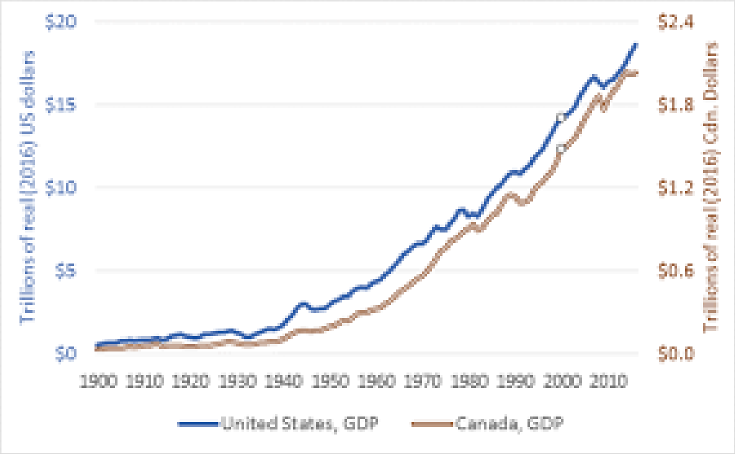
Linear Time Trend Model

John C. Chao

Econ 721 Lecture Notes

September 1, 2020

- Many economic time series, especially macroeconomic time series, are very persistent; that is, they exhibit certain trend-like behavior.



- Some Important Questions from an Econometric Perspective
  1. How do we model this trending behavior?
  2. How does our model of trends affect the statistical properties of our estimators?
- Two Alternative Models of Trending Behavior:

$$\begin{aligned} y_t &= \alpha + \beta t + u_t \qquad \text{(TS)}, \\ y_t &= \alpha + y_{t-1} + u_t \qquad \text{(DS)}, \end{aligned}$$

where $\{u_t\}$ may be taken to be strictly stationary, ergodic process.
- **Remark:** These two alternative models of trending behavior are not as dissimliar as they might appear at first sight.

- Note that we can rewrite the difference stationary (DS) model as follows.

$$\Delta y_s = y_s - y_{s-1} = \alpha + u_s$$

Summing both sides of the equation above from $t = 1$ to t, we obtain

$$\begin{aligned} y_t - y_0 &= \sum_{s=1}^{t} \Delta y_s \\ &= \sum_{s=1}^{t} \alpha + \sum_{s=1}^{t} u_s \\ &= \alpha t + \sum_{s=1}^{t} u_s \end{aligned}$$

- Adding y_0 to both sides of the equation, we further obtain

$$y_t = \underbrace{\alpha t}_{\text{linear trend}} + \underbrace{\sum_{s=1}^{t} u_s}_{\text{stochastic trend}} + y_0$$

Hence, we see that a DS model with an intercept is actually a model that has both a deterministic linear trend and a stochastic trend.

Time Series Model with Deterministic Trend

- Consider the simple linear trend model

$$\begin{aligned} y_t &= \alpha + \beta t + u_t \\ &= x_t' \theta + u_t \end{aligned}$$

where

$$\theta = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \text{ and } x_t = \begin{pmatrix} 1 \\ t \end{pmatrix}$$

- **Assumptions:**

(i) $\{u_t\} \equiv i.i.d. \left(0, \sigma^2\right)$, $0 < \sigma^2 < \infty$.

(ii) $E\left[u_t^4\right] < \infty$

Time Series Model with Deterministic Trend

- **Remark:** The linear trend model looks like a simple, classical linear regression. However, suppose we try to prove consistency of the OLS estimator $\widehat{\theta}_T$ of the parameter vector θ, we will see that we need to modify the arguments we gave previously for classical linear models a bit.
- To see this, write

$$\begin{aligned} \widehat{\theta}_T &= \left(\sum_{t=1}^{T} x_t x_t' \right)^{-1} \sum_{t=1}^{T} x_t y_t \\ &= \left(\sum_{t=1}^{T} x_t x_t' \right)^{-1} \sum_{t=1}^{T} x_t \left(x_t' \theta + u_t \right) \\ &= \theta + \left(\sum_{t=1}^{T} x_t x_t' \right)^{-1} \sum_{t=1}^{T} x_t u_t \end{aligned}$$

