Financial Econometrics | Lecture 8

Time Series and Forecasting: What Can and Cannot Be Predicted

In Class Group Activity

· SJTU, Summer School 2026 · Fatih Kansoy

How this block runs. You work in your home group of three or four, textbook open. Roles rotate: **Scribe**, **Reporter**, **Sceptic**, **Timekeeper**. All money today is in US dollars ($), pounds sterling (£) and euros (€).

- **Phase 1, Concept checks (~10 min).** Vote, talk for 90 seconds, vote again.
- **Phase 2, Desk problem (~20 min).** Forecast with an AR(1), with an interval.
- **Phase 3, Case (~25 min).** One real story, ending hidden.

A short word list. The **autocorrelation function (ACF)** gives the correlation of a series with its own past at each lag. An **AR(1)** is $x_t = c + \phi x_{t-1} + u_t$; ϕ measures **persistence**, the long-run mean is $c/(1-\phi)$. **BIC** and **AIC** choose the lag length, BIC the more parsimonious. The **efficient-markets** result is that returns are unforecastable (AR(1) near zero), so a low R^2 here is the *correct* answer. **Volatility clustering** means returns have a near-zero ACF but squared returns are persistent: risk is forecastable even when the return is not. **Granger** prediction is not causation. Overlapping or serially correlated data need **HAC** standard errors.

Phase 1 – Concept checks

~10 MIN

ConcepTest 1. You fit an AR(1) to monthly US excess stock returns and get $\widehat{\phi} = 0.08$, $t = 1.4$ ($p = 0.17$), $R^2 = 0.007$. A classmate says "this model is useless, the R^2 is basically zero, throw it out." The best reading is:

(a) The model is broken; a 0.7% R^2 means you specified it wrong, so add more lags until R^2 rises.

(b) The result is correct and informative: a near-zero, insignificant slope and a tiny R^2 are exactly what market efficiency predicts. The finding is "returns are unforecastable", not "the model failed".

(c) Because $\widehat{\phi}$ is not exactly zero, returns *are* forecastable; trade on last month's return.

(d) The low R^2 proves the market is irrational, because a rational market would be predictable.

ConcepTest 2. For the same returns, the autocorrelations of the returns are $\rho_{1..4} = 0.08, -0.01, -0.10, 0.02$ (near zero), but the autocorrelations of the *squared* returns are large and positive at every lag, decaying slowly. Which is right?

(a) A contradiction: if the returns have no autocorrelation, their squares cannot either.

(b) The squared-return result must be a coding error, because returns are unforecastable.

(c) No contradiction: the *direction* of the return is unforecastable, but the *size* is forecastable. Calm follows calm and turbulence follows turbulence: volatility clustering, so risk is forecastable even when the return is not.

(d) It means you can forecast next month's return after all, using squared returns.

ConcepTest 3. Regressing next year's industrial-production growth on today's term spread gives a slope of 1.19, HAC $t = 3.3$, $R^2 = 0.08$, and an inverted curve has preceded almost every US recession since 1960. A newspaper writes "the bond market causes recessions." The correct reading is:

(a) Correct: a significant predictor is a cause.

(b) The spread Granger-*predicts* output: its past helps forecast future growth. That is predictability, not causation; the curve anticipates the economy rather than driving it, and it gives false signals too.

(c) Because it has given false signals, the predictor is worthless and should be ignored.

(d) The high t proves there is no omitted third factor, so the link must be causal.

Phase 2 – The desk problem

~20 MIN

Forecasting a retailer's sales growth

The setting. You have 240 months of a German grocery retailer's annualised like-for-like sales growth g_t (%). OLS gives the AR(1)

$$g_t = 1.20 + 0.60\, g_{t-1} + u_t, \qquad t_{\text{slope}} = 11.5,\ R^2 = 0.36,\ SER = 2.00.$$

The latest value is $g_T = 4.00$.

(a) One-step forecast. Compute $\widehat{g}_{T+1}$. Is it above or below the last value? Why?

(b) Long-run mean and persistence. Compute the long-run mean $c/(1-\phi)$ and the half-life of a shock, $\ln(0.5)/\ln(\phi)$. In one sentence, how much memory does this series have?

(c) Forecast interval. Build the 95% one-step interval, $\widehat{g}_{T+1} \pm 1.96\, SER$. Why is it wide, and why is even this interval slightly too *narrow*?

(d) Lag length. You refit AR(1) to AR(4):

	AR(1)	AR(2)	AR(3)	AR(4)
BIC	512.0	508.5	507.0	509.0
AIC	505.0	498.0	493.5	492.5

Which lag length does each criterion choose? Why is BIC more parsimonious?

(e) The efficient-markets contrast. You fit the same AR(1) to a stock's monthly returns and get $r_t = 0.50 + 0.03\, r_{t-1} + u_t$, $t = 0.6$, $R^2 = 0.001$, mean return 0.52. If last month was $r_T = 6.00\%$, what is the best forecast of next month's return? Why is that essentially the mean, and why is that *good* news?

Phase 3 – The case

~25 MIN

The AlphaSignal letter: a backtest too good to be true

How to run this. Read the story with your group. The ending is hidden on purpose. Commit a written position before the reveal. Do not search the internet.

The story. It is early 2007. A London quant, Daniel Roche, launches a paid newsletter, **AlphaSignal**. His pitch: he has found a leading indicator that forecasts the FTSE 100. Each month he regresses the FTSE's next-12-month total return on a single predictor he builds from macro data, using monthly data 1985 to 2006 (about 260 months). His headline regression: slope +2.4, ordinary $t = 4.8$, $R^2 = 0.41$ ("explains 41% of the market"), and the fit line tracks the FTSE beautifully. Subscribers pour in at £500 a year. A pension consultant asks his junior analyst to check three things: (1) Roche's predictor is itself a slow-moving, trending macro series that looks like a random walk; (2) the dependent variable is a 12-month-ahead *overlapping* return, and the standard errors are ordinary OLS, not HAC; (3) Roche only ever reports *in-sample* fit, with no out-of-sample test.

The story stops here.

Four discussion questions (answer all four; take a position):

1. Roche advertises $R^2 = 0.41$ and $t = 4.8$. Name **two** reasons these in-sample numbers can be misleading here.
2. The analyst recomputes the t with HAC standard errors and gets 1.5. Does the point estimate (2.4) change? What changes, and why does it matter?
3. Roche says "twenty years and never wrong." Why is a perfect *in-sample* record almost worthless as evidence? Name the one test that would settle it.
4. **Commit.** Should the consultant recommend AlphaSignal to clients? State your position and your single strongest reason.

End of activity handout. Solutions are distributed after the block.