Università Commerciale
Luigi Bocconi

MSc. Finance/CLEFIN
2017/2018 Edition

FINANCIAL ECONOMETRICS AND EMPIRICAL FINANCE - MODULE 2

Mock Question 3 (total 17 points, out of 50 from 3 questions)

Time Advised: 24 minutes (for this question)

Question 3.A (13 points)

Define a stochastic trend and indicate what is the relationship between a stochastic trend and a random walk, with and without drift, for the special case of a I(1) process. For this case, comment on (or show, as you deem most appropriate) the stationarity or lack thereof of a random walk and explain why this may represent a problem in empirical work. Indicate how would you proceed to make a I(d) time series, $\{y_t\}$, with $d \geq 2$, stationary. Would the choice of considering $\{y_t - y_{t-d}\}$ instead of $\{y_t\}$ be an appropriate one? Make sure to carefully explain your answers.

Debriefing:

Trends in Time Series

- The long-term forecast of y_t will converge to the trend line, δt, so that this type of model is said to be trend stationary

② Stochastic trends, which characterize all processes that can be written as: $y_t = y_0 + \mu t + \sum_{\tau=1}^{t} \varepsilon_\tau$

  - Because $y_{t+1} = y_0 + \mu t + \sum_{\tau=1}^{t} \varepsilon_\tau + \mu + \varepsilon_{t+1} = \mu + y_t + \varepsilon_{t+1}$, presence of a stochastic trend, implies a random walk with drift: $y_{t+1} = \mu + y_t + \epsilon_{t+1}$
- Therefore a stochastic trend is not a RW, but it implies its presence
  - A RW is the non-stationary variant of AR(1) with $\mu = \phi_0$ and $\phi_1 = 1$
- A determinist trend also implies the presence of a stochastic trend, but stochastic trends may arise on their own
- Since all values of ϵ_τ carry a coefficient of unity, the effect of each shock on the intercept term is permanent, which is indeed the intrinsic nature of a trend
- If shocks are never forgotten and time series have infinite memory $\Longrightarrow$ both deterministic and stochastic trends are non-stationary, denoted as I(d), $d \geq 1$

Lecture 5: Unit Roots, Cointegration and Error Correction Models – Prof. Guidolin 4

The Random Walk Process

- The RW is the key (but not only) type of non-stationary process
  - While its conditional mean is well-defined, its unconditional mean explodes: $E_t[y_{t+1}] = \mu + E_t[y_t] + E_t[\epsilon_{t+1}] = \mu + y_t$

$$E_t[y_{t+s}] = \mu + E_t[y_{t+s-1}] + E_t[\epsilon_{t+s}] = 2\mu + E_t[y_{t+s-2}] + E_t[\epsilon_{t+s} + \epsilon_{t+s-1}]$$
$$= \ldots = s\mu + y_t \qquad (s > t)$$
$$E[y_t] = E_t[y_0 + \mu t + \textstyle\sum_{\tau=1}^{t} \varepsilon_\tau] = y_0 + \mu t$$

  so it depends on t and if there is no drift ($\mu = 0$), then $E_t[y_{t+1}] = E_t[y_{t+s}]$
  - Also unconditional variance depends on time and it explodes as $t \longrightarrow \infty$

$$Var[y_t] = Var\left[y_0 + \sum_{\tau=1}^{t} \varepsilon_\tau\right] = \sum_{\tau=1}^{t} Var[\varepsilon_\tau] = t\sigma^2$$

  - Also autocovariances and autocorrelations display pathological patterns:

$$E[(y_t - y_0)(y_{t+s} - y_0)] = E\left[\sum_{\tau=1}^{t} \varepsilon_\tau, \sum_{\tau=1}^{t+s} \varepsilon_\tau\right] = E\left[\sum_{\tau=1}^{t} \varepsilon_\tau^2\right] = t\sigma^2$$
$$Corr[y_t, y_{t+s}] = \frac{E[(y_t - y_0)(y_{t+s} - y_0)]}{\sqrt{Var[y_t]Var[y_{t+s}]}} = \frac{t\sigma^2}{\sqrt{t\sigma^2(t+s)\sigma^2}} = \sqrt{\frac{t}{t+s}}$$