Time Series Model with Deterministic Trend

- so, following the standard argument for classical linear regression, we would write

$$
\begin{aligned}
\widehat{\theta}_T - \theta &= \begin{pmatrix} \widehat{\alpha}_T - \alpha \\ \widehat{\beta}_T - \beta \end{pmatrix} \\
&= \left(\frac{1}{T} \sum_{t=1}^{T} x_t x_t' \right)^{-1} \frac{1}{T} \sum_{t=1}^{T} x_t u_t \\
&= \left(\frac{1}{T} \sum_{t=1}^{T} \begin{bmatrix} 1 \\ t \end{bmatrix} \begin{bmatrix} 1 & t \end{bmatrix} \right)^{-1} \frac{1}{T} \sum_{t=1}^{T} \begin{bmatrix} 1 \\ t \end{bmatrix} u_t \\
&= \begin{pmatrix} T^{-1} \sum_{t=1}^{T} 1 & T^{-1} \sum_{t=1}^{T} t \\ T^{-1} \sum_{t=1}^{T} t & T^{-1} \sum_{t=1}^{T} t^2 \end{pmatrix}^{-1} \begin{pmatrix} T^{-1} \sum_{t=1}^{T} u_t \\ T^{-1} \sum_{t=1}^{T} t u_t \end{pmatrix}
\end{aligned}
$$

Time Series Model with Deterministic Trend

- However, this is the wrong way to standardize the "numerator" and the "denominator" of the OLS estimator in this case since, amongst other things, we know, by an elementary summation formula that

$$\frac{1}{T}\sum_{t=1}^{T} t = \frac{1}{T}\frac{T\,(T+1)}{2} \to \infty \text{ as } T \to \infty.$$

- **Remark:** Hence, one challenge with analyzing models with trends is that sample averages associated with different components of the model may have different order of magnitude. As we will see, this will, in turn, lead to estimators of different parameters possibly having different rates of convergence.
- To analyze the asymptotic properties of $\widehat{\theta}_T = \left(\widehat{\alpha}_T, \widehat{\beta}_T\right)'$, the OLS estimator of θ, we, thus, introduce the diagonal matrix

$$D_T = \begin{pmatrix} \sqrt{T} & 0 \\ 0 & T^{3/2} \end{pmatrix}.$$

Time Series Model with Deterministic Trend

- Next, write

$$D_T\left(\widehat{\theta}_T-\theta\right)=\left[D_T^{-1}\sum_{t=1}^{T}x_tx_t'D_T^{-1}\right]^{-1}D_T^{-1}\sum_{t=1}^{T}x_tu_t$$

- Now,

$$\begin{aligned}&D_T^{-1}\sum_{t=1}^{T}x_tx_t'D_T^{-1}\\=&\begin{pmatrix}T^{-1/2}&0\\0&T^{-3/2}\end{pmatrix}\begin{pmatrix}T&\sum_{t=1}^{T}t\\\sum_{t=1}^{T}t&\sum_{t=1}^{T}t^2\end{pmatrix}\begin{pmatrix}T^{-1/2}&0\\0&T^{-3/2}\end{pmatrix}\\=&\begin{pmatrix}1&T^{-2}\sum_{t=1}^{T}t\\T^{-2}\sum_{t=1}^{T}t&T^{-3}\sum_{t=1}^{T}t^2\end{pmatrix}\end{aligned}$$

- Hence, for the off-diagonal terms, we have, as $T \to \infty$,

$$\frac{1}{T^2}\sum_{t=1}^{T} t = \frac{1}{T^2}\frac{T\left(T+1\right)}{2} = \frac{1}{2} + \frac{1}{2T} \to \frac{1}{2}$$

