F500 Problem Set 3 - Solutions

Dr Tom Auld*

March 11, 2019

Question 1

The model is

$$
\begin{aligned}
r_{t+1} &= x_t + \varepsilon_{t+1} \\
x_{t+1} &= \mu + \phi x_t + \xi_{t+1}
\end{aligned}
$$

The innovations ε_t and ξ_t are i.i.d., mutually independent and mean zero. $|\phi| < 1$. x_t is a stationary AR(1) process about a non-zero mean.

Moments:

$$E(r_t) = E(x_t) + E(\varepsilon_t) = E(x_t)$$

but

$$E(x_{t+1}) = \mu + \phi E(x_t) + E(\xi_{t+1}) = \mu + \phi E(x_t)$$

Since x_t is stationary $E(x_t) = E(x_{t+1}) \Rightarrow$

$$E(r_t) = E(x_t) = \frac{\mu}{1-\phi}$$

Variance:

$$VAR(r_t) = VAR(x_t) + VAR(\varepsilon_{t+1}) = VAR(x_t) + \sigma_\varepsilon^2 \qquad \text{since } \varepsilon_t \text{ and } \xi_t \text{ are i.i.d.}$$

but

$$
\begin{aligned}
VAR(x_{t+1}) &= VAR(\mu + \phi x_t + \xi_{t+1}) \\
&= VAR(\phi x_t) + VAR(\xi_{t+1}) + 2COV(\phi x_t, \xi_{t+1}) \\
&= \phi^2 VAR(x_t) + \sigma_\xi^2 \qquad \text{because } \xi_{t+1} \text{ occurs "after" } x_t
\end{aligned}
$$

again stationarity of $x_t \Rightarrow$

$$VAR(x_t) = \frac{\sigma_\xi^2}{1-\phi^2}$$

*Comments and corrections to tja20@cam.ac.uk

so

$$VAR(r_t) = \frac{\sigma_\xi^2}{1-\phi^2} + \sigma_\varepsilon^2$$

Conditional Moments:

$$\begin{aligned} E_t(r_t|x_t) &= E_t(x_t|x_t) + E_t(\varepsilon_{t+1}|x_t) = x_t \\ VAR(r_t|x_t) &= VAR(x_t + \varepsilon_{t+1}|x_t) = VAR(\varepsilon_{t+1}|x_t) = \sigma_\varepsilon^2 \end{aligned}$$

ACF:

$$COV(r_t, r_{t-k}) = COV(x_{t-1} + \varepsilon_t, x_{t-1-k} + \varepsilon_{t-k}) = COV(x_t, x_{t-k})$$

since ε_t is i.i.d and independent of x_t. Expand x_t is terms of past innovations x_{t-k}:

$$\begin{aligned} x_t &= \mu + \phi x_{t-1} + \xi_t \\ &= \text{constant} + \phi(\phi x_{t-2} + \xi_{t-1}) + \xi_t \\ &= \text{constant} + \phi^2 x_{t-2} + \phi \xi_{t-1} + \xi_t \\ &\qquad\vdots \qquad\qquad \text{iterating } k \text{ times} \\ &= \text{constant} + \phi^k x_{t-k} + \phi^{k-1}\xi_{t-k+1} + \cdots + \xi_t \end{aligned}$$

so

$$COV(r_t, r_{t-k}) = COV(x_t, x_{t-k}) = COV(\text{constant} + \phi^k x_{t-k} + \phi^{k-1}\xi_{t-k+1} + \cdots + \xi_t, x_{t-k})$$

the ξ_{t-l} terms all occur "after" x_{t-k} so are independent and thus vanish $\Rightarrow$

$$COV(r_t, r_{t-k}) = \phi^k VAR(x_{t-k}) = \phi^k \cdot \frac{\sigma_\xi^2}{1-\phi^2}$$

dividing through by the variance yields

$$\rho_k = \phi^k \cdot \frac{\sigma_\xi^2}{\sigma_\xi^2 + (1-\phi^2)\sigma_\varepsilon^2}$$

When returns are slowly varying shocks die out slowly and $\phi \approx 1$ and

$$\rho_k \approx \phi^k$$

This is NOT consistent with the empirical evidence of very small and often statistically insignificant autocorrelations.