As t $\longrightarrow \infty$, there is perfect memory, $Corr[y_t, y_{t+s}] \simeq 1$

Lecture 5: Unit Roots, Cointegration and Error Correction Models – Prof. Guidolin 6

Stochastic trend: all series that can be written as

$$y_t = y_0 + \mu t + \sum_{\tau=1}^{t} \epsilon_\tau$$

- $Var[\epsilon_t] = \sigma^2$ and $y_{t+1} = y_0 + \mu t + \sum_{\tau=1}^{t} \epsilon_\tau + \mu + \epsilon_{t+1} = \mu + y_t + \epsilon_{t+1}$
  $\Rightarrow$ The series follows a random walk with drift process.

or more generally, y_t contains a stochastic trend $\iff$ it can be decomposed as

$$y_t = y_0 + \mu t + \sum_{\tau=1}^{t} \eta_\tau$$

1

where η_τ follows any ARMA(p, q) stationary process.
This expression is centered on the first-differences of y_t.

- All deterministic trends can be converted into stochastic trends, while the opposite is not true.
- A series $\{y_t\}$ can be decomposed as y_t=trend + stationary component+noise=(deterministic trend+stochastic trend)+stationary component+noise.
- In large samples, the deterministic time trend induced by the drift component dominates the time series, while in small samples it is not always easy to discern the difference between a driftless RW and a model with drift.

Expectation of a Random Walk:

1. If $\mu = 0$

$$E[y_t] = E[y_{t+s}] = y_0 + E\Big[\sum_\tau\Big]\epsilon_\tau = y_0 + \sum_\tau E[\epsilon_\tau] = y_0$$

that is constant equal to the initial value. However, all stochastic shocks have non-decaying effects on the series.

If $\mu \neq 0$

$$E[y_t] = E_t[y_0 + \mu t + \sum_{\tau=1}^{t}\epsilon_\tau] = y_0 + \mu t$$

2. If $\mu = 0$

$$E_t[y_{t+s}] = E_t[y_{t+s-1} + \epsilon_{t+s}] = E_t[y_{t+s-2} + \epsilon_{t+s} + \epsilon_{t+s-1}] = \ldots = E_t[y_t] + E_t[\sum_{i=1}^{s}\epsilon_{t+i}] = y_t$$

that is for any $s \geq 2$ the conditional means for all values of y_{t+s} are equivalent and the current y_t represents the minimum mean-squared loss function forecast of all future values of the series.

If $\mu \neq 0$

$$E_t[y_{t+s}] = E_t[y_0 + \mu(t+s) + \sum_{\tau=1}^{t}\epsilon_t] = y_0 + \mu s$$

that is the forecast function changes deterministically with time.

Expectation of a driftless Random Walk ($\mu = 0$):

(a)

$$Var[y_t] = Var[y_0 + \sum_{\tau=1}^{t}\epsilon_\tau] = \sum_{\tau=1}^{t} Var[\epsilon_\tau] = t\sigma^2$$

the variance is time-dependent and therefore a RW is not covariance stationary

(b)

$$Var[y_{t+s}] = Var[y_0 + \sum_{\tau=1}^{t+s}\epsilon_\tau] = \sum_{\tau=1}^{t+s} Var[\epsilon_\tau] = (t+s)\sigma^2$$

as $s \to \infty$ the variance approaches infinity.

(c) Given that the mean is constant

$$E[(y_t - y_0)(y_{t+s} - y_0)] = E[\sum_{\tau=1}^{t}\epsilon_\tau \cdot \sum_{\tau=1}^{t+s}\epsilon_\tau] = E[\sum_{\tau=1}^{t}\epsilon_\tau^2] = t\sigma^2$$

De-Trending a Series: Deterministic vs. Stochastic

Definition **(Deterministic De-Trending)** De-trending entails regressing a variable on a deterministic (polynomial) function of time and saving the residuals, $\{\hat{\varepsilon}_t\}$, that come then to represent the new, de-trended series,

$$\hat{y}_{t+1} = \sum_{j=0}^{Q}\hat{\delta}_j t^j + \hat{\varepsilon}_{t+1},$$

where the coefficients can be simply estimated by OLS.