- More generally, we have, for positive integer v,

$$\frac{1}{T^{v+1}}\sum_{t=1}^{T} t^v = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{t}{T}\right)^v \to \int_0^1 r^v dr = \left.\frac{r^{v+1}}{v+1}\right|_0^1 = \frac{1}{v+1}$$

- Applying this to the case where $v = 2$, we have

$$\frac{1}{T^3}\sum_{t=1}^{T} t^2 \to \frac{1}{3} \text{ as } T \to \infty$$

Time Series Model with Deterministic Trend

- Putting everything together, we then obtain

$$
\begin{aligned}
& D_T^{-1} \sum_{t=1}^{T} x_t x_t' D_T^{-1} \\
= & \begin{pmatrix} 1 & T^{-2} \sum_{t=1}^{T} t \\ T^{-2} \sum_{t=1}^{T} t & T^{-3} \sum_{t=1}^{T} t^2 \end{pmatrix} \\
\to & \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix} = Q \quad (say)
\end{aligned}
$$

- Next, we turn our attention to the "numerator"

$$
\begin{aligned}
D_T^{-1} \sum_{t=1}^{T} x_t u_t &= \begin{pmatrix} T^{-1/2} & 0 \\ 0 & T^{-3/2} \end{pmatrix} \begin{pmatrix} \sum_{t=1}^{T} u_t \\ \sum_{t=1}^{T} t u_t \end{pmatrix} \\
&= \begin{pmatrix} T^{-1/2} \sum_{t=1}^{T} u_t \\ T^{-3/2} \sum_{t=1}^{T} t u_t \end{pmatrix}
\end{aligned}
$$

Time Series Model with Deterministic Trend

- To show that

$$D_T^{-1} \sum_{t=1}^{T} x_t u_t$$

converges to an asymptotic normal distribution, we make use of the Cramér-Wold device and a central limit theorem for martingale difference array.

- **Cramér-Wold Device:** Let $Z_1, Z_2, ...$ be a sequence of $k \times 1$ random vectors. Then, the following statements are equivalent

(i) $Z_n \xrightarrow{d} Z$ as $n \to \infty$

(ii) $\lambda' Z_n \xrightarrow{d} \lambda' Z$ as $n \to \infty$ for all $\lambda \in \mathbb{R}^k$ such that $\lambda \neq 0$

(iii) $\lambda' Z_n \xrightarrow{d} \lambda' Z$ as $n \to \infty$ for all $\lambda \in \mathbb{R}^k$ such that $\|\lambda\| = 1$.

Time Series Model with Deterministic Trend

- To proceed, let $\lambda = (\lambda_1, \lambda_2)'$ be a (nonrandom) 2×1 vector such that $\|\lambda\| = 1$ and construct the linear combination

$$\lambda' D_T^{-1} \sum_{t=1}^{T} x_t u_t = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left[\lambda_1 + \lambda_2 \left(\frac{t}{T} \right) \right] u_t = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} W_{t,T}$$

where

$$W_{t,T} = \left[\lambda_1 + \lambda_2 \left(\frac{t}{T} \right) \right] u_t$$

- **Remark:** Note that the elements of the sum above depend on both t and T. Hence, to obtain a large sample approximation for the standardized sum above, we need to talk about a central limit theorem for triangular array.

Review of Martingale Difference Array and MDACLT

- Let $\{W_{tT} : t = 0, ..., T, T \geq 1\}$ be a triangular array of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ and let $\{\mathcal{F}_{tT} : t = 0, ..., T, T \geq 1\}$ be a triangular array of nested σ-fields such that

$$\mathcal{F}_{0T} \subseteq \mathcal{F}_{1T} \subseteq \mathcal{F}_{2T} \subseteq \cdots \subseteq \mathcal{F}_{TT} \subseteq \mathcal{F}.$$

  This is known as a filtration. If W_{tT} is $\mathcal{F}_{tT}$-measurable for every t and T; then, the triangular array $\{W_{tT}\}$ is said to be adapted to $\{\mathcal{F}_{tT}\}$ and the pair $\{W_{tT}, \mathcal{F}_{tT}\}$ is called an adapted triangular array.

