Non-Stationary Time Series Notes

Structural Breaks

- A structural break occurs at time τ iff the parameters of the model change abruptly at τ. A structural break is a source of non-stationarity.

Chow Breakpoint Test

- Suppose that $\{Y_t\}$ follows some $ADL(p,q)$ model, then the Chow breakpoint test is a test for a structural break at τ.
- A hypothesis that a break occurs at τ can be informed by inspection of the time series data, or by knowledge of economic conditions proximate to that date.
- The breakpoint dummy variable given some hypothesised time τ is $D_t(\tau) := [t > \tau] = \begin{cases} 1 & \text{if } t > \tau \\ 0 & \text{if } t \le \tau \end{cases}$.
- Then, the Chow breakpoint test is a test of
    - $H_0 : \gamma_0 = \gamma_1 = \ldots = \gamma_p = \theta_1 = \ldots = \theta_q = 0$, against
    - $H_1 : \exists \gamma \in \{\gamma_0, \gamma_1, \ldots, \gamma_p, \theta_1, \ldots, \theta_q\} : \gamma \ne 0$, in
    - $$Y_t = \beta_0 + \sum_{i=1}^{p} \beta_i Y_{t-i} + \sum_{i=1}^{q} \delta_i X_{t-i}$$
      $$+ \gamma_0 D_t(\tau) + D_t(\tau) \sum_{i=1}^{p} \gamma_i Y_{t-i} + D_t(\tau) \sum_{i=1}^{q} \theta_i X_{t-i} \text{, where}$$
      $$+ u_t$$
    - $\mathbb{E}[u_t | \mathcal{Y}_{t-1}, \mathcal{X}_{t-1}] = 0$ by construction.
    - Supposing that $\{X_t\}$ and $\{Y_t\}$ are jointly stationary, the critical values for this test are drawn from the usual $F_{q,\infty}$ distribution for a test of q restrictions.
    - A rejection of the null can be interpreted as evidence of a structural break at τ, and that at least one of the parameters changes at τ.
- Suppose that the objective of modelling is to forecast Y_t. Then, if a structural break is detected at τ, split the data and estimate an $ADL(p,q)$ model on the most recent epoch. Data from the most recent epoch is more plausibly described by a stable model (than data from both epochs taken together), so a model fitted on data from only the most recent epoch is likely to have better forecasting performance.
    - Fitting the augmented (with the breakpoint dummy variable) model and estimating the optimal forecast using the estimated coefficients is equivalent.

Quandt Likelihood Ratio Test

- In practice, the precise time τ of a structural break is not known, hence the hypothesis of interest is not that a structural break occurred at some specific τ, but that a structural break occurred at some τ in some hypothesised interval $\{\tau_0, \tau_0 + 1, \ldots, \tau_1\}$.
    - A naive test of this hypothesis is to repeat the Chow breakpoint test for each $\tau \in \{\tau_0, \ldots, \tau_1\}$ and reject the null hypothesis that no structural break occurs in this interval if the null hypothesis of any of the Chow breakpoint tests is rejected.
    - This naive test is unsatisfactory because the significance of this test (i.e. the type I error rate, or the probability of rejecting a true null) is greater than the significance of each individual Chow test. When there is in fact no break, this test concludes that some break exists with greater probability than the significance level of each individual Chow test.
- The procedure for the Quandt likelihood ratio test is as follows.
    - Suppose that $\{Y_t\}$ follows some $ADL(p,q)$ model.
    - For each $\tau \in \{\tau_0, \ldots, \tau_1\}$, compute the F statistic $F(\tau)$ of the associated Chow breakpoint test.
    - The Quandt likelihood ratio statistic is $QLR := \max_{\tau \in \{\tau_0, \ldots, \tau_1\}} F(\tau)$.
    - Compare QLR against a suitable critical value.
        - This critical value depends on (1) the number q of parameters restricted by the null, (2) the number k of slope coefficients in the unrestricted model, and (3) the locations of the endpoints $\frac{\tau_0}{T}$ and $\frac{\tau_1}{T}$ relative to the length of the sample T.
        - The critical values of this test are larger than the critical values of the Chow breakpoint test, drawn from the $F_{q,\infty}$ distribution. This is because QLR is the maximum of several F statistics, so the probability that QLR is larger (in