- In trend-stationary case, de-trending is simply done by OLS
- For stochastic trends, consider the RW with drift process and take its first difference: $\Delta y_{t+1} \equiv y_{t+1} - y_t = (\mu + y_t + \varepsilon_{t+1}) - y_t = \mu + \varepsilon_{t+1}$
- The result is a white noise plus a constant intercept (the drift)
- This approach is more general:

Definition **(Unit Root Process and *d*th Order Integration)** When a time series process $\{y_t\}$ needs to be differenced d times before being reduced to the sum of constant terms plus a white noise process, $\{y_t\}$ is said to contain d unit roots or to be **integrated of order *d***; we also write that $y_t \sim I(d)$.

(d)

$$Corr[y_t, y_{t+s}] = \frac{E[(y_t - y_0)(y_{t+s} - y_0)]}{\sqrt{Var[y_t]Var[y_{t+s}]}} = \frac{t\sigma^2}{\sqrt{t\sigma^2(t+s)\sigma^2}} = \sqrt{\frac{t}{t+s}}$$

- A driftless RW meanders without exhibiting any tendency to increase or decrease. In fact, for any fixed real number M, the random walk and its absolute will exceed M with probability 1 as the series lengthens.
- For sufficiently large samples, $Corr[y_t, y_{t+s}] \simeq 1$, but as s grows $Corr[y_t, y_{t+s}]$ declines below 1 $\Rightarrow$ We cannot distinguish between a unit root process and a stationary process with an autoregressive coefficient that is close to unity using the ACF.

- The de-trended process can then be modeled using traditional methods

Unit Root Process and dth Order Integration: When a time series process $\{y_t\}$ needs to be differenced d times before being reduced to the sum of constant terms plus a white noise process, $\{y_t\}$ is said to contain d unit roots or to be integrated of order d; we also write that $y_t \sim I(d)$.

Example: RW with drift process.
Take its first-difference:

$$\Delta y_{t+1} \equiv y_{t+1} - y_t = (\mu + y_t + \epsilon_{t+1}) - y_t = \mu + \epsilon_{t+1}$$

$\Rightarrow$ white noise series plus a constant intercept. It is a $I(1)$ and it contains one unit root.

- If $\{y_t\}$ contains d unit roots, then $\{y_t\}$ is non-stationary, while the opposite does not hold (e.g. explosive autoregressive process: $y_{t+1} = \mu + \phi y_t + \epsilon_{t+1}$).
- Typically $\phi = 1$ is used to characterize non-stationarity because it describes accurately many time series.
- All $I(d)$ processes (with $d > 0$) are non-stationary, but there are non-stationary processes that do not fit the definition of unit root processes.

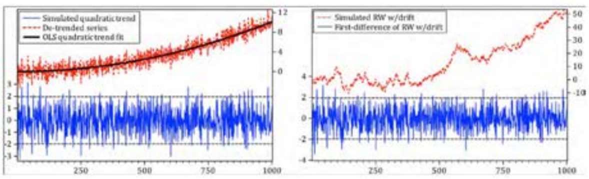


Figure 4: De-trended and First-Differenced Series

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob	Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.030	0.030	0.9193	0.338			1	0.032	0.032	1.0573	0.304
		2	-0.045	-0.046	2.9546	0.228			2	-0.043	-0.044	2.9340	0.231
		3	0.023	0.026	3.5007	0.321			3	0.025	0.028	3.5634	0.313
		4	0.019	0.016	3.8700	0.424			4	0.020	0.017	3.9801	0.409
		5	-0.021	-0.020	4.3076	0.506			5	-0.020	-0.019	4.3632	0.498
		6	-0.014	-0.012	4.5094	0.608			6	-0.013	-0.011	4.5462	0.603
		7	-0.049	-0.051	6.9761	0.431			7	-0.048	-0.050	6.8865	0.441
		8	-0.074	-0.072	12.495	0.130			8	-0.073	-0.071	12.272	0.139
		9	-0.041	-0.040	14.185	0.116			9	-0.040	-0.039	13.892	0.126
		10	0.024	0.023	14.791	0.140			10	0.026	0.024	14.554	0.149
		11	-0.008	-0.009	14.860	0.189			11	-0.007	-0.007	14.603	0.201
		12	-0.032	-0.027	15.899	0.196			12	-0.031	-0.025	15.548	0.213