- **Definition:** An adapted triangular array $\{W_{tT}, \mathcal{F}_{tT}\}$ is called a martingale difference array (or m.d.a. for short) if
  (a) There exists positive constant C such that $E\,|W_{tT}| \leq C < \infty$ for all t and T.
  (b) $E\left[W_{tT} | \mathcal{F}_{t-1,T}\right] = 0$ $a.s.$for all t and T.

Remarks of Martingale Difference Array and MDACLT

- Note that conditions (a) and (b) and the law of iterated expectations imply that
$$E\left[W_{tT}\right] = 0 \text{ for all } t \text{ and } T.$$
- Along with an appropriate moment condition, conditions (a) and (b) also imply that
$$Cov\left(W_{tT}, \phi\left(W_{t-1,T}, W_{t-2,T}, ..., W_{0,T}\right)\right) = 0$$
for any Borel-measurable, integrable function of the arguments.
- It is tempting to view the martingale difference property as being intermediate between uncorrelatednesss and independence. However, the definition of m.d.a. allows for a possible asymmetry with respect to time. While reversing the direction of time will not affect the concept of uncorrelatedness or independence, reversing the time ordering on a m.d.a. (i.e, treating the future as the past and vice versa) will not in general result in a m.d.a.

- The importance of m.d.a.to modern probability theory has to do with the fact that it allows us to introduce some dependence and yet still be able to, under rather general conditions, prove central limit results, very similar to the “classical" ones.

Review of Martingale Difference Array and MDACLT

- **Theorem:** Let $\{W_{tT}, \mathcal{F}_{tT}\}$ be a martingale difference array with finite unconditional variances $\{\sigma^2_{tT}\}$, where $\sigma^2_{tT} = E\left[W^2_{tT}\right]$. Let

$$\omega^2_T = \sum_{t=1}^{T} \sigma^2_{tT}$$

Suppose that, as $T \to \infty$, the following conditions hold

(a)

$$\frac{1}{\omega^2_T} \sum_{t=1}^{T} W^2_{tT} \xrightarrow{p} 1;$$

(b) for some $\delta > 0$

$$\sum_{t=1}^{T} E\left[\left|\frac{W_{tT}}{\omega_T}\right|^{2+\delta}\right] \to 0$$

Then,

$$S_T = \frac{1}{\omega_T} \sum_{t=1}^{T} W_{tT} \xrightarrow{d} N(0,1) \text{ as } T \to \infty.$$

- Note that condition (b) above is sometimes known as Liapunov's condition. The condition is stronger than the well-known Lindeberg's condition, i.e., for any $\varepsilon > 0$

$$\sum_{t=1}^{T} E\left[\frac{W_{tT}^2}{\omega_T^2}\mathbb{I}\left\{\left|\frac{W_{tT}}{\omega_T}\right| > \varepsilon\right\}\right] \to 0 \text{ as } T \to \infty$$

This can be seen from the fact that for every t and T

$$\begin{aligned} & \sum_{t=1}^{T} E\left[\frac{W_{tT}^2}{\omega_T^2}\mathbb{I}\left\{\left|\frac{W_{tT}}{\omega_T}\right| > \varepsilon\right\}\right] \\ \leq & \sum_{t=1}^{T} E\left[\left|\frac{W_{tT}}{\omega_T}\right|^2 \left|\frac{W_{tT}/\omega_T}{\varepsilon}\right|^{\delta}\mathbb{I}\left\{\left|\frac{W_{tT}}{\omega_T}\right| > \varepsilon\right\}\right] \\ = & \frac{1}{\varepsilon^{\delta}}\sum_{t=1}^{T} E\left[\left|\frac{W_{tT}}{\omega_T}\right|^{2+\delta}\mathbb{I}\left\{\left|\frac{W_{tT}}{\omega_T}\right| > \varepsilon\right\}\right] \end{aligned}$$