magnitude) than the $F_{q,\infty}$ critical value is greater than the probability that any individual F statistic is greater than that critical value.
- The difference in magnitude of the Quandt likelihood ratio critical value and the $F_{q,\infty}$ critical value is increasing in the magnitude of $\tau_1 - \tau_0$.

Unit Roots

Unit Root Autoregressive Models

- An $AR(p)$ model $Y_t = \beta_0 + \sum_{i=1}^{p} \beta_i Y_{t-i} + u_t$ has a unit root iff $\sum_{i=1}^{p} \beta_i = 1$.
  - If $\{Y_t\}$ follows such a model, then $\{Y_t\}$ is a unit root autoregressive process.
  - Iff $\{Y_t\}$ is a unit root autoregressive process with order p, then $\{\Delta Y_t\}$ is an autoregressive process with order $p - 1$, i.e. $\{\Delta Y_t\}$ follows an $AR(p-1)$ model.
  - Note the technicality that having a unit root is a property of random variables, but that being a unit root autoregressive process or following a unit root autoregressive model is a property of series.
- A unit root $AR(1)$ model can be decomposed into (1) the initial value, (2) a deterministic trend, and (3) a stochastic trend.
  - A unit root $AR(1)$ model is some $Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t$, where $\mathbb{E}[u_t | \mathcal{Y}_{t-1}] = 0$ by construction, and where $\beta_1 = 1$.
  - By substitution, $Y_t = \beta_0 + Y_{t-1} + u_t$.
  - By recursive substitution, $Y_t = Y_0 + t\beta_0 + \sum_{s=1}^{t} u_s = Y_0 + t\beta_0 + U_t$, where $U_t := \sum_{s=1}^{t} u_s$.
  - Y_0 is the initial value, $t\beta_0$ is a deterministic trend, $U_t := \sum_{s=1}^{t} u_s$ is a stochastic trend.
  - Heuristically, a deterministic trend is some cumulation of determinate quantities, and a stochastic trend is some cumulation of random variables.
    - U_t is a random walk, which is a type of stochastic trend. A random walk is a process formed from the cumulation of serially uncorrelated, stationary, mean-zero random variables.
  - Which of the components of Y_t dominates depends on the relative magnitude of Y_0, β_0, σ_u, where σ_u is the standard deviation of u.
    - Recall that the variation of u is attributable to the "real" randomness of the entity observed, which is supposed to have a range of counterfactual histories.
- A unit root $AR(1)$ model is unrealistic for economic time series because the differences of an $AR(1)$ process are unforecastable, but the differences of economic time series are typically forecastable.
  - By construction of an $AR(1)$ model, by definition of a unit root, $Y_t = \beta_0 + Y_{t-1} + u_t$.
  - Then, the difference of Y_t is $\Delta Y_t := Y_t - Y_{t-1} = \beta_0 + u_t$, where
  - u_t is unforecastable, by construction of an $AR(1)$ model.
- A unit root $AR(p)$ model can similarly be decomposed into (1) the initial value, (2) a deterministic trend, and (3) a stochastic trend.
  - A unit root $AR(p)$ model is some $Y_t = \beta_0 + \sum_{i=1}^{p} \beta_i Y_{t-i} + u_t$, where $\mathbb{E}[u_t | \mathcal{Y}_{t-1}] = 0$ by construction, and where $\sum_{i=1}^{p} \beta_i = 1$.
  - It can be verified that iff $\{Y_t\}$ follows a unit root $AR(p)$ process, then $\{\Delta Y_t\}$ follows an $AR(p-1)$ process.
  - Suppose for simplicity that $\{\Delta Y_t\}$ is stationary. Then $\mathbb{E}\Delta Y_t = \mu$ for all t.
  - Let $v_t := \Delta Y_t - \mu$. Then $\mathbb{E}v_t = \mathbb{E}(\Delta Y_t - \mathbb{E}\Delta Y_t) = \mathbb{E}\Delta Y_t - \mathbb{E}\Delta Y_t = 0$.
  - $$Y_t = Y_0 + \sum_{s=1}^{t} \Delta Y_s$$
    $$= Y_0 + t\mu + \sum_{s=1}^{t} v_s$$
    $$=: Y_0 + t\mu + V_t$$
  - Y_0 is the initial value, $t\mu$ is a deterministic trend, and V_t is a stochastic trend.
- Note that any Y_t with a stochastic trend forms a non-stationary series because the variance of the stochastic trend component is increasing in t, hence the variance of Y_t, in general, is increasing in t and not constant.

