Master MIF

Financial Econometrics

Nour Meddahi (Toulouse School of Economic)

Lecture 1 - Part A

Introduction to Financial Econometrics

I. Examples

During these lectures, we will be interested in characterizing the relationships between financial variables. These relationships are sometimes implied by the financial theory like the **CAPM** or the **term structure of interest rates**. Some times, the relationships are just statistical ones (no theory behind the relationship). For instance, this is the case when one wants to forecast the market's return by using the dividend-price ratio.

In both case, these relationships depend on **unknown** parameters. In these case, one has to use the available data to **estimate** the unknown parameters, i.e., one has to extract the information in the data about the unknown parameters. We will study several estimation methods, and compare their properties.

In many examples, one does not know whether a theory is supported by the data (like the CAPM) or whether a particular variable should appear in the relationship (e.g., whether the interest rates help to predict the markets' return). Again, one will use the data and **test** these assumptions. We will study several testing procedures, and compare their properties.

Example 1: Capital Asset Pricing Model (CAPM).

• Define the price of an individual stock as S_t^i and the market index by M_t. Denote the corresponding arithmetic returns as

$$r_t^i = \frac{S_t^i}{S_{t-1}^i} - 1,\ r_t^M = \frac{M_t}{M_{t-1}} - 1$$

Likewise, r_t^B as the risk-free rate of interest (in practice the return on 3 month bonds and assumed non-stochastic).

• Then, the basic CAPM implies

$$E[r_{t+1}^i - r_{t+1}^B] = \beta_i(E[r_{t+1}^M - r_{t+1}^B]),\ \beta_i = \frac{Cov[r_{t+1}^i, r_{t+1}^M]}{Var[r_{t+1}^M]},$$

which means

$$r_{t+1}^i - r_{t+1}^B = \beta_i(r_{t+1}^M - r_{t+1}^B) + \varepsilon_{t+1},\ E[\varepsilon_{t+1}] = 0,\ Cov[r_{t+1}^M, \varepsilon_{t+1}] = 0.$$

• Consequently, by considering the linear regression model

$$r_{t+1}^i - r_{t+1}^B = \alpha_i + \beta_i(r_{t+1}^M - r_{t+1}^B) + \varepsilon_{t+1},$$

the constant α_i should be zero, which is a testable restriction.

• Fama and French suggested that one should add two other variables in the regression.

Example 2: Dynamic Term Structure of Interest Rates.

• Assume that the short term interest rate x_t is characterized by the following dynamics under the Q-measure:

$$dx_t = \kappa(\mu - x_t)dt + \sigma dW_t$$

where W_t is a standard Brownian process. In what follows, θ is $\theta = (\mu, \kappa, \sigma)^\top$.

• One can show that the discretization of this process leads to

$$\forall \Delta > 0,\ x_{t+\Delta} = \mu + \exp(-\kappa\Delta)(x_t - \mu) + \varepsilon_{t+\Delta},\ \text{where}$$
$$\varepsilon_{t+\Delta}\ i.i.d.\ \sim \mathcal{N}\left(0, \sigma^2 \frac{(1 - \exp(-2\kappa\Delta))}{2\kappa}\right).$$

• The Q-measure is a probability measure such that the price of a financial asset equals the expected value of the (discounted) future cash flows.

• One can show that there exist two functions $a(\cdot)$ and $b(\cdot)$ such that the price at time t of an asset that gives you one pound at time $t + h$ is given by

$$B(t, h) = E\left[\exp\left(-\int_t^{t+h} x_s ds\right) |x_\tau, \tau \leq t\right] = \exp\left(a(h, \theta)x_t + b(h, \theta)\right).$$

• Hence, when one knows θ, one can characterize the **term structure of interest rates**, i.e. the function $y(t, h) = -\ln(B(t, h))/h$ when h varies.

Example 3: Forecasting the Market's return.

• When one wants to invest in a stock, she/he often tries to understand the future behavior of the stock price.

• A simple way to do it is to try to predict or forecast the future returns of the stock by using the available data.

• The predictability of asset returns is quite small when one considers short horizons (day, week). However, it increases when one considers long-horizons (up to five years).

• A common variable used to predict future stock returns is the dividend-price ratio. More precisely, one is interested in the regression

$$r_{t:t+k} = a + b\frac{D_t}{S_t} + \varepsilon_{t+1}, \text{ where } r_{t:t+k} = \frac{S_{t+k}}{S_t} - 1.$$

• Note that this relation is not impled by a financial theory; it is a statistical (or probabilistic) relationship.

• The critical assumption is whether $b \neq$ or not. This restriction should be tested.

• One could also add other variables in the equation like the short term interest rate or some macroeconomic variables like the "CAY" variable of Lettau and Ludvigson.

II. Stylized Facts of Asset Returns

- We can consider the following list of so-called stylized facts which apply to most stochastic returns.

- We will use daily returns on the S&P500 from 1/1/97 to 12/31/01 to illustrate each of the features (from the book Christoffersen (2003)).

- We will also use some equity returns from 1926 to 1997 (from the book Tsay (2002)).

- We start the analysis by defining some quantities of interest.

Defining Asset Returns:

- The daily geometric or "log" return on an asset is defined as

$$R_t = \ln(S_t) - \ln(S_{t-1}).$$

- The arithmetic return is instead defined as

$$r_t = \frac{S_t - S_{t-1}}{S_{t-1}} = \frac{S_t}{S_{t-1}} - 1.$$

- The two returns are typically fairly similar, as can be seen from

$$R_t = \ln\left(\frac{S_t}{S_{t-1}}\right) = \ln\left(\frac{S_t}{S_{t-1}} - 1 + 1\right) = \ln(r_t + 1) \approx r_t,$$

the approximation holds because $\ln(1+x) \approx x$ when x is close to 0.