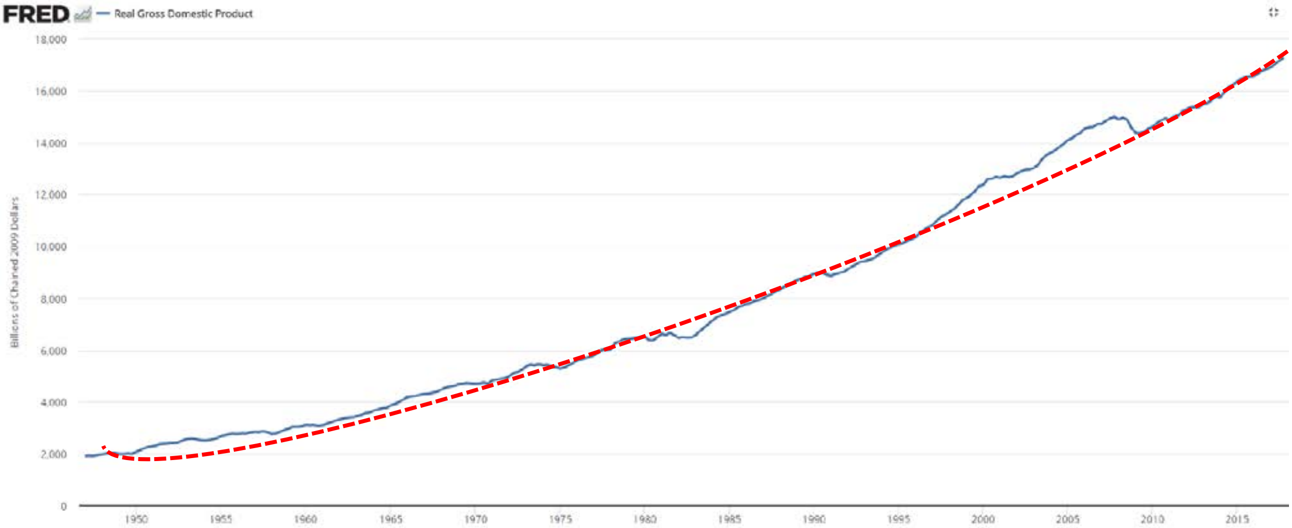
Figure 5: Sample ACF of De-trended and First-Differenced Series

The answer to the last sub-point is negative because we know that $\Delta^d y_t \neq y_t - y_{t-d}$, while $\Delta^d y_t$ consists of taking a number of d of successive differences of the series under consideration, i.e.,

$$\Delta^2 y_t = \Delta(\Delta y_t) \qquad \Delta^3 y_t = \Delta(\Delta^2 y_t) \quad \ldots$$

Question 3.B (2.5 points)

An analyst at Charles Thomas and Associates has just downloaded the following series of data on the quarterly US real GDP (in constant dollars, expressed as 2009 billions).



He has proposed to make this series stationary by first fitting (by simple OLS) a quadratic function of time (shown as a dashed red line in the picture) and then replace the time series of real GDP with the OLS residuals from such a quadratic trend regression. What are the risks that the analysts is exposing himself and his firm to by adopting this simple procedure?

Debriefing:

As commented in the lectures, such a procedure does not really make an obviously trending, non-stationary series any stationary when the series contains a stochastic unit root, which should be tested as a first order of matter. The risks are that, because the adopted method is ineffective, the residuals will be then treated as I(0) while they are in fact I(1) or worse.

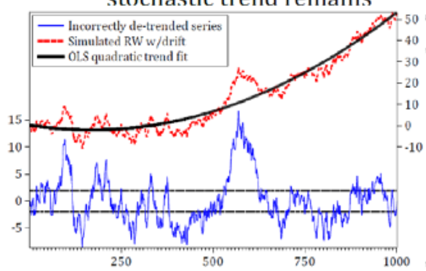
Pitfalls in De-Trending Applications

- Serious damage—in a statistical sense—can be done when the inappropriate method is used to eliminate a trend

  ① When a time series is I(d) but an attempt is made to remove its stochastic trend by fitting deterministic time trend functions, the OLS residuals will still contain one or more unit roots