Dickey-Fuller Test

- The Dickey-Fuller test is a test for a unit root autoregressive process.
- The parameters of a unit root autoregressive process can be consistently estimated by OLS regression.
  - But the central limit theorem does not apply to ensure the asymptotic normality of the OLS estimators of the parameters of a non-stationary process (and a unit root autoregressive process is non-stationary). Then, the standard methods for performing hypothesis tests and constructing confidence intervals do not apply.

- Additionally, "the OLS estimator suffers from non-trivial biases that become particularly pronounced in the vicinity of a unit root" and "has a markedly non-normal distribution".
- Then, if it is known that some $AR(p)$ process has a unit root, instead of estimating the parameters of that process by OLS, we should estimate the implied $AR(p-1)$ model of the differences, compute forecasts of differences using this estimated model, and then recover the implied forecast of the original series.

Dickey-Fuller (Constant Only) Test

- The Dickey-Fuller (constant only) test is a test of
  - $H_0 : \delta = \delta_0$, where $\delta_0 = 0$, against
  - $H_1 : \delta < \delta_0$, in
  - $\Delta Y_t = \beta_0 + \delta Y_{t-1} + u_t$, where
  - $\delta = \beta_1 - 1$, and
  - $\mathbb{E}[u_t | \mathcal{Y}_{t-1}] = 0$ by construction.
  - Note that $\Delta Y_t = \beta_0 + \delta Y_{t-1} + u_t$ is equivalent to $Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t$, hence $\delta = 0$ is equivalent to $\beta_1 = 1$, i.e. that the $AR(1)$ model $Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t$ has a unit root.
  - The Dickey-Fuller test is a one-sided hypothesis test because in general, for economic time series where a unit root is plausible, the relevant competing model is a stationary model with $\beta_1 < 1$. $\beta_1 > 1$ implies an explosive behaviour that is seldom observed in economics, particularly after taking logarithms.
  - A one-sided test is more powerful against the alternative that the time series is stationary, at the same level of significance.
  - The test statistic used in this test is the standard t statistic $t(\delta_0) = \frac{\hat{\delta} - \delta_0}{se(\hat{\delta})}$.
  - Under the null, $t(\delta_0)$ converges in distribution to the DF_{cm} distribution. The t statistic does not converge in distribution to the standard normal because of the presence of a unit root under the null.
  - A rejection of the null can be interpreted as a finding that $\{Y_t\}$ is a stationary (in the substantive sense of being such that $\beta_1 \in (-1, 1)$) autoregressive process and not a unit root autoregressive process.