- One advantage of the log return is that we can easily calculate the compounded return for K-days

$$R_{t:t+k} = \ln(S_{t+k}) - \ln(S_t) = \sum_{i=1}^{k} \ln(S_{t+i}) - \ln(S_{t+i-1}) = \sum_{i=1}^{k} R_{t+i}.$$

Moments of a random variable:

$$\text{Expectation or Mean} = E[X] = \mu$$

$$\text{Variance} = E[(X-\mu)^2] = E[X^2] - (E[X])^2 = \sigma^2$$

$$\text{Volatility} = \sigma$$

$$\text{Skewness coefficient} = \frac{E[(X-\mu)^3]}{(Var[X])^{3/2}}$$

$$\text{Kurtosis coefficient} = \frac{E[(X-\mu)^4]}{(Var[X])^2}$$

Dependence between two random variables:

$$\text{Covariance} = Cov[X,Y] = E[(X-\mu_x)(Y-\mu_y)] = E[XY] - \mu_X\mu_y$$

$$\text{Correlation} = Corr[X,Y] = \frac{Cov[X,Y]}{\sqrt{Var[X]Var[Y]}}.$$

Sample counterparts. One observes $x_1, x_2, ..., x_n$:

$$\text{sample mean of X} : \bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\text{sample variance of X} : \hat{V}ar[X] = \frac{1}{n}\sum_{i=1}^{n} x_i^2 \; - \; (\bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X})^2$$

$$\text{sample skewness of X} : \hat{S}kew[X] = \frac{1}{(\hat{V}ar[X])^{3/2}}\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X})^3$$

$$\text{sample kurtosis of X} : \hat{K}urt[X] = \frac{1}{(\hat{V}ar[X])^2}\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X})^4$$

$$\text{sample covariance of X and Y} : \hat{Cov}[X, Y] = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})$$

$$\text{sample correlation of X and Y} : \hat{Corr}[X, Y] = \frac{\hat{Cov}[X, Y]}{\sqrt{\hat{Var}[X]\hat{Var}[Y]}}$$

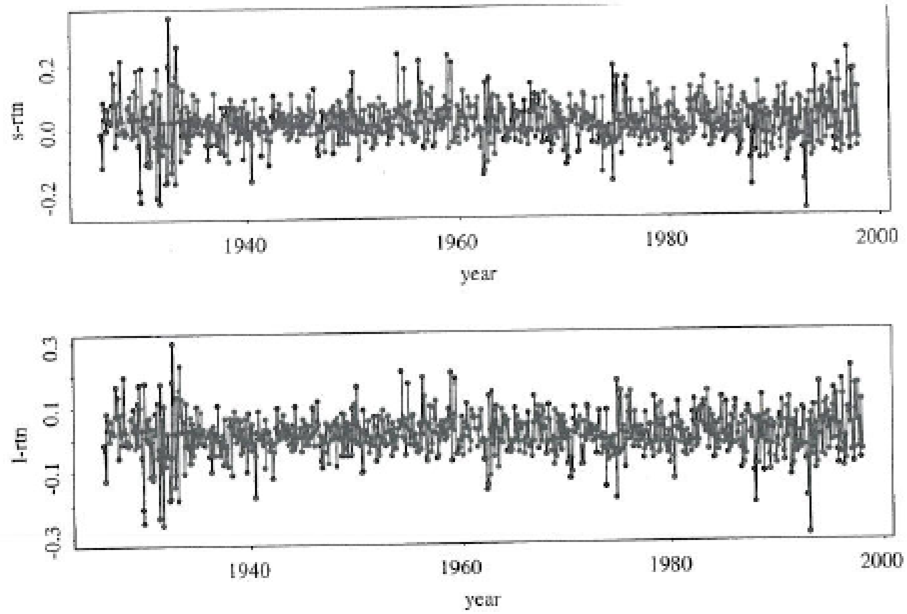


Figure 1.2. Time plots of monthly returns of IBM stock from January 1926 to December 1997. The upper panel is for simple net returns, and the lower panel is for log returns.

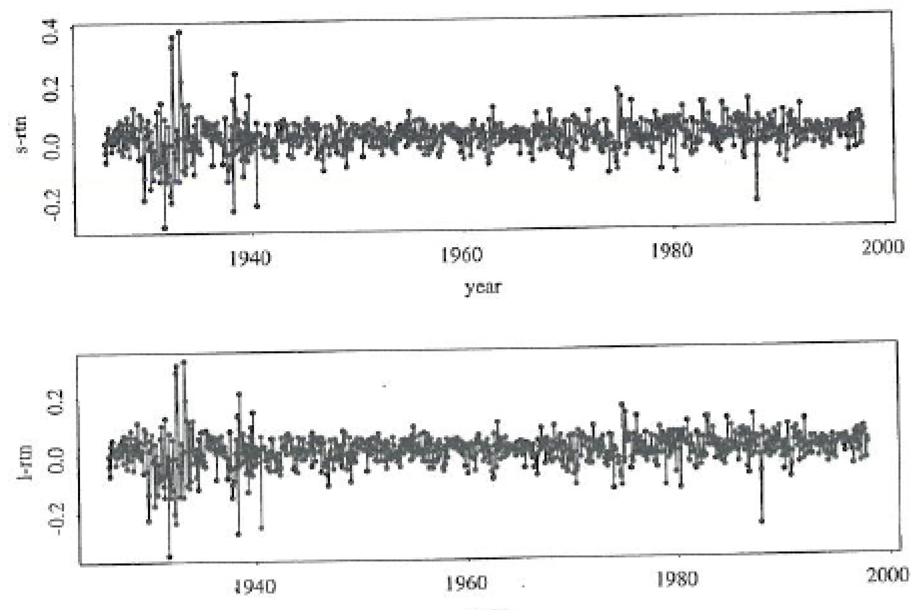


Figure 1.3. Time plots of monthly returns of the value-weighted index from January 1926 to December 1997. The upper panel is for simple net returns, and the lower panel is for log returns.

