Financial Econometrics
Predictability of Asset Returns

Nour Meddahi
Toulouse School of Economics

Winter 2025

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:

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

Recent extension to the multivariate case by Hong, Linton and Zhang (2015).

- Box-Pierce Type Tests:

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

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

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

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

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

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.

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

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.

Serial Correlation in the Disturbances: The HAC Estimator

While one does not use the GLS estimator when Ω is unknown, one has to estimate consistently $Var[\hat{\beta}^{OLS}]$. Under heteroskedasticity, one should use the Eicker-White estimator. However, the Eicker-White estimator is not consistent when the disturbances u_t are serially correlated. There are two leading examples:

1) Multi-horizon forecasting: $r_{t+1:t+k} = x_t'\beta + \varepsilon_{t+k}$. Due to the overlapping of periods, the disturbances ε_{t+k} are correlated. The OLS estimator is still consistent, biased in finite sample and not asymptotically. We need to estimate $Var[\hat{\beta}]$.

2) We want to estimate the mean of the short term interest rate r_t, $\bar{r}$, and a variance of $\bar{r}$. The problem is that the short term interest rate is highly correlated with unknown correlation (if we do not specify a model).

Let us focus on the second example.

$$Var[\bar{r}] = Var\left[\frac{1}{n}\sum_{t=1}^{n} r_t\right] = \frac{1}{n^2}\sum_{1\leq i,j\leq n} Cov[r_i, r_j]$$
$$= \frac{1}{n}Var[r_t] + \frac{2}{n}\sum_{l=1}^{n-1}\left(1-\frac{l}{n}\right)Cov[r_t, r_{t+l}],$$

under the assumption $E[r_i] = E[r_{i+h}]$ and $Cov[r_i, r_j] = Cov[r_{i+h}, r_{j+h}]$ for any i, j, h. In this case, we will say that the process r_t is a second order stationary process.

One can show that

$$lim_{n\to\infty} nVar[\bar{r}] = Var[r_t] + 2\sum_{l=1}^{\infty} Cov[r_t, r_{t+l}].$$

A potential estimator of $Var[\sqrt{n}\bar{r}]$ is

$$\hat{V}ar[r_t] + 2\sum_{l=1}^{\infty} \hat{C}ov[r_t, r_{t+l}],$$

where

$$\hat{C}ov[r_t, r_{t+l}] = \frac{1}{n-k}\sum_{t=1}^{n-l}(r_t - \bar{r})(r_{t+l} - \bar{r}).$$

There are three problems.

1. We have finite sample, we will not be able to estimate an infinite number of parameters.
2. We should estimate a small number of parameters, otherwise the quality of the estimators is poor.
3. We have to be sure that the estimator is positive (univariate case) or positive definite (regression case).

A solution has been proposed by Newey and West. They show that the following estimator is positive and consistent (under some assumptions)

$$\hat{V}ar[\sqrt{n}\bar{r}] = \hat{V}ar[r_t] + 2\sum_{l=1}^{L}\left(1 - \frac{l}{L}\right)\hat{C}ov[r_t, r_{t+l}].$$

- Such estimator is called a Heteroskedasticity and Autocorrelation Consistent (HAC) estimator of the standard errors.
- The parameter L is called the truncation parameter of the HAC estimator.
- L must be chosen such that it is large in large samples, although still much less than n.
- A good guideline is $L = 0.75n^{1/3}$.

In the regression case, form the formula $\hat{\beta} = \beta + [\mathbf{X'X}]^{-\mathbf{1}}\mathbf{X'}\varepsilon$, one gets

$$\begin{aligned} Var[\hat{\beta} \mid X] &= [\mathbf{X'X}]^{-\mathbf{1}} Var[\mathbf{X'}\varepsilon][\mathbf{X'X}]^{-\mathbf{1}} \\ &= \left[\sum_{t=1}^{n} \mathbf{x_t x_t'}\right]^{-1} Var\left[\sum_{t=1}^{n} x_t \varepsilon_t\right] \left[\sum_{t=1}^{n} \mathbf{x_t x_t'}\right]^{-1} \\ &= \frac{1}{n}\left[\frac{1}{n}\sum_{t=1}^{n} \mathbf{x_t x_t'}\right]^{-1} Var\left[\frac{1}{\sqrt{n}}\sum_{t=1}^{n} x_t \varepsilon_t\right] \left[\frac{1}{n}\sum_{t=1}^{n} \mathbf{x_t x_t'}\right]^{-1}. \end{aligned}$$

The Newey and West estimator of $Var\left[\sqrt{n}\sum_{t=1}^{n} x_t\varepsilon_t\right]$ is given by

$$\hat{\Sigma}_{x\varepsilon} = \hat{V}ar[x_t\varepsilon_t] + \sum_{l=1}^{L}\left(1 - \frac{l}{L}\right)\left(\hat{C}ov[x_t\varepsilon_t, x_{t+l}\varepsilon_{t+l}] + \hat{C}ov[x_t\varepsilon_t, x_{t+l}\varepsilon_{t+l}]'\right),$$

where

$$\hat{C}ov[x_t\varepsilon_t, x_{t+l}\varepsilon_{t+l}] = \frac{1}{n-k}\sum_{t=1}^{n-l}(x_t\varepsilon_t - \bar{x\varepsilon})(x_{t+l}\varepsilon_{t+l} - \bar{x\varepsilon})'.$$

Then, a positive definite estimator of the variance of $\hat{\beta}$ is

$$Var[\hat{\beta}] = \frac{1}{n}\left[\frac{1}{n}\sum_{t=1}^{n}\mathbf{x_t}\mathbf{x_t'}\right]^{-1}\hat{\Sigma}_{x\varepsilon}\left[\frac{1}{n}\sum_{t=1}^{n}\mathbf{x_t}\mathbf{x_t'}\right]^{-1}.$$

Observe that the Eicker-White estimator is a special case of the HAC estimators.

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.

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

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

	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 $*$.