Question 2

The bubble process, whilst in effect, has exponential growth at a rate of $(1+R)/\pi$ but collapses each period with probability $1-\pi$. The average duration of the bubble is $(1-\pi)^{-1}$. As we saw in lectures this is not an irrational process whilst expectations of future growth at this rate persist.

Moments (we denote the information set at time t by $\mathcal{F}_t$):

$$E_t(B_{t+1}) = \pi \cdot E_t(\frac{1+R}{\pi}B_t + \eta_{t+1}|\mathcal{F}_t) + (1-\pi)\cdot E_t(\eta_{t+1}|\mathcal{F}_t) = (1+R)B_t$$

$$\begin{aligned}
E_t(B_{t+1}^2) &= \pi \cdot E_t\left[(\frac{1+R}{\pi}B_t + \eta_{t+1})^2|\mathcal{F}_t\right] + (1-\pi)\cdot E_t\left[(\eta_{t+1})^2|\mathcal{F}_t\right] \\
&= \pi E_t\left[(\frac{1+R}{\pi}B_t)^2|\mathcal{F}_t\right] + \pi E_t\left[2\frac{1+R}{\pi}B_t\eta_{t+1}|\mathcal{F}_t\right] + \pi E_t\left[\eta_{t+1}^2|\mathcal{F}_t\right] + (1-\pi)\cdot E_t\left[(\eta_{t+1})^2|\mathcal{F}_t\right] \\
&= \frac{(1+R)^2}{\pi}B_t^2 + 1
\end{aligned}$$

so

$$\begin{aligned}
VAR_t(B_{t+1}) &= E_t(B_{t+1}^2) - E_t(B_{t+1})^2 \\
&= \frac{(1+R)^2}{\pi}B_t^2 + 1 - (1+R)^2B_t^2 \\
&= 1 + \frac{(1-\pi)}{\pi}(1+R)^2B_t^2
\end{aligned}$$

This conditional variance is the variance from one period to the next given the current value of the bubble B_t and is higher when the probability of collapse $(1-\pi)$ is higher. This is because, for that one period, there is greater variability if the bubble does indeed collapse.

The probability that the bubble lasts for more than 5 periods is the probability that the first 6 observations retain the bubble. This is analogous to tossing a biased coin 6 times in a row where the probability of heads (or bubble) is π. The probability is of course π^6.

Testing for the presence of a bubble. Notice:

$$\triangle P_t = u_t + \triangle B_t$$

In the *absence* of a bubble $\triangle P_t = u_t + \eta_t - \eta_{t-1}$. This is a stationary series with zero autocovariance at lags above 1 (and zero everywhere if $VAR(\eta_t) = 0$). Autocorrelations can be tested in the usual way. If the bubble is in operation price returns are explosive and non-stationary. Stationarity can be tested using the KPSS test (look this up on Wikipedia for more information). If non-stationary we reject the NULL of no bubble.

Question 3

Mathematics aside, the point of this question is to contrast the properties of an MA(1) process with that of an AR(1) process, particularly in terms of the autocorrelation function of the squares of the return. The process is:

$$r_t = \varepsilon_t + \theta\varepsilon_{t-1} \qquad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2) \qquad |\theta| < 1$$