Table 1.2. Descriptive Statistics for Daily and Monthly Simple and Log Returns of Selected Indexes and Stocks. Returns Are in Percentages, and the Sample Period Ends on December 31, 1997. The Statistics Are Defined in Equations (1.10) to (1.13), and VW and EW Denote Value-Weighted and Equal-Weighted Indexes.

Security	Start	Size	Mean	Stan. Dev.	Skew.	Excess Kurt.	Min.	Max.
				(a) Daily simple returns (%)				
VW	62/7/3	8938	0.049	0.798	−1.23	30.06	−17.18	8.67
EW	62/7/3	8938	0.083	0.674	−1.09	18.09	−10.48	6.95
I.B.M.	62/7/3	8938	0.050	1.479	0.01	11.34	−22.96	12.94
Intel	72/12/15	6329	0.138	2.880	−0.17	6.76	−29.57	26.38
3M	62/7/3	8938	0.051	1.395	−0.55	16.92	−25.98	11.54
Microsoft	86/3/14	2985	0.201	2.422	−0.47	12.08	−30.13	17.97
Citi-Grp	86/10/30	2825	0.125	2.124	−0.06	9.16	−21.74	20.75
				(b) Daily log returns (%)				
VW	62/7/3	8938	0.046	0.803	−1.66	40.06	−18.84	8.31
EW	62/7/3	8938	0.080	0.676	−1.29	19.98	−11.08	6.72
I.B.M.	62/7/3	8938	0.039	1.481	−0.33	15.21	−26.09	12.17
Intel	72/12/15	6329	0.096	2.894	−0.59	8.81	−35.06	23.41
3M	62/7/3	8938	0.041	1.403	−1.05	27.03	−30.08	10.92
Microsoft	86/3/14	2985	0.171	2.443	−1.10	19.65	−35.83	16.53
Citi-Grp	86/10/30	2825	0.102	2.128	−0.44	10.68	−24.51	18.86
				(c) Monthly simple returns (%)				
VW	26/1	864	0.99	5.49	0.23	8.13	−29.00	38.28
EW	26/1	864	1.32	7.54	1.65	15.24	−31.23	65.51
I.B.M.	26/1	864	1.42	6.70	0.17	1.94	−26.19	35.12
Intel	72/12	300	2.86	12.95	0.59	3.29	−44.87	62.50
3M	46/2	623	1.36	6.46	0.16	0.89	−27.83	25.77
Microsoft	86/4	141	4.26	10.96	0.81	2.32	−24.91	51.55
Citi-Grp	86/11	134	2.55	9.17	−0.14	0.47	−26.46	26.08
				(d) Monthly log returns (%)				
VW	26/1	864	0.83	5.48	−0.53	7.31	−34.25	32.41
EW	26/1	864	1.04	7.24	0.34	8.91	−37.44	50.38
I.B.M.	26/1	864	1.19	6.63	−0.22	2.05	−30.37	30.10
Intel	72/12	300	2.03	12.63	−0.32	3.20	−59.54	48.55
3M	46/2	623	1.15	6.39	−0.14	1.32	−32.61	22.92
Microsoft	86/4	141	3.64	10.29	0.29	1.32	−28.64	41.58
Citi-Grp	86/11	134	2.11	9.11	−0.50	1.14	−30.73	23.18

Table from Tsay (2002): Analysis of Financial Time Series

Normal and Student distributions

- **Normal distribution.** If $X \sim \mathcal{N}(\mu, \sigma^2)$, with density function. $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$. Then,

$$E[X] = \mu, \quad Var[X] = \sigma^2, \quad \text{Skew}[X] = 0, \quad \text{Kurt}[X] = 3.$$

- **Chi-squared distribution:** Let X_i for $i = 1, 2, \ldots, n$ be a set of independent standard normal random variables. Let $Y \equiv \sum_{i=1}^{n} X_i^2$. Then Y is Chi-squared distribution with n degrees of freedom.

- **Student distribution $\mathcal{T}(n)$.** Suppose that X is standard normal, Y is Chi-squared with n degrees of freedom, and X and Y are independent. Then $Z = X/\sqrt{Y/n}$ is T-distributed and has the density function:

$$f(x) = \frac{1}{\sqrt{\pi n}} \frac{\Gamma\left((n+1)/2\right)}{\Gamma(n/2)} \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2}, \text{ where } \Gamma(w) \equiv \int_0^\infty s^{w-1} \exp[-s] ds.$$

- The moments of order $m \geq n$ do not exist. When the moments exist, one has

$$E[\mathcal{T}(n)] = 0, \; Var[\mathcal{T}(n)] = \frac{n}{n-2}, \; Skew[\mathcal{T}(n)] = 0, \; Kurt[\mathcal{T}(n)] = 3\frac{n-2}{n-4}$$

- When $n \to +\infty$, we have $\mathcal{T}(n) \to \mathcal{N}(0, 1)$. In practice, $n = 30$.

Stylized Fact 1:

• Daily returns have very little autocorrelation. We can write

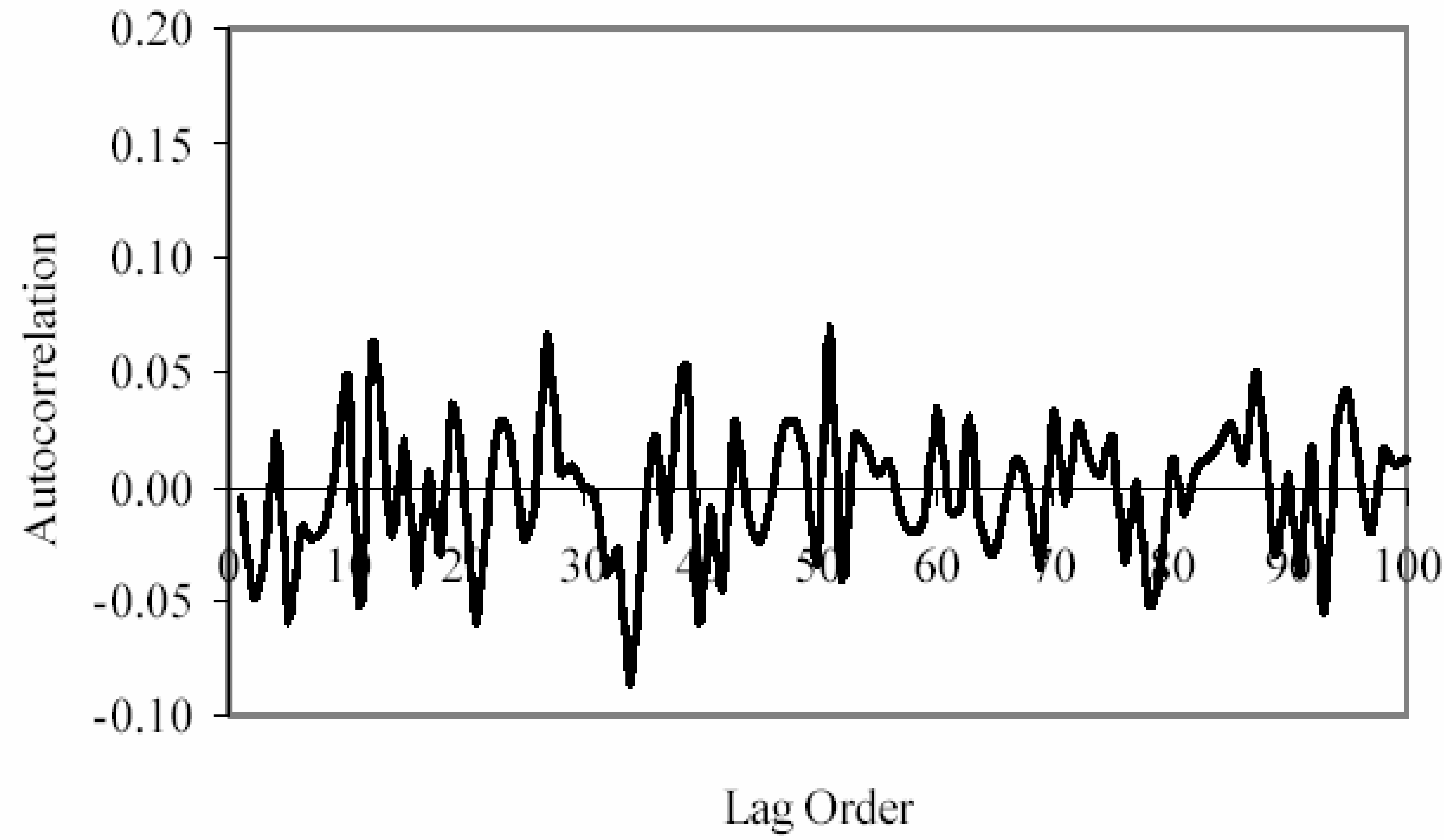
$$Corr[R_t, R_{t-i}] \approx 0, \;\; i = 1, 2, ...$$

• Returns are almost impossible to predict from their own past. Fig 1.1 shows the correlation of daily S&P500 returns with returns lagged from one to 100 days.

• We will take this as evidence that the conditional mean is roughly constant when one considers daily or weekly returns.

• However, when one decreases the frequency, some serial correlations appear in the returns. This serial correlation does conflict with modern financial theory.

Autocorrelations of Daily S&P Returns for Lags 1 through 100 1/1/97-12/31/01 Figure 1.1



Stylized Fact 2:

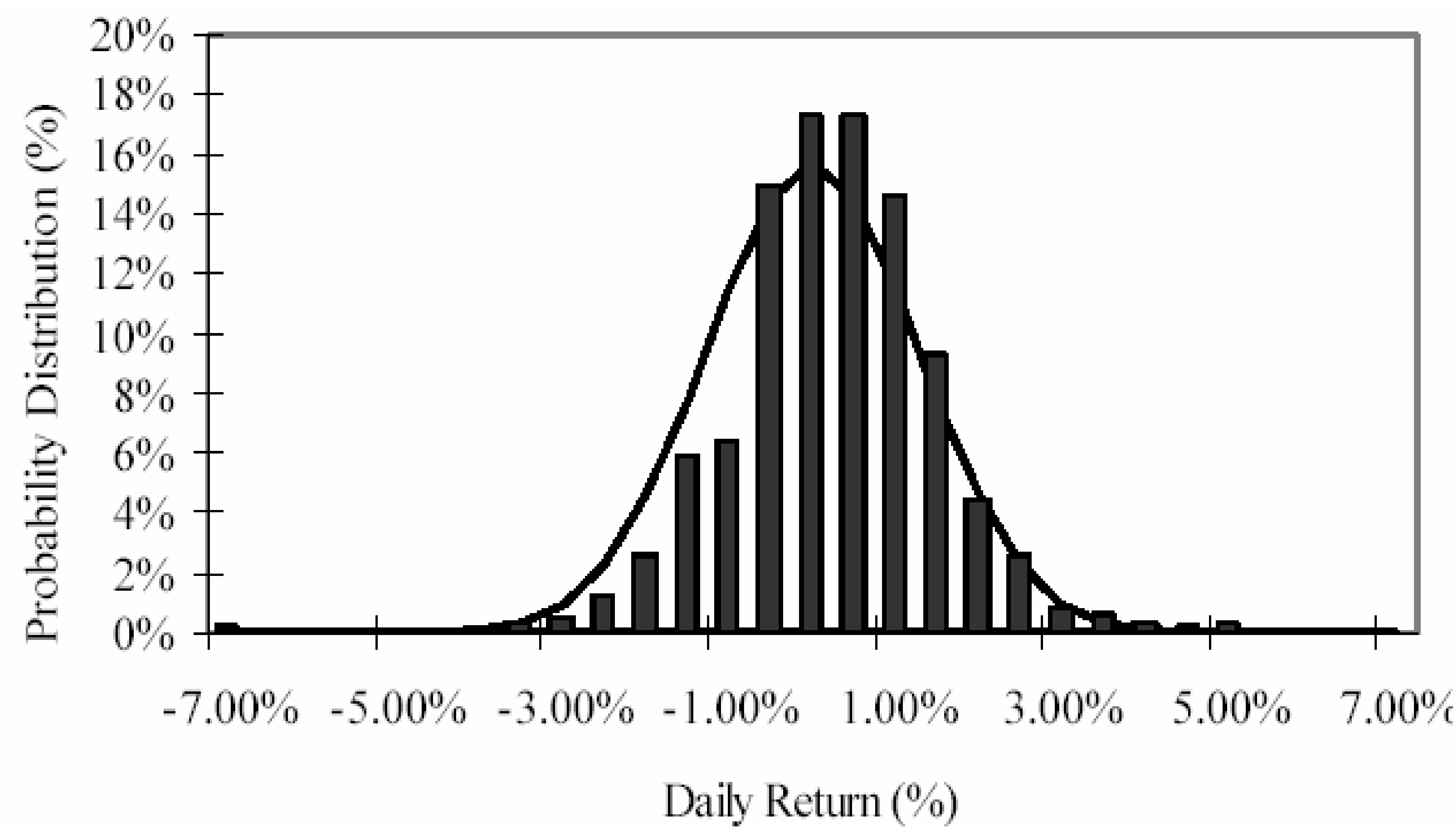
• The unconditional distribution of daily returns have fatter tail than the normal distribution.

