Sheet 8

I don't get why this term doesn't imply that as $t \to \infty$, $\Delta X_t \to \infty$. There must be some cancelling-out happening?

I think it does.

1a) $\Delta X_t = \beta_0 + \alpha t + \delta X_{t-1} + u_t$

~~If $X_t \sim I(1)$, it's the cumulation of a stationary series. So then $\Delta X_t := X_t - X_{t-1} = \beta_0 + \alpha t + (\beta_1 - 1)X_{t-1} + u_t$ will have $\delta := \beta_1 - 1 = 0$ since $\{X_t\}$ is $AR(p)$ with unit root $\Leftrightarrow$ $\{\Delta X_t\}$ is $AR(p-1)$.~~

~~So the relevant $H_0: \delta = 0$, $H_1: \delta < 0$.~~

Let X_t be an $AR(p)$ process, i.e. $X_t = \beta_0 + \alpha t + \beta_1 X_{t-1} + \dots + u_t$

Then $\Delta X_t := X_t - X_{t-1} = \beta_0 + \alpha t + \delta X_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta X_{t-i} + u_t$

and define $\delta := \sum \beta_i - 1$.

The slides sum up to p, but the notes only to $p-1$, in $\sum_{i=1} \gamma_i \Delta X_{t-i}$. Seems like $\gamma_p = 0$?

we'll clear this up in class

So we can regress ΔX_t onto X_{t-1} and $p-1$ lags of ΔX_t, and obtain δ as above by OLS.

If our alternative $X_t \sim I(0)$ holds, then X_t itself is stationary. This implies that $\sum \beta_i < 1$ (by the requirements for stationarity of $AR(p)$ processes) and thus $\delta < 0$.

Conversely, if $X_t \sim I(1)$ then it has a unit root (exactly one), i.e. $\sum \beta_i = 1$ $\Leftrightarrow \delta = 0$.

So $H_0: \delta = 0$; $H_1: \delta < 0$. We can then calculate the test statistic $t := \frac{\hat{\delta}}{se(\hat{\delta})} \xrightarrow{d} DF_{tr}$ and find the appropriate c_α to compare it against, rejecting H_0 iff $t < c_\alpha$.

Great

b) Advantage: without including an αt term, we exclude the possibility under H_0 of $\{X_t\}$ containing a deterministic trend, because H_1 corresponds to stationarity, where by definition μ_t must be time-invariant. So we might incorrectly fail to reject H_0, (i.e. we increase our Type II errors) when $\{X_t\}$ does have drift.

Disadvantage: the DF_{tr} distribution is shifted left even further than ~~our~~ DF_{cn}, so the critical values are ~~smaller~~ more negative and thus we lose power.

c) You could use an information criterion such as BIC or AIC and find the p^* which minimises loss. Alternatively, you could use stepwise testing from p_{max} on the hypothesis $H_0: p_{max} = 0$, and then reducing p until you reject H_0.

d) AIC suggests Y_t is AR(2) and C_t is AR(3), since we want $\arg\min_p AIC(p)$. Stepwise testing at the 1% level gives us the same conclusion, looking at the t-prob column.

$c_{1\%} = -3.96$ and in both cases we fail to reject that X_t has a unit root but do reject that ΔX_t has a unit root, so both are $I(1)$, since we look for $t^{obs} < c_\alpha$.

e) i. We've already verified that $Y_t, C_t \sim I(1)$. So now we can proceed to ~~regressing Y_t onto C_t and~~ calculating the residual $\{\varepsilon_t\}$ in the model $Y_t = C_t + \varepsilon_t$. If Y_t and C_t really are cointegrated with $\theta = 1$, then ε_t will be $I(0)$, so we ~~can~~ should perform an ADF test to check this, with a null that $\varepsilon_t \sim I(1)$ because Y_t and C_t aren't cointegrated with $\theta = 1$ and thus ε_t inherited a stochastic trend.

ii. ~~The~~ Similarly to above, except we need to estimate $\hat{\theta}$ by OLS with the model $Y_t = \theta C_t + \varepsilon_t$ before computing residuals. For our ADF test, use the adjusted Engle-Granger c_α.