Dickey-Fuller (Constant and Trend) Test

- Consider the Dickey-Fuller (constant only) test. Under the null, the process Y_t is a unit root autoregressive process. Under the alternative, the process Y_t is a stationary autoregressive process. Because a time series is not necessarily an autoregressive process, the null hypothesis and alternative hypothesis are not collectively exhaustive.
- For example, consider a trend stationary process, i.e. a process that can be decomposed into a linear trend and a stationary stochastic component.
  - Such a process does not have a unit root because it does not have a stochastic trend.
  - Such a process is better accommodated under the null hypothesis of the Dickey-Fuller (constant only) test because a unit root $AR(1)$ model has a deterministic trend (although it also has a stochastic trend), whereas a stationary $AR(1)$ model exhibits reversion to the mean $\mu = \frac{\beta_0}{1 - \beta_1}$ and has no deterministic trend. Then, it is possible that we fail to reject the false null in such cases because of the deterministic trend.
- The Dickey-Fuller (constant and trend) test is a test of
  - $H_0 : \delta = \delta_0$, where $\delta_0 = 0$, against
  - $H_1 : \delta < \delta_0$, in
  - $\Delta Y_t = \beta_0 + \alpha t + \delta Y_{t-1} + u_t$, where
  - $\delta = \beta_1 - 1$, and
  - $\mathbb{E}[u_t | \mathcal{Y}_{t-1}] = 0$ by construction.
  - The test statistic used in this test is the standard t statistic $t(\delta_0) = \frac{\hat{\delta} - \delta_0}{se(\hat{\delta})}$.
  - Under the null, $t(\delta_0)$ converges in distribution to the DF_{tr} distribution.
  - A rejection of the null can be interpreted as a finding that $\{Y_t\}$ is a (trend) stationary process and does not have a unit root.

Augmented Dickey-Fuller Test

- The augmented Dickey-Fuller test generalises the Dickey-Fuller test to test the null that $\{Y_t\}$ is a unit root $AR(p)$ process against the alternative that it is a (trend) stationary $AR(p)$ process.
- The augmented Dickey-Fuller (constant only) test is a test of

- $H_0 : \delta = \delta_0$, where $\delta_0 = 0$, against
- $H_1 : \delta < \delta_0$, in
- $\Delta Y_t = \beta_0 + \delta Y_{t-1} + \sum_{i=1}^{p_\Delta} \gamma_i \Delta Y_{t-i} + u_t$, which is derived from
- $Y_t = \beta_0 + \sum_{i=1}^{p} \beta_i Y_{t-i} + u_t$, where
  - $\delta = \left(\sum_{i=1}^{p} \beta_i\right) - 1$,
  - $p_\Delta = p - 1$, and
  - $\mathbb{E}[u_t | \mathcal{Y}_{t-1}] = 0$ by construction.
- The test statistic used in this test is the standard t statistic $t(\delta_0) = \frac{\hat{\delta} - \delta_0}{se(\hat{\delta})}$.
- Under the null, $t(\delta_0)$ converges in distribution to the DF_{cn} distribution.
- A rejection of the null can be interpreted as a finding that $\{Y_t\}$ is an $AR(p)$ stationary process and does not have a unit root.

- The augmented Dickey-Fuller (constant and trend) test is a test of
  - $H_0 : \delta = \delta_0$, where $\delta_0 = 0$, against
  - $H_1 : \delta < \delta_0$, in
  - $\Delta Y_t = \beta_0 + \alpha t + \delta Y_{t-1} + \sum_{i=1}^{p_\Delta} \gamma_i \Delta Y_{t-i} + u_t$, which is derived from
    - $Y_t = \beta_0 + \alpha t + \sum_{i=1}^{p} \beta_i Y_{t-i} + u_t$, where
    - $\delta = \left(\sum_{i=1}^{p} \beta_i\right) - 1$,
    - $p_\Delta = p - 1$, and
    - $\mathbb{E}[u_t | \mathcal{Y}_{t-1}] = 0$ by construction.
  - The test statistic used in this test is the standard t statistic $t(\delta_0) = \frac{\hat{\delta} - \delta_0}{se(\hat{\delta})}$.
  - Under the null, $t(\delta_0)$ converges in distribution to the DF_{tr} distribution.
  - A rejection of the null can be interpreted as a finding that $\{Y_t\}$ is a (trend) stationary $AR(p)$ process and not a unit root $AR(p)$ process.
- The augmented Dickey-Fuller test is identical to the Dickey-Fuller test except in augmenting the regression with $\sum_{i=1}^{p_\Delta} \gamma_i \Delta Y_{t-i}$.
- Choice of p_Δ can be guided by information criteria or a stepwise testing down procedure. It is common practice to report the augmented Dickey-Fuller test statistics for several values of p_Δ to verify the robustness of the result of the augmented Dickey-Fuller test to the choice of p_Δ.