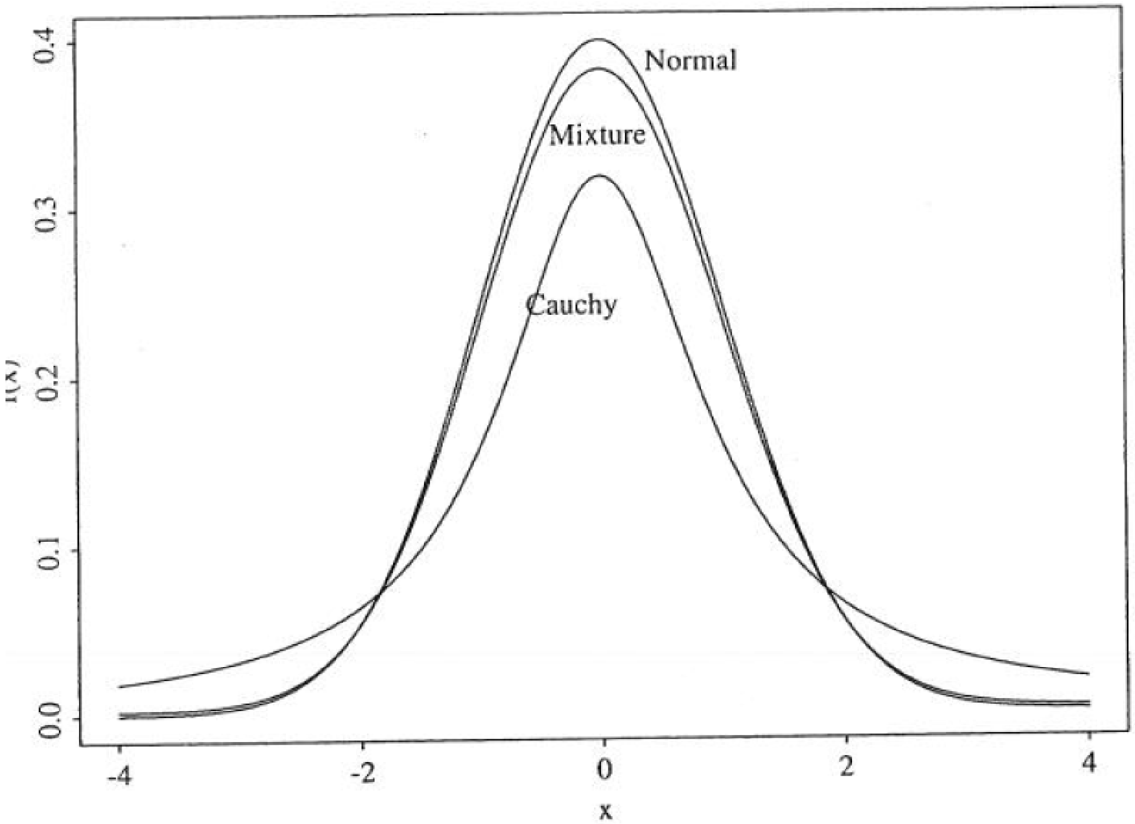
• Fig.1.2 shows a histogram of the daily S&P500 return data with the normal distribution imposed.

• Notice how the histogram has longer and fatter tails, in particular in the left side, and how it is more peaked around zero than the normal distribution.

• Fatter tails mean a higher probability of large losses than the normal distribution would suggest.

Histogram of Daily S&P Returns Superimposed on the Normal Distribution 1.1.97 – 12.31.01 – Fig.1.2





Normal
Mixture
Cauchy
f(x)
0.0
0.1
0.2
0.3
0.4
-4
-2
0
2
4
x

Stylized Fact 3:

- The stock market exhibits occasional, very large drops but not equally large up-moves.
- Consequently the return distribution is asymmetric or negatively skewed.
- This is clear from Figure 1.2 as well.
- Other markets such as that for foreign exchange tend to show less evidence of skewness.

Stylized Fact 4:

- The standard deviation of returns completely dominates the mean of returns at short horizons such as daily.
- It is typically not possible to statistically reject a zero mean return.
- Our S&P500 data has a daily mean of 0.0353% and a daily standard deviation of 1.2689%.