$$\begin{aligned} COV(r_t^2, r_{t-1}^2) &= COV(\varepsilon_t^2 + 2\theta\varepsilon_t\varepsilon_{t-1} + \theta^2\varepsilon_{t-1}^2, \varepsilon_{t-1}^2 + 2\theta\varepsilon_{t-1}\varepsilon_{t-2} + \theta^2\varepsilon_{t-2}^2) \\ &= \theta^2 VAR(\varepsilon_{t-1}^2) + COV(2\theta\varepsilon_t\varepsilon_{t-1}, \varepsilon_{t-1}^2) + \\ &\qquad COV(2\theta\varepsilon_t\varepsilon_{t-1}, 2\theta\varepsilon_{t-1}\varepsilon_{t-2}) + COV(\theta^2\varepsilon_{t-1}^2, 2\theta\varepsilon_{t-1}\varepsilon_{t-2}) \end{aligned} \tag{1}$$

since ε_t is i.i.d. and we can ignore covariances which do not share the same ε_{t-s} type terms. Also

$$COV(\varepsilon_t\varepsilon_{t-1}, \varepsilon_{t-1}^2) = E(\varepsilon_t\varepsilon_{t-1}^3) - E(\varepsilon_t\varepsilon_{t-1})E(\varepsilon_{t-1}^2) = E(\varepsilon_t)E(\varepsilon_{t-1}^3) - E(\varepsilon_t)E(\varepsilon_{t-1})E(\varepsilon_{t-1}^2) = 0$$

Similarly all terms except $VAR(\varepsilon_{t-1}^2)$ in equation 1 vanish as expectations factorise into a $E(\varepsilon_{t-s})$ type term which is zero. For a normal distribution $VAR(\varepsilon_t^2) = 2VAR(\varepsilon_t)^2 = 2\sigma_\varepsilon^4$ so

$$COV(r_t^2, r_{t-1}^2) = 2\theta^2\sigma_\varepsilon^4$$

For the variance note r_t is the sum of 2 uncorrelated normal distributed random variables (ε_t and ε_{t-1}) so it itself normal. It has variance $(1+\theta^2)\sigma_\varepsilon^2$. Thus

$$VAR(r_t^2) = 2VAR(r_t)^2 = 2(1+\theta^2)^2\sigma_\varepsilon^4$$

Noting that $COV(r_t^2, r_{t-s}^2) = 0 \; \forall s > 1$ since then r_t^2 and r_{t-s}^2 share no similar ε_t terms. Then the result is:

$$CORR(r_t^2, r_{t-k}^2) = \begin{cases} \frac{\theta^2}{(1+\theta^2)^2} & k = 1 \\ 0 & k > 1 \end{cases}$$

as required.

For the AR(1) process the ACF of returns die of at a rate of ρ whereas the ACF of r_t^2 die of at the faster rate of ρ^2. The ACF for the MA(1) model is $\frac{\theta}{(1+\theta^2)}$ at lag 1 and zero elswhere. The ACF of r_t^2 is also non-zero only at lag one. However, since $\theta < 1$ the magnitude of $CORR(r_t^2, r_{t-1}^2) = \frac{\theta^2}{2(1+\theta^2)^2} < \frac{\theta}{(1+\theta^2)}$ is also smaller than that of the ACF of r_t. Thus the MA(1) model does not solve the issues that AR(1) has, it just limits the significant part of the ACFs to lag 1. A better description of the ACFs observed in financial time series (statistically significant positive serial correlation in squared returns with little or no serial correlation in return) is given by the GARCH model (Question 5).

Question 4

This question concerns the estimation of contemporaneous correlation between stocks that trade at different times - the parameter of interest is the covariance γ (or in practice, $3\times\gamma$ for the 24 hour measure). The trading day is split into thirds and we observe only one price per day for the 2 assets (which are at different times of the day). The assumptions are that the stocks are contemporaneously correlated and have i.i.d. returns. For the purposes of answering this question this means that returns from one 1/3 of a day to the next are independent. This also implies that the return from stock i is independent from the return from stock j during different non-overlapping periods. Assuming otherwise would contradict the i.i.d. of the individual stock returns.