[Do you ever include lags of C_t / Y_t, in the regression being done here?]

→ Not for what we'll be concerned with.

2a) We have $\pi_t = \beta_0 + \beta_1 \pi_{t-1} + u_t$ and want to test $H_0: \beta_1 = 0$ vs $H_1: \beta_1 \neq 0$. Our t-statistic $t^{obs} = 21$.

What is the dist. of t under the null?

We wouldn't expect the $N(0,1)$ c_α to be appropriate here, because ~~our residuals~~ we're dealing with a unit root process and so CLTs

Hmm, so P_t is unit root, but likely not π_t, given this regression output.

do not apply. However, the t-statistic is extremely large so we can surely reject H_0 at the 5% level, whatever the limiting distribution.

b) $\pi_{2017Q2|2017Q1} = 0.68 + 0.84 \times 2 =$ 2.36%

Because β_1 is close to 1, our estimate $\hat{\beta}_1 = 0.84$ is heavily biased downwards, and this affects accuracy/reliability.

c) · Our β_is are now further from 1, so the bias should be reduced.

In principle, run an F test.

· However, it seems surprising that ~~the effect size is larger for~~ ~~than of~~ the second and third lags of $\Delta\pi_t$ have larger effects than the first lag, and that the direction of the effect switches. [Maybe interpreting this is finnicky because of the Δs though]

· The effect sizes are small relative to the SEs and some of the 95% CIs look likely to cross 0.

d) Note that we could also express her model as

$$\pi_t = 0.467 + 0.706\,\pi_{t-1} - 0.067\,\pi_{t-2} + 0.45\,\pi_{t-3} - 0.196\,\pi_{t-4} + \hat{u}_t$$

since $\Delta\pi_{t-i} := \pi_{t-i} - \pi_{t-i-1}$.

seems like we could potentially test here whether $\sum_i \beta_i = 1$, though calculating the variance of the sum would be hard because of dependence.

$\Delta\pi_t$ is AR(3) $\iff$ π_t is AR(4) and has a unit root. (*)

Suppose $\pi_t = \beta_0 + \beta_1\pi_{t-1} + \beta_2\pi_{t-2} + \beta_3\pi_{t-3} + \beta_4\pi_{t-4} + u_t$.

Then
$$\begin{aligned}\Delta\pi_t := \pi_t - \pi_{t-1} &= \beta_0 + (\beta_1 - 1)\pi_{t-1} + \beta_2(\pi_{t-1} - \Delta\pi_{t-1}) \\ &\quad + \beta_3(\pi_{t-1} - \Delta\pi_{t-1} - \Delta\pi_{t-2}) + \beta_4(\pi_{t-1} - \Delta\pi_{t-2} \cdots) \\ &\quad + u_t \\ &= \beta_0 + (\beta_1 + \beta_2 + \beta_3 + \beta_4 - 1)\pi_{t-1} + (-\beta_2 - \beta_3 - \beta_4)\Delta\pi_{t-1} \\ &\quad + (-\beta_3 - \beta_4)\Delta\pi_{t-2} - \beta_4\Delta\pi_{t-3} + u_t \\ &=: \beta_0 + \gamma\pi_{t-1} + \sum_{i=1}^{3}\gamma_i\Delta\pi_{t-i} + u_t.\end{aligned}$$

I guess this is really just a proof of (*)?

The unit root condition $\sum_{i=1}^{4}\beta_i = 1$ can thus be captured by

Looks right, (?)

Am confused here.

the statement $\delta = 0$. We're effectively performing an ADF test (?), and so have
$H_0 : \delta = 0$ vs $H_1 : \delta < 0$ corresponding to $\Delta\pi_t$ being AR(3) vs ???.

Our test statistic $t^{obs} = \frac{-0.107}{0.041} = -2.610$
And $t \xrightarrow{d} DF_{cn}$, so $c_{10\%} = -2.57$, $c_{5\%} = -2.86$.
(I think we're just doing a constant-only ADF test?)