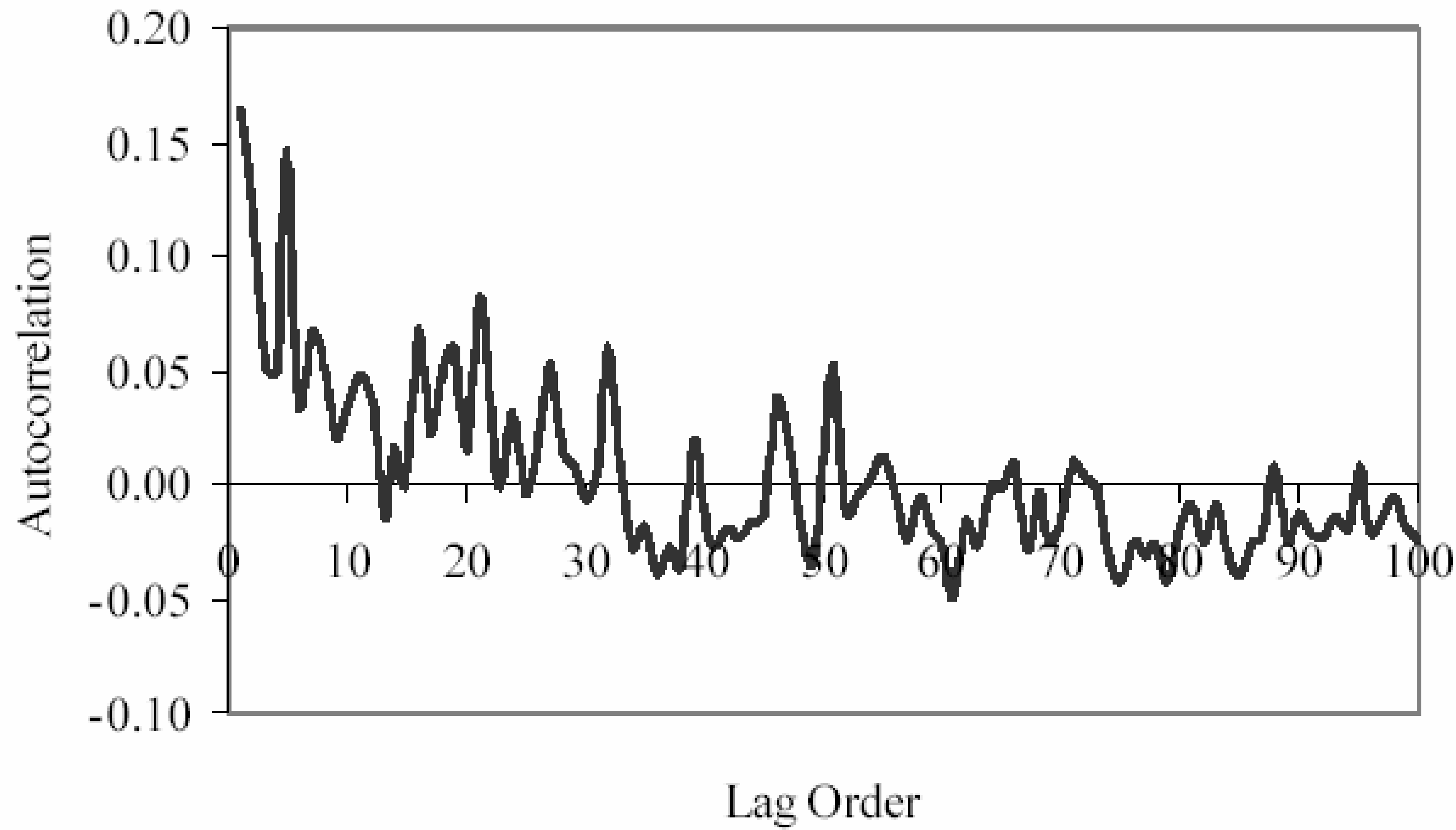
Stylized Fact 5:

- Variance measured for example by squared returns, displays positive correlation with its own past.
- This is most evident at short horizons such as daily or weekly.
- Fig 1.3 shows the autocorrelation in squared returns for the S&P500 data, that is

$$Corr[R_t^2, R_{t-i}^2] > 0 \text{ for small } i > 0.$$

- Models which can capture this variance dependence will be presented in SAFE II.

Autocorrelation of Squared Daily S&P500 Returns for Lags 1 through 100 1.1.97 – 12.31.01



Stylized Fact 6:

• Equity and equity indices display negative correlation between variance and returns.

• This often termed the leveraged effect, arising from the fact that a drop in stock price will increase the leverage of the firm as long as debt stays constant.

• This increase in leverage might explain the increase variance associated with the price drop.

Stylized Fact 7:

• Correlation between assets appears to be time varying.

• Importantly, the correlation between assets appear to increase in highly volatile down-markets and extremely so during market crashes.

Stylized Fact 8:

• Even after standardizing returns by a time-varying volatility measure, they still have fatter than normal tails.

• We will refer to this as evidence of conditional non-normality.

Stylized Fact 9:

• As the return-horizon increases, the unconditional return distribution changes and looks increasingly like the normal distribution.

Asset Return Model:

• Based on the above of stylized facts, a common model of individual asset returns will take the generic form

$$R_{t+1} = \mu_t + \sigma_t z_{t+1}, \text{ where } z_{t+1} \text{ i.i.d. } \sim D(0,1).$$

• The conditional mean return is thus μ_t and the conditional variance σ_t^2. A good empirical model for σ_t^2 is a GARCH(1,1) specification. μ_t is constant for short horizons and time-varying when one considers long horizons.

• The random variable z_{t+1} is an innovation term, which we assume is identically and independently distributed (i.i.d.) as D(0,1). A good empirical example is a Student $\mathcal{T}(\nu)$ distribution.

Stylized Facts of other financial data.

Returns on foreign exchange returns: Like equities but symmetric (no skewness, no leverage effect). The kurtosis is also smaller.