- so that

$$
\begin{aligned}
& \sum_{t=1}^{T} E\left[\frac{W_{tT}^2}{\omega_T^2}\mathbb{I}\left\{\left|\frac{W_{tT}}{\omega_T}\right| > \varepsilon\right\}\right] \\
\leq \quad & \frac{1}{\varepsilon^{\delta}}\sum_{t=1}^{T} E\left[\left|\frac{W_{tT}}{\omega_T}\right|^{2+\delta}\mathbb{I}\left\{\left|\frac{W_{tT}}{\omega_T}\right| > \varepsilon\right\}\right] \\
\leq \quad & \frac{1}{\varepsilon^{\delta}}\sum_{t=1}^{T} E\left[\left|\frac{W_{tT}}{\omega_T}\right|^{2+\delta}\right]
\end{aligned}
$$

- The situations where the Liapunov condition and the Lindeberg condition are designed to rule out are ones where the behavior of a finite subset of the sequence (or array) elements dominates the behavior of all others, even in the limit. Suppose, for example, a substantial number of σ_{tT}^2 is zero or tends toward zero such that

$$\omega_T^2 = \sum_{t=1}^{T} \sigma_{tT}^2$$

is bounded in T. In such cases, the specified condition would not hold.

Asymptotic Normality of OLS Estimator of Linear Trend Model

- We can verify the conditions of this central limit theorem for martingale difference array to show that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\lambda_1+\lambda_2\left(\frac{t}{T}\right)\right]u_t \xrightarrow{d} N\left(0,\sigma^2\lambda' Q\lambda\right)$$

for any $.\lambda=(\lambda_1,\lambda_2)'\in\mathbb{R}^2$ such that $\|\lambda\|=1$. As defined previously,

$$Q=\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix}.$$

- The Cramér-Wold device then allows us to deduce that

$$D_T^{-1}\sum_{t=1}^{T}x_t u_t=\begin{pmatrix} T^{-1/2}\sum_{t=1}^{T}u_t \\ T^{-1/2}\sum_{t=1}^{T}\left(\frac{t}{T}\right)u_t \end{pmatrix}\xrightarrow{d} N\left(0,\sigma^2 Q\right)$$

Asymptotic Normality of OLS Estimator of Linear Trend Model

- Finally, applying the continuous mapping theorem, we obtain, as $T \to \infty$

$$
\begin{aligned}
& D_T\left(\widehat{\theta}_T - \theta\right) \\
= & \left[D_T^{-1}\sum_{t=1}^{T} x_t x_t' D_T^{-1}\right]^{-1} D_T^{-1}\sum_{t=1}^{T} x_t u_t \xrightarrow{d} Q^{-1} N\left(0, \sigma^2 Q\right)
\end{aligned}
$$

More succinctly, we have

$$
D_T\left(\widehat{\theta}_T - \theta\right) \xrightarrow{d} N\left(0, \sigma^2 Q^{-1}\right) \quad \text{as } T \to \infty.
$$

Asymptotic Normality of OLS Estimator of Linear Trend Model

- **Remark:** Since

$$D_T\left(\widehat{\theta}_T-\theta\right)=\left[\begin{array}{c}\sqrt{T}\left(\widehat{\alpha}_T-\alpha\right)\\ T^{3/2}\left(\widehat{\beta}_T-\beta\right)\end{array}\right]$$

this result shows that

$$\begin{aligned}\widehat{\alpha}_T-\alpha &= O_p\left(\frac{1}{\sqrt{T}}\right),\\ \widehat{\beta}_T-\beta &= O_p\left(\frac{1}{T^{3/2}}\right),\end{aligned}$$

so that $\widehat{\alpha}_T$ and $\widehat{\beta}_T$ have different rates of convergence.