We fail to reject H_0 at the 5% level.

e) We're interested in whether the parameters of our time series process might have changed as we entered/exited the 1979–1987 epoch. Specifically, perhaps more aggressive responsiveness in MP to inflation deviations meant that ~~previous period~~ lags of π_t were less useful in predicting π_t.

We could use the QLR test to investigate this: for a variety of $\tau \in \{1979Q3, \dots, 1987Q3\}$, fit models with added breakpoint dummies of the form $\gamma_0 D_t(\tau)$, $\gamma_i D_t(\tau)\pi_{t-i}$, where $D_t(\tau) = 1$ if $t \geq \tau$ and 0 otherwise. Then we'd (implicitly) perform an F test of whether some γ_i is nonzero for each choice of τ, and then check for significance at the τ which maximises $F(\tau)$. Rejecting H_0 would mean we had evidence that Volker's tenure led to structural changes in how inflation evolves.

We might also simply look at a plot of π_t and see whether there are any obvious changes to its variance, ACF, or mean in this epoch.

[Maybe you'd want to allow for multiple breakpoints, incl. some outside his tenure. But probably then you'd end up with extremely low power?]

3a) $Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + u_t$, where $\{u_t\}$ is stationary and unforecastable ~~white noise~~ with $\mu=0$

a) $\{Y_t\}$ has a unit root iff $\beta_1 + \beta_2 = 1$.

b) $\Delta Y_t := Y_t - Y_{t-1} = \beta_0 + (\beta_1 - 1) Y_{t-1} + \beta_2 (Y_{t-1} - \Delta Y_{t-1}) + u_t$

$$= \beta_0 + (\beta_1 + \beta_2 - 1) Y_{t-1} - \beta_2 \Delta Y_{t-1} + u_t$$

but since $\{Y_t\}$ is unit root, $\beta_1 + \beta_2 - 1 = 0$

$$= \beta_0 \underbrace{- \beta_2}\Delta Y_{t-1} + u_t =: \gamma_0 + \underbrace{\gamma_1}\Delta Y_{t-1} + u_t$$

[Hmm, this Q makes me think my answer to 2d) must've been needlessly long-winded]

c) If ΔY_t is stationary, then one differencing of Y_t was required to reach stationarity. So by definition, $Y_t \sim I(1)$, since it's the cumulation of an $I(0)$ process — namely, ΔY_t.

Note that given $|\gamma_1| < 1$, it straightforwardly follows that $\{\Delta Y_t\}$ is stationary, by the condition for AR(1) processes (with suitable conditions on ΔY_0), and vice versa. So we don't really need both parts of the "suppose" statement here.

d) $Y_t = Y_0 + \sum_{j=1}^{t} \Delta Y_j$ as stated above, and $\mathbb{E}[\Delta Y_t] =: \mu$ the unconditional, time-invariant mean.

$$= Y_0 + \mu t + \sum_{j=1}^{t} (\Delta Y_j - \mu)$$

$$=: Y_0 + \mu t + \sum_{s=1}^{t} v_s$$, with v_s inheriting stationarity from ΔY_t, and demeaned.

i. Since $\mathbb{E}[\Delta Y_t] = \mathbb{E}[\Delta Y_{t-1}]$, we have

$\mathbb{E}[\Delta Y_{t}] = \mathbb{E}[\gamma_0 + \gamma_1 \Delta Y_{t-1} + u_t]$

$\therefore \; \mathbb{E}[\Delta Y_{t-1}] = \gamma_0 + \gamma_1 \mathbb{E}[Y_{t-1}] + \cancel{\mathbb{E}[u_t]}^{\text{mean zero}}$ by linearity

$\therefore \; \mathbb{E}[\Delta Y_{t-1}] =: \mu = \dfrac{\gamma_0}{1-\gamma_1} = \dfrac{\beta_0}{1+\beta_2}$

ii. $v_t := \Delta Y_t - \mathbb{E}[\Delta Y_t]$ (is that it?)