Interest rates and bond returns: Highly persistent (high autocorrelation). We will study these data during the lectures on Time Series.

Master MIF

Financial Econometrics

Nour Meddahi (Toulouse School of Economic)

Lecture 1 - Part B

Predictability of Asset Returns

Serial Correlation in Asset Returns

$$R_t = \log p_t - \log p_{t-1} = \mu_{t-1} + \varepsilon_t$$

Three types of dependence:

1. i.i.d.,
2. martingale difference sequences,
3. uncorrelated returns.

- Variance Ratio Tests:

$$\frac{Var[\sum_{k=1}^{q} r_{t+k}]}{qVar[r_t]} = 1 + 2\sum_{k=1}^{q-1}\left(1 - \frac{k}{q}\right)\rho(k)$$

- Box-Pierce Type Tests: $E[\varepsilon_t] = 0$, $Var[\varepsilon_t] = \sigma^2$, $E[\varepsilon_t^4] = \eta\sigma^4$, $\gamma(k) = Cov(r_t, r_{t-k})$

$$\sqrt{T}\,[\hat{\gamma}(0) - \gamma(0) \quad \hat{\gamma}(1) - \gamma(1) .... \hat{\gamma}(m) - \gamma(m)]^\top \sim^a \mathcal{N}(0, V)$$

$$V = [v_{ij}], \quad v_{ij} \equiv (\eta - 3)\gamma(i)\gamma(j) + \sum_{l=-\infty}^{l=+\infty} [\gamma(l)\gamma(l-i+j) + \gamma(l+j)\gamma(l-i)].$$

$$\sqrt{T}\,[\hat{\rho}(0) - \rho(0) \quad \hat{\rho}(1) - \rho(1) .... \hat{\rho}(m) - \rho(m)]^\top \sim^a \mathcal{N}(0, G)$$

$$G = [g_{ij}], \quad g_{ij} \equiv \sum_{l=-\infty}^{l=+\infty} [\rho(l)\rho(l-i+j) + \rho(l+j)\rho(l-i) - 2\rho(j)\rho(l)\rho(l-i)$$
$$- 2\rho(i)\rho(l)\rho(l-j) + 2\rho(i)\rho(j)\rho^2(l)].$$

Special example: uncorrelated noise.

Box-Pierce: $$T\sum_{k=1}^{m} \hat{\rho}^2(k) \;\sim\; \chi^2(m)$$

Ljung-Box: $$T(T+2)\sum_{k=1}^{m} \frac{\hat{\rho}^2(k)}{T-k} \;\sim\; \chi^2(m)$$

ARMA(p,q): The asymptotic distribution is $\chi^2(m-p-q)$ with $m > p + q$.

Predictive Regressions

$$y_t = \alpha + \beta x_{t-1} + u_t$$

$$x_t = \theta + \rho x_{t-1} + v_t$$

$$Cov\left([u_t, v_t]^\top, [u_t, v_t]\right) = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix}$$

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = (X^\top X)^{-1} X^\top Y.$$

Main statistical problem: $E[U \mid X] \neq 0$: The OLS estimator is biased in finite sample but consistent (asymptotic bias equals 0).

Under normality of $(u_t, v_t)^\top$, one has

$$E[\hat{\beta} - \beta] = \frac{\sigma_{uv}}{\sigma_v^2} E[\hat{\rho} - \rho] \approx -\frac{\sigma_{uv}}{\sigma_v^2}\left(\frac{1 + 3\rho}{T}\right) + O(1/T^2).$$

When one considers the dividend-price ratio, one has $\sigma_{uv} < 0$) and ρ close to one.

Other statistical issues:

Persistence of x_t (local-to-unity type asymptotics).

Estimation of $Var[OLS]$

1. One-step ahead forecast: Eicker-White robust method.
2. Multi-steps ahead forecasts: Newey-West estimator.

Table 1
Finite-sample properties of $\hat{\beta}$

The table reports finite-sample properties of the ordinary least squares (OLS) estimator $\hat{\beta}$ in the regression

$$y_t = \alpha + \beta x_{t-1} + u_t.$$

The sampling properties are computed under the assumption that x_t obeys the process

$$x_t = \theta + \rho x_{t-1} + v_t,$$

where $\rho^2 < 1$ and $[u_t\ v_t]'$ is distributed N(0, Σ), identically and independently across t. The true bias and higher-order moments depend on ρ and Σ (with distinct elements σ_u^2, σ_v^2, and σ_{uv}). For each sample period, those parameters are set equal to the estimates obtained when y_t is the continuously compounded return in month t on the value-weighted NYSE portfolio, in excess of the one-month T-bill return, and x_t is the dividend–price ratio on the value-weighted NYSE portfolio at the end of month t. The moments in the standard setting are conditioned on $x_0, \ldots, x_{T-1}$ and ignore any dependence of u_t on those values. The p-values are associated with a test of $\beta = 0$ versus $\beta > 0$

	Sample period			
	1927–1996	1927–1951	1952–1996	1977–1996
A. True properties				
Bias	0.07	0.18	0.18	0.42
Standard deviation	0.16	0.33	0.27	0.45
Skewness	0.71	0.83	0.98	1.29
Kurtosis	3.84	4.14	4.62	5.83
p-value for $\beta = 0$	0.17	0.42	0.15	0.64
B. Properties in the standard regression setting				
Bias	0	0	0	0
Standard deviation	0.14	0.27	0.20	0.30
Skewness	0	0	0	0
Kurtosis	3	3	3	3
p-value for $\beta = 0$	0.06	0.22	0.02	0.26
C. Sample characteristics and parameter values				
$\hat{\beta}$	0.21	0.21	0.44	0.19
T	840	300	540	240
ρ	0.972	0.948	0.980	0.987
$\sigma_u^2 \times 10^4$	30.05	54.46	16.42	17.50
$\sigma_v^2 \times 10^4$	0.108	0.247	0.029	0.033
$\sigma_{uv} \times 10^4$	− 1.621	− 3.360	− 0.651	− 0.715