Order of Integration

- Recall that the Dickey-Fuller test for a unit root is motivated by the difficulty in estimating a unit root autoregressive model by OLS. Then, if the Dickey-Fuller test rejects the null of a unit root, there is no difficulty in estimating the autoregressive model by OLS. If the Dickey-Fuller test fails to reject the null of a unit root, then the difference of the original $AR(p)$ process is itself a $AR(p-1)$ process. It remains possible that this difference itself has a unit root, so a further Dickey-Fuller test is required to determine if it is stationary.
- The order of integration of a process $\{Y_t\}$ is the smallest $d \in \mathbb{N}_{\geq 0}$ such that $\{\Delta^d Y_t\}$ is stationary. "$Y_t \sim I(d)$" denotes that $\{Y_t\}$ has order of integration d. If $\{Y_t\}$ is stationary, then $Y_t \sim I(0)$. If $\{Y_t\}$ has a unit root and $\{\Delta Y_t\}$ is stationary, then $Y_t \sim I(1)$, and so on.
  - Note the technicality that stationarity is a property of series, not of random variables, and that the order of integration is a property of series, not of random variables, but that the "number of roots" is a property of random variables.
- The procedure for estimating the order of integration of series $\{Y_t\}$ is as follows.
  - Perform the augmented Dickey-Fuller test for a unit root in Y_t.
  - If the null hypothesis that Y_t has a unit root is rejected, then conclude that $\{Y_t\}$ has order of integration 0, i.e. $Y_t \sim I(0)$ and the number of roots of Y_t is 0.
  - If the null hypothesis that Y_t has a unit root is not rejected, then test for a unit root in ΔY_t.
    - Note that differencing eliminates deterministic trends, so the constant-only test is appropriate when testing for a unit root in $\Delta^d Y_t$ for $d \geq 1$.
  - If the null hypothesis that ΔY_t has a unit root is rejected, then conclude that $\{Y_t\}$ has order of integration 0, i.e. $Y_t \sim I(1)$ and the number of roots of Y_t is 1.
  - If the null hypothesis that ΔY_t has a unit root is not rejected, then test for a unit root in $\Delta^2 Y_t$ and so on.
- The order of integration is well defined also for processes that are not autoregressive.

Spurious Regression

- Spurious regression is the tendency to find statistically significant regression relationships between $I(1)$ series, even where they are entirely independent.
- Consider two series $\{Y_t\}, \{X_t\}$. Suppose that the series are independent and are not (descriptively) related to each other, and that $Y_t \sim I(1)$ and $X_t \sim I(1)$. Then,
    - the OLS estimator $\hat{\beta}_1$ of the coefficient on X in a linear regression of Y on X is not consistent for zero, and is a (nondegenerate) random variable, even as T increases indefinitely,
    - the t statistic $t_T(0) = \frac{\hat{\beta}_1 - 0}{se(\hat{\beta}_1)}$ does not converge in distribution to the standard normal or any other distribution, but diverges in magnitude as sample size T increases, such that the probability $\mathbb{P}\{|t_T(0)| > c\}$ that the null $H_0 : \beta_1 = 0$ is rejected converges to 1,
    - R^2 does not converge to 0 and is a (nondegenerate) random variable, that has a high value with non-negligible probability.
- Spurious regression arises because $I(1)$ series have a stochastic trend, hence a high tendency to exhibit long "swings" which entail that, with surprising regularity, movements in one such series can be "matched" against those in another.