$\quad = \gamma_0 + \gamma_1 \Delta Y_{t-1} + u_t - \cancel{\tfrac{\gamma_0}{1-\gamma_1}}\, \mu$, ~~where~~ and $\Delta Y_{t-1} = v_{t-1} + \mu$

so $v_t = \gamma_0 + \gamma_1 v_{t-1} + (\gamma_1 - 1)\mu + u_t$

$\quad = \gamma_1 v_{t-1} + u_t$

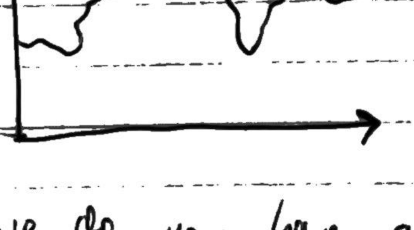
different in practice? (ofc uncorr $\not\Rightarrow$ indep., but does it matter?)

e) $\sum_{s=1}^{t} v_s$ is a random walk iff it's the sum of ~~independent~~, serially uncorrelated, mean zero stationary random variables. For this to be, we need $\operatorname{cov}(v_t, v_{t+h}) = 0$ for all h. (we know it's time-invariant already, by stationarity).

Recall that for an AR(1) stationary process $X_t = \delta_0 + \delta_1 X_{t-1} + \varepsilon_t$, the $ACF(h) = |\delta_1|^h$. In this case, then, we require $\gamma_1 = 0$, i.e. $\beta_2 = 0$.

This means that Y_t would in fact be AR(1), ΔY_t is AR(0), and our partial sum $\sum_{s=1}^{t} v_s$ is simply adding up white noise (since we demean any constant term anyway).

f) i. $\mu = 0 \iff \gamma_0 = 0 \iff \beta_0 = 0$. There's no deterministic trend ("drift") in Y_t and its mean remains constant over time.



It probably looks similar to a random walk, but smoother since we may have autocorrelation in shocks.

ii. We do now have a deterministic trend, as well as a stochastic one.

4a) They're both $I(1)$: we fail to reject H_0 of a unit root for the levels but do reject it (against H_1 of stationarity) for their first difference, at the 1% level with $c_\alpha = -3.96$ (constant + trend).

b) $\log(US/UK) - \log(US/EU) = \log(US/UK \div US/EU) = \log(EU/UK)$ which is the log of our rate of interest, €:£.

Seems plausible that we have deterministic drift in the rate over time due to real macro changes, so let's use the ADF_{tr} results. At the 5% level we cannot reject H_0 of a unit root in the levels ($-3.113 \nless -3.41 = c_\alpha$), but we can reject H_0 (at the 1% level, even) for the differenced series. DL~ - DL~

For these Qs, you could just pick $\alpha = 10\%$ and then argue the opposite...

So the Euro-Pound exchange rate is $I(1)$.

~~Normally when we test $I(d)$ of $Y_t - X_t$ we're looking to see whether they're cointegrated with~~ [This is different to an ADF test as part of a cointegration test, right? But there is a similarity in that we're sort of assuming $\theta = 1$, finding the residual - i.e. the €:£ rate - and seeing that it's not $I(0)$, so we fail to reject the null of no cointegration? Except surely \$:£ and \$:€ are!]

huh, actually now I think it isn't different: having done c)

Not sure what this means about MP independence, probably evidence for it? The €:£ rate wanders around, so BoE isn't pegging rates to Eurozone? (*)

(*) But couldn't you tell a story where lasting within-US shocks affect both USEU and UKUS rates? Then you'd expect to see cointegration, but it doesn't mean UK lacks independent MP, right?

c) No, these results are highly doubtful. They're much more likely to be due to a spurious correlation from the presence of ~~determinis~~ stochastic trends in both series, giving the illusion of a structural relationship when there in fact is none. This accounts for the high R^2 and t-statistic, which $\xrightarrow{P} \infty$ in large samples in θ, if we assume unity cointegration.

As noted, our residual is $I(1)$, so this is strong evidence of the regression relationship being spurious - otherwise, the stochastic trend would've been elimated through the regression! You could verify the same using the analyst's values of θ.

d) Since our differenced series have the stochastic trends stripped out, you could test for Granger causality between DLUSEU and DLUSUK.

Maybe you can also do something like holding out part of the data for validation, or doing regressions on several splits + comparing for robustness, since you might get very different effect sizes (+ directions) due to spuriousness.