Eq. (3), and Σ is set equal to the sample covariance matrix of the least-squares residuals from (1) and (3). Those values for ρ and Σ, as well as the sample size T and the realized sample value of $\hat{\beta}$, are given in Part C of Table 1. Part B of the table reports the corresponding moments and p-values implied by the standard regression model. The standard deviations in Part B depend on σ_u^2 and are

Table 7.1. *Long-horizon regressions of log stock returns on the log dividend-price ratio.*

$$r_{t+1} + \cdots + r_{t+K} = \beta(K)(d_t - p_t) + \eta_{t+K,K}$$

	Forecast Horizon (K)					
	1	3	12	24	36	48
1927 to 1994						
$\hat{\beta}(K)$	0.012	0.044	0.191	0.383	0.528	0.654
$R^2(K)$	0.004	0.015	0.068	0.144	0.209	0.267
$t(\hat{\beta}(K))$	1.221	1.400	2.079	4.113	4.631	3.943
1927 to 1951						
$\hat{\beta}(K)$	0.015	0.059	0.274	0.629	0.880	1.050
$R^2(K)$	0.003	0.014	0.074	0.207	0.322	0.424
$t(\hat{\beta}(K))$	0.660	0.844	1.677	4.521	2.967	3.783
1952 to 1994						
$\hat{\beta}(K)$	0.024	0.079	0.329	0.601	0.776	0.863
$R^2(K)$	0.015	0.047	0.190	0.344	0.428	0.432
$t(\hat{\beta}(K))$	2.733	3.055	3.228	3.225	3.315	3.561

r is the log real return on a value-weighted index of NYSE, AMEX, and NASDAQ stocks. $(d-p)$ is the log ratio of dividends over the last year to the current price. Regressions are estimated by OLS, with Hansen and Hodrick (1980) standard errors, calculated from equation (A.3.3) in the Appendix setting autocovariances beyond lag $K-1$ to zero. Newey and West (1987) standard errors with $q = K-1$ or $q = 2(K-1)$ are very similar and typically are slightly smaller than those reported in the table.

statistics also increase dramatically with the forecast horizon, although they are fairly stable within the range 3.0 to 3.5 in the postwar subsample.

It is interesting to compare the results in Table 7.1 with those obtained when stock returns are regressed onto the stochastically detrended short-term interest rate in Table 7.2. The regressions reported in Table 7.2 are run in just the same way as those in Table 7.1. Once again almost identical results are obtained if real returns are replaced by excess returns over the one-month Treasury bill rate.

Table 7.2 shows that, like the dividend-price ratio, the stochastically detrended short rate has some ability to forecast stock returns. However this forecasting power is very different in two respects. First, it is concentrated in the postwar subsample; this is not surprising since short-term interest rates were pegged by the Federal Reserve during much of the 1930s and 1940s, and so the detrended short rate hardly varies in these years. Second, the forecasting power of the short rate is at much shorter horizons than the

Table 7.2. *Long-horizon regressions of log stock returns on the stochastically detrended short-term interest rate.*

$$r_{t+1} + \cdots + r_{t+K} = \beta(K)(y_{1,t} - \textstyle\sum_{i=0}^{11} y_{1,t-i}/12) + \eta_{t+K,K}$$

	Forecast Horizon (K)					
	1	3	12	24	36	48
1927 to 1994						
$\hat{\beta}(K)$	−5.468	−17.181	−41.663	−4.492	−26.148	−20.129
$R^2(K)$	0.005	0.016	0.023	0.000	0.004	0.002
$t(\hat{\beta}(K))$	−2.292	−2.582	−1.564	−0.164	−1.341*	−0.838*
1927 to 1951						
$\hat{\beta}(K)$	3.144	−6.183	73.712	158.989	−67.505	−50.900
$R^2(K)$	0.000	0.000	0.012	0.031	0.005	0.002
$t(\hat{\beta}(K))$	0.222	−0.165	0.520	1.662	−0.637*	−0.580*
1952 to 1994						
$\hat{\beta}(K)$	−6.547	−18.621	−56.406	−26.115	−26.573	−25.894
$R^2(K)$	0.019	0.047	0.103	0.013	0.010	0.008
$t(\hat{\beta}(K))$	−3.263	−3.206	−2.741	−1.354	−1.555*	−1.092*

r is the log real return on a value-weighted index of NYSE, AMEX, and NASDAQ stocks. $y_{1,t}$ is the 1-month nominal Treasury bill rate. Regressions are estimated by OLS, with Hansen and Hodrick (1980) standard errors, calculated from equation (A.3.3) in the Appendix setting autocovariances beyond lag $K-1$ to zero. Newey and West (1987) standard errors, with $q = (K-1)$, are used when the Hansen and Hodrick (1980) covariance matrix estimator is not positive definite. The cases where this occurs are marked $*$.

forecasting power of the dividend-price ratio. The postwar R^2 statistics are comparable to those in Table 7.1 at horizons of one or three months, but they peak at 0.10 at a horizon of one year and then rapidly decline. The regression t-statistics are likewise insignificant beyond a one-year horizon.

How can we understand the hump-shaped pattern of R^2 statistics and t-statistics in Table 7.2 and the strongly increasing pattern in Table 7.1? At one level, the results in Table 7.1 can be understood by recalling the formula relating the log dividend-price ratio to expectations of future returns and dividend growth rates, given above as (7.1.24):

$$d_t - p_t \;=\; \mathrm{E}_t\left[\sum_{j=0}^{\infty} \rho^j \left[-\Delta d_{t+1+j} + r_{t+1+j}\right]\right].$$

This expression shows that the log dividend-price ratio will be a good proxy for market expectations of future stock returns, provided that expectations of future dividend growth rates are not too variable. Moreover, in general