Proceed by splitting up the relevant daily price differences into intra-period price differences:

$$\begin{aligned}
COV(p_{i4}-p_{i1},p_{j5}-p_{j2}) &= COV(p_{i4}-p_{i1},p_{j5}-p_{j2}) \\
&= COV(\,(p_{i4}-p_{i2})+(p_{i2}-p_{i1})\,,\,(p_{j5}-p_{j4})+(p_{j4}-p_{j2})\,) \\
&= COV((p_{i4}-p_{i2}),(p_{j4}-p_{j2}))
\end{aligned}$$

using the fact that non-overlapping returns are independent. Continuing:

$$\begin{aligned}
COV(p_{i4}-p_{i1},p_{j5}-p_{j2}) &= COV(p_{i4}-p_{i2},p_{j4}-p_{j2}) \\
&= COV(\,(p_{i4}-p_{i3})+(p_{i3}-p_{i2})\,,\,(p_{j4}-p_{j3})+(p_{j3}-p_{j2})\,) \\
&= COV(p_{i4}-p_{i3},p_{j4}-p_{j3})+COV(p_{i3}-p_{i2},p_{j3}-p_{j2}) \\
&= 2COV(p_{i2}-p_{i1},p_{j2}-p_{j1}) \\
&= 2\gamma \qquad (2)
\end{aligned}$$

because the covariances are stationary and do not very from one 8 hour period to the next.

Since i trades for the first third of the day and j trades for the second half of the day, daily returns for day t are defined as follows:

$$\begin{aligned}
r_{it} &= p_{i,3t+1}-p_{i,3t-2} \\
r_{jt} &= p_{j,3t+2}-p_{j,3t-1}
\end{aligned}$$

So (generalizing the result in equation 2 above)

$$\begin{aligned}
COV(r_{it},r_{jt}) &= COV(p_{i,3t+1}-p_{i,3t-2},p_{j,3t+2}-p_{j,3t-1}) \\
&= 2\gamma
\end{aligned}$$

Calculating $COV(r_{it},r_{j,t-1})$ by again splitting up into overlapping returns and ignoring terms involving non-overlapping periods:

$$\begin{aligned}
COV(r_{it},r_{j,t-1}) &= COV(p_{i,3t+1}-p_{i,3t-2},p_{j,3t-1}-p_{j,3t-4}) \\
&= COV(p_{i,3t-1}-p_{i,3t-2},p_{j,3t-1}-p_{j,3t-2}) \\
&= \gamma
\end{aligned}$$

So the observed daily return of stock i is both correlated with the daily return of stock j as well as the previous day's return of that stock. Either of $\frac{1}{2}\widehat{COV}(r_{it},r_{jt})$ or $\widehat{COV}(r_{it},r_{j,t-1})$ are consistent estimators of the 8 hour contemporaneous correlation γ of the stocks but the estimator:

$$\hat{\gamma} = \frac{1}{3}\widehat{COV}(r_{it},r_{jt}) + \frac{1}{3}\widehat{COV}(r_{it},r_{j,t-1})$$

is more efficient as it uses both information from the correlation of same day stock returns as well as that from j's one day lagged return.

Question 5

The GARCH model was developed to explain both the fat tails of the unconditional return distribution and serial correlation in squared returns observed in financial time series.

$$
\begin{aligned}
r_t &= h_t^{1/2}\eta_t \\
h_t &= \omega + \beta h_{t-1} + \gamma r_{t-1}^2
\end{aligned}
$$

First notice

$$E(r_t^2|h_t) = h_t$$

so h_t is of course the conditional variance. Expanding:

$$
\begin{aligned}
h_t &= \omega + \gamma r_{t-1}^2 + \beta h_{t-1} \\
&= \omega + \gamma r_{t-1}^2 + \beta(\omega + \gamma r_{t-2}^2 + \beta h_{t-2}) \\
&= \omega(1+\beta) + \gamma(r_{t-1}^2 + \beta r_{t-2}^2) + \beta^2 h_{t-2} \\
&\qquad\vdots \qquad\qquad \text{iterating } s \text{ times} \\
&= \omega \sum_{s=1}^{n} \beta^s + \gamma \sum_{s=1}^{n} \beta^{s-1} r_{t-s}^2 + \beta^n h_{t-n} \\
&= \frac{\omega}{1-\beta} + \gamma \sum_{s=1}^{n} \beta^{s-1} r_{t-s}^2 + lim_{n\to\infty}\beta^n h_{t-n} \\
&= \frac{\omega}{1-\beta} + \gamma \sum_{s=1}^{n} \beta^{s-1} r_{t-s}^2 \qquad\qquad \text{if } |\beta| < 1
\end{aligned}
$$

