Remember that the ACF of an AR(1) decays geometrically.

- Figure 6 shows the ACF of fractionally intergrated processes and of a stationary AR(1).

---

\(^{13}\text{See, H94, p. 448.}\)
Figure 6: Autocorrelation functions $\rho(\ell)$ of fractionally integrated processes
Rescaled Range (R/S) statistic

- The original statistical measure of long memory proposed by the hydrologist Hurst (1951)\(^{14}\) is the R/S statistic. Mandelbrot (1971)\(^{15}\) was the first who used this measure to analyze long-range dependence in asset markets.

- Let

\[
\{r_t\}_{t=1}^T, \quad \text{with} \quad \bar{r}_T = \frac{1}{T} \sum_{t=1}^T r_t, \quad s_T = \sqrt{\frac{1}{T} \sum_{t=1}^T (r_t - \bar{r}_T)^2},
\]

denote a time series of returns. Then, the R/S-statistic is given by:

\[
Q_T = \frac{1}{s_T} \left[ \max_{1 \leq t \leq T} \sum_{j=1}^t (r_j - \bar{r}_T) - \min_{1 \leq t \leq T} \sum_{j=1}^t (r_j - \bar{r}_T) \right] = w_{1t}
\]

where:

\[
w_{1t} = \text{maximum of the partial sums of the first } t \text{ deviations of } r_j \text{ from its mean } \bar{r}_T.
\]

Since \(\sum_{j=1}^T (r_j - \bar{r}_T) = 0\), \(w_{1t}\) is non-negative.

\[
w_{2t} = \text{corresponding minimum of the partial sums, which is non-positive.}
\]

Hence, the difference between \(w_{1t}\) and \(w_{2t}\) is non-negative and hence: \(Q_T \geq 0\).

- The R/S-Statistic was originally employed by Hurst (1951) to study reservoir storage where the \(r\)'s are the variable inflows from a river in successive time periods and \(\bar{r}_T\) was the constant outflow.

- To illustrate the basic principle of this measure consider the following example

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r_t)</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>partial sum (\sum_{j=1}^t r_j)</td>
<td>1</td>
<td>4</td>
<td>8</td>
<td>10</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>(t \cdot \bar{r}_T = t \cdot 3)</td>
<td>3</td>
<td>6</td>
<td>9</td>
<td>12</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>(\sum_{j=1}^t r_j - t\bar{r}_T)</td>
<td>-2</td>
<td>-2</td>
<td>-1</td>
<td>-2</td>
<td>-2</td>
<td>0</td>
</tr>
</tbody>
</table>

Hence, the maximum of partial sums is \(w_{1t} = 0\) and the minimum \(w_{2t} = -2\) such that \(Q_T = 2\).


- Feller (1951)\textsuperscript{16} has shown that the asymptotic distribution of the standardized R/S–statistic $Q_T \cdot T^{-1/2}$ for iid variables is that of the range of a Brownian Bridge. The expectation of this asymptotic distribution is

$$\text{Asy.E} \left[ \frac{1}{\sqrt{T}} Q_T \right] = \sqrt{\frac{\pi}{2}}.$$ 

- Hence, under the iid-assumption we obtain approximately

$$\mathbb{E} [Q_T] \approx \sqrt{\frac{\pi}{2}} T^{1/2}, \quad \text{as} \quad T \to \infty,$$

and a R/S–statistic which significantly exceeds this expectation:

$$\sqrt{\frac{\pi}{2}} T^{1/2} \ll Q_T$$

is interpreted as evidence of long memory.

- Since

$$\ln \mathbb{E} [Q_T] \approx \frac{1}{2} \ln \left( \frac{\pi}{2} \right) + \frac{1}{2} \ln(T), \quad \text{as} \quad T \to \infty,$$

a test of the Null of iid returns against the alternative of long–range dependent returns can be implemented using the following auxiliary regression:

$$\ln Q_T = \alpha + \beta \ln \tau + \eta_T; \quad \tau = T - k, \ldots, T - 1, T.$$

- If the LS–estimate of $\beta$ significantly exceeds $1/2$ one can reject the RW–hypothesis for the log–price in favor of long–range dependent returns.

- Figure 7 plots $Q_T$ and $\mathbb{E} [Q_T] \approx \sqrt{\frac{\pi}{2}} T^{1/2}$ against $\tau$ using a log scale for three return series (Dollar/DM, Daimler, DAX).

The asymptotic approximation appears to be valid for $\tau > 30$ (Dollar/DM) and for $\tau > 40$ (Daimler, DAX).

However, for the Dollar/DM–series, we obtain for $\tau > 1000$ significant deviations between $Q_T$ and its expected values under the iid assumption, indicating some evidence of long–memory.

- Properties of the R/S–statistic (see, CLM97):

• The R/S-statistic can detect long-range dependence in highly non-Gaussian time-series with large skewness and/or kurtosis, or with non-existing variance, observed for many return series.
   (The autocorrelation and variance-ratio need not be well defined for a process with infinite variance.)

• An important shortcoming of the R/S-statistic is its sensitivity to short-range dependence: symptoms of short-range memory are sometimes interpreted as long-range dependence.
Figure 7: Observed and Predicted R/S-statistic for the return series of the Daimler-Benz stock, the German stock index DAX, and the spot exchange rate for the US-Dollar/Deutsche Mark using a log scale.