Cointegration

- Y_t, X_t are cointegrated iff $Y_t, X_t \sim I(1)$ and there exists $\theta \in \mathbb{R}$ such that $\{\xi_t\} := \{Y_t - \theta X_t\}$ has order of integration 0, i.e. iff some linear combination of the two series is stationary.
    - If Y_t, X_t are cointegrated, then the $\theta \in \mathbb{R}$ such that $\{\xi_t\} := \{Y_t - \theta X_t\}$ has order of integration 0 is unique (i.e. for all $\theta' \neq \theta$, $\{\xi'_t\} := \{Y_t - \theta' X_t\}$ is not stationary), and is referred to as the cointegrating coefficient.
    - The series $\{\xi_t\}$ is referred to as the equilibrium error.
    - In general, $\{\xi_t\}$ exhibits serial correlation.
    - Note the technicality that cointegration is a property of random variables and not a property of series.
- This definition extends straightforwardly to three or more series. n series $Y_t^1, \ldots, Y_t^n$ are cointegrated iff $Y_t^1, \ldots, Y_t^n \sim I(1)$ and there exists $\theta_2, \ldots, \theta_n \in \mathbb{R}$ such that $\{\xi_t\} := \{Y_t^1 - \theta_2 Y_t^2 - \ldots - \theta_n Y_t^n\}$ has order of integration 0.
- **If Y_t, X_t are cointegrated, OLS regression of Y_t on X_t consistently estimates the cointegrating coefficient θ. Then, the estimated equilibrium error $\hat{\xi}_t := Y_t - \hat{\theta} X_t$ is consistent for the equilibrium error ξ_t, which has order of integration 0, and is stationary.**
    - Suppose that Y_t, X_t are cointegrated. Then, there exists $a, b \in \mathbb{R}$ such that $Y_t - a - bX_t$ contains no stochastic and/or deterministic trend. Noting that the variance of a stochastic trend component is large and increasing in t, and that the magnitude of a deterministic trend is large and non-zero, informally, minimising the sum of squared errors $\sum_{t=1}^{T}(Y_t - a - bX_t)^2$ "requires" eliminating the stochastic and/or deterministic trend components. So the OLS estimator, which minimises the sum of squared errors, selects a, b (specifically b) to eliminate the stochastic and/or deterministic trend components, which is when $b = \theta$.

Cointegration Test

- The principal motivation for a test of cointegration is to determine if a regression involving $I(1)$ variables is potentially spurious, or if the regression estimates a cointegrating relationship. Heuristically, such a test can be used to distinguish between spurious and genuine findings.
- The procedure for testing for cointegration (that is suggested by the above discussion) is as follows.
    - Perform an augmented Dickey-Fuller test on each of $\{Y_t\}, \{X_t\}$ to determine the order of integration of each series.
    - If $Y_t \sim I(0)$ or $X_t \sim I(0)$, then there is no possibility of cointegration, but also no possibility of spurious regression.
    - If $Y_t, X_t \sim I(1)$, then continue to test for spurious regression.
    - If there is a hypothesised value of θ, compute $\xi_t := Y_t - \theta X_t$, and perform an augmented Dickey-Fuller test (for a suitable range of p_Δ) to determine the order of integration of $\{\xi_t\}$.
        - If the null is rejected, then conclude that $\xi_t \sim I(0)$, then Y_t, X_t are cointegrated with cointegrating coefficient θ.
        - If the null is not rejected, then Y_t, X_t are still potentially cointegrated, but not with the cointegrating coefficient θ.
    - If there is no hypothesised value of θ, estimate θ by OLS regression of Y_t on X_t, then compute the estimated equilibrium error $\hat{\xi}_t = Y_t - \hat{\theta} X_t$. Then, perform an augmented Dickey-Fuller test (for a suitable range of p_Δ) to determine the order of integration of $\hat{\xi}_t$, using adjusted (Engle-Granger) critical values.
        - If the null is rejected, then conclude that Y_t, X_t are cointegrated.
        - If the null is not rejected, then conclude that Y_t, X_t are not cointegrated.