As in lectures h_t can be expanded as an EMA of past squared returns with decay factor β plus a level $(\frac{\omega}{1-\beta})$. This gives an intuitive explanation as to why volatility in this model clusters as well as the fact that variance is expected to be higher after a “big move”.

a) We require h_t to be stationary for the variance to exist. We also need to ensure that h_t is non-negative with probability 1. The latter requires that $\omega, \beta, \gamma \geq 0$ (assuming general support for η_t). From the above expansion of h_t as an EMA of past squared returns it is apparent that β<1 is a condition of stationarity. The stronger condition $\beta + \gamma < 1$ is also required to ensure stationarity. This can be demonstrated by expanding future expectations in terms of today's squared return and conditional volatility:

$$
\begin{aligned}
E_t(h_{t+k}|r_t,h_t) &= E_t\left(\omega+\gamma r^2_{t+k-1}+\beta h_{t+k-1}\,|r_t,h_t\right)\\
&= E_t\left(\omega+(\gamma+\beta)h_{t+k-1}\,|r_t,h_t\right) && \because E_t(r^2_{t+k-1})=E_t(h_{t+k-1})\\
&= E_t\left(\omega+(\gamma+\beta).\,(\omega+\gamma r^2_{t+k-2}+\beta h_{t+k-2})\,|r_t,h_t\right)\\
&= E_t\left(\omega(1+\gamma+\beta)+(\gamma+\beta)^2.\,h_{t+k-1})\,|r_t,h_t\right) && \because E_t(r^2_{t+k-2})=E_t(h_{t+k-2})\ k>2\\
&\quad\vdots && \text{iterating } k \text{ times}\\
&= E_t\left[\omega\left(\sum_{s=0}^{k}(\gamma+\beta)^s\right)+\left((\gamma+\beta)^{k-1}.\,(\gamma r_t^2+\beta h_t)\right)\,|r_t,h_t\right]\\
&= \omega\left(\frac{1-(\gamma+\beta)^{k+1}}{1-\gamma-\beta}\right)+\left(\boldsymbol{(\gamma+\beta)^{k-1}}.\,(\gamma r_t^2+\beta h_t)\right) && \text{using}\sum_{r=0}^{r=k}x^r=\frac{1-x^{k+1}}{1-x}
\end{aligned}
\tag{3}
$$

So if $\gamma+\beta\geq 1$ today observed squared returns will have an explosive and persistent impact on future volatility. This explains the stationarity condition.

b)

$$
\begin{aligned}
E(h_t) &= E(\omega+\beta h_{t-1}+\gamma r^2_{t-1})\\
&= \omega+\beta E(h_{t-1})+\gamma E(r^2_{t-1})\\
&= \omega+\beta E(h_t)+\gamma E(h_t) && E(r^2_{t-1})=E(h_{t-1})=E(h_t)
\end{aligned}
$$

so

$$
E(h_t)=\frac{\omega}{1-\beta-\gamma}
$$

c)

ω affects the long run level of the variance but does not affect the persistence of any return shock. From equation 3 we see that the sum $\gamma+\beta$ is the rate of decay of the effect of r_t^2 shocks on future expectations of volatility. γ has a greater effect on the level than β as

$$
E_t(h_{t+k}|r_t,h_t)=\cdots+\gamma\times(\gamma+\beta)^{k-1}\times r_t^2+\cdots
$$