  - Deterministic de-trending does not remove the stochastic trends
  - For instance: $y_t - \hat{y}_t^{de-trend} = y_0 + \mu t + \sum_{\tau=1}^{t} \varepsilon_\tau - \hat{\delta}_0 - \hat{\delta}_1 t$
  - Even when $\mu = \delta$, the stochastic trend remains $= (y_0 - \hat{\delta}_0) + (\mu - \hat{\delta}_1)t + \sum_{\tau=1}^{t} \varepsilon_\tau$

Linear trend



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.974	0.974	952.57	0.000
		2	0.946	-0.044	1852.9	0.000
		3	0.921	0.033	2706.8	0.000
		4	0.895	-0.040	3513.1	0.000
		5	0.867	-0.030	4271.6	0.000
		6	0.841	0.006	4985.6	0.000
		7	0.816	-0.001	5657.5	0.000
		8	0.793	0.039	6292.9	0.000
		9	0.774	0.061	6898.8	0.000
		10	0.757	0.031	7479.3	0.000
		11	0.739	-0.033	8032.9	0.000
		12	0.721	-0.001	8561.1	0.000

Question 3.C (1.5 points)

You know that a time series $\{y_t\}$ was originally suspected to be I(d) with $d \geq 1$. A fellow quant analyst, Ms. Maria Delas, has then transformed it by differentiating three times, in the attempt to make it stationary and delivered the series to you. Upon your own analysis, you determine that the series contains now 2 unit roots in its MA component (i.e., the residuals need to be differentiated twice for them to be "well-behaved", that we may have called invertible). What do you know about the d characterizing the original series?

Debriefing:

The original series was I(1): differentiating it three times—well more than what is needed—"messes" it stochastic structure up, by creating two unit roots in its MA component. In short, if d – 3 = -2, then it must have been d = 1.

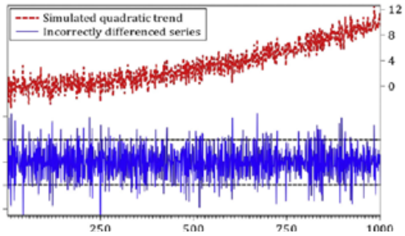
Pitfalls in De-Trending Applications

② When a time series contains a deterministic trend but is otherwise I(0), (trend-stationary) and an attempt is made to remove the trend by differentiating the series d times, the resulting differentiated series will contain d unit roots in its MA components

- It will therefore be not *invertible*
- Differentiating a trend- stationary series, creates new stochastic trends that are shifted inside the shocks of the series

$$\Delta y_t \equiv y_t - y_{t-1} = \left(\sum_{j=0}^{Q} \delta_j t^j + \varepsilon_{t+1}\right) - \left(\sum_{j=0}^{Q} \delta_j (t-1)^j + \varepsilon_t\right)$$

$$= \sum_{j=0}^{Q} \delta_j [t^j - (t-1)^j] + (\varepsilon_{t+1} - \varepsilon_t)$$



Autocorrelation	Partial correlation		AC	PAC	Q-Stat	Prob
		1	-0.460	-0.460	212.08	0.000
		2	-0.075	-0.363	217.71	0.000
		3	0.036	-0.259	219.02	0.000
		4	0.020	-0.176	219.44	0.000
		5	-0.025	-0.156	220.05	0.000
		6	0.021	-0.101	220.50	0.000
		7	-0.005	-0.076	220.53	0.000
		8	-0.030	-0.102	221.44	0.000
		9	-0.019	-0.152	221.80	0.000
		10	0.053	-0.102	224.65	0.000
		11	-0.005	-0.076	224.68	0.000
		12	-0.042	-0.116	226.51	0.000

Pitfalls in De-Trending Applications

- Even when the trend-stationary component is absent, if the time series is I(0) but it is incorrectly differenced d times, the resulting differentiated series will contain d unit roots in its MA components

- What if $y_t \sim I(d)$ but by mistake we differentiate it $d + r$ times?

③a -- If $r > 0$, we are over-differencing the series, and as such ② applies, that is, the resulting over-differentiated series will contain r unit roots in its MA components and will therefore be not invertible

③b -- If $r < 0$, we are not differencing the series enough and the resulting series will still contain $d - r$ and will remain nonstationary