Econometrics 2b: Time Series Data
Handout #3

Matthew J. Lindquist*

March 2009

1 Univariate Time Series Models

- This section will address univariate time series modeling
- Imagine we have the following, general univariate time series model

$$Y_t = f\left(Y_{t-1}, Y_{t-2}, ..., u_t\right).$$

- To make this model operational we must specify three things;
  1. the functional form of $f\left(\right)$,
  2. the number of lags,
  3. and the structure of the disturbance term, u_t.
- Rationale for univariate analysis
  1. Purely statistical (atheoretical) models can oftentimes be extremely useful for summarizing information about a time series and for making reliable short-term forecasts.
  2. Look at the individual data series BEFORE running your regressions! They may be able to tell you a lot.
     (a) Seasonal patterns?
     (b) Long-run trends?
     (c) Structural breaks and/or unusual historical events?

*Department of Economics, Stockholm University, SE-106 91 Stockholm, Sweden. Tel. 46+8+163831. E-mail: Matthew.Lindquist@ne.su.se.

3. Theoretical models with lagged dependent variables can many time be reduced to meaningful univariate time series models.

Example: A simple macroeconomic model

$$\begin{aligned} c_t &= \beta_0 + \beta_1 y_t + \beta_2 c_{t-1} + \varepsilon_t && (1) \\ y_t &= c_t + i_t && (2) \\ i_t &= s y_t && (3) \end{aligned}$$

where $0 < s < 1$ is a constant saving rate, i_t is investment, c_t is consumption, y_t is income and $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$. Some algebra gives us

$$\begin{aligned} c_t &= \beta_0 + \beta_1 y_t + \beta_2 c_{t-1} + \varepsilon_t && (4) \\ y_t &= c_t + s y_t. && (5) \end{aligned}$$

Substituting y_t from Equation 5 into Equation 4 gives us

$$\begin{aligned} c_t &= \beta_0 + \frac{\beta_1 c_t}{1-s} + \beta_2 c_{t-1} + \varepsilon_t \Longrightarrow \\ c_t - \frac{\beta_1 c_t}{1-s} &= \left(1 - \frac{\beta_1}{1-s}\right) c_t = \beta_0 + \beta_2 c_{t-1} + \varepsilon_t \Longrightarrow \\ c_t &= \alpha_0 + \alpha_1 c_{t-1} + u_t \end{aligned}$$

where $u_t \sim N(0, \sigma_u^2)$. Thus, the macro model reduces to a univariate, $AR(1)$ process for consumption.

4. Time series models are important for theoretical, analytical and numerical simulation methods. Not just for econometrics.

1.1 Some Basic Concepts

Stochastic Process A stochastic process is a collection of random variables ordered in time.

- Let annual US GDP from 1960 to 2002 $\equiv \{Y_t\}_{1960}^{2002} = \{Y_{1960}, Y_{1961}, ..., Y_{2002}\}$. Each Y_t is a random variable (i.e., a single realization drawn from an infinite number of alternative realities). The stochastic process $\{Y_t\}_{1960}^{2002}$ helps us to describe and draw inferences concerning the development of the US GDP over time.

Strict Stationarity implies that the joint probability distribution of a stochastic variable is invariant over time.

Weak Stationarity implies that the first two moments of the joint probability distribution of a stochastic variable (i.e., the mean and variance-covariance matrix) are invariant over time, that is

$$E\{Y_t\} = \mu < \infty$$

$$V\{Y_t\} = E\{(Y_t - \mu)^2\} = \gamma_0 < \infty$$

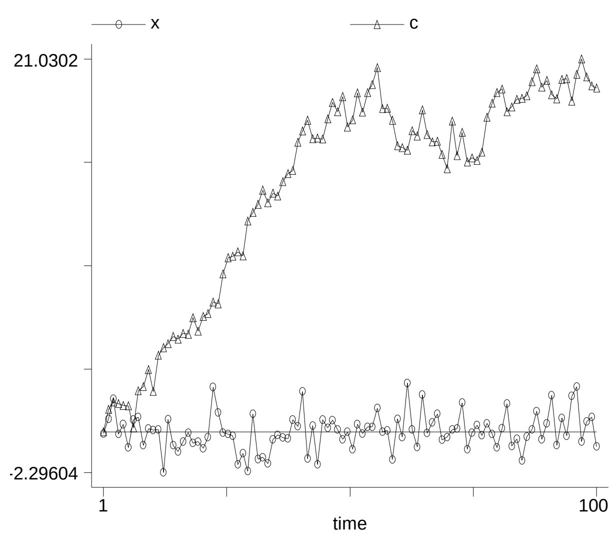
$$cov\{Y_t, Y_{t-k}\} = E\{(Y_t - \mu)(Y_{t-k} - \mu)\} = \gamma_k,\ k = 1, 2, 3...$$

- When are weak- and strict stationarity equivalent?
- For now, we need only be concerned with the concept of weak stationarity (henceforth stationarity). It suffices for most all of our needs (i.e., it allows us to proceed with our standard estimation, prediction and inference tools intact).

Nonstationary Process A nonstationary process will have a time-varying mean and/or a time-varying variance. This prevents us from making inferences outside of the sample period.

Purely Random Process A stochastic process which has a zero mean, constant variance, and is serially uncorrelated is called a purely random process, or, "white noise". We will quite frequently assume that the shocks in our econometric models are purely random processes (and then test whether this is a good approximate or not), i.e., we will assume

$$\varepsilon_t \sim IID(0, \sigma^2).$$



Variable X is Stationary and Variable C is Nonstationary.

- Question: What type of economic variables look like X? What type of economic variables look like C?

- Question: In *theory*, what type of economic variables should look like X and what type of variables should look like C?

1.2 Moving Average Processes

- An $MA(1)$ Process

$$Y_t = \mu + \varepsilon_t + \alpha\varepsilon_{t-1}$$
$$\varepsilon_t \sim IID(0, \sigma^2).$$

- Calculate the mean of Y_t

$$\begin{aligned} E\{Y_t\} &= E\{\mu + \varepsilon_t + \alpha\varepsilon_{t-1}\} \\ &= E\{\mu\} + E\{\varepsilon_t\} + E\{\alpha\varepsilon_{t-1}\} \\ &= \mu \end{aligned}$$

- Calculate the variance of Y_t

$$\begin{aligned} V\{Y_t\} &= E\{(Y_t - \mu)^2\} \\ &= E\{(\varepsilon_t + \alpha\varepsilon_{t-1})^2\} \\ &= E\left\{\varepsilon_t^2 + 2\alpha\varepsilon_t\varepsilon_{t-1} + \alpha^2\varepsilon_{t-1}^2\right\} \\ &= E\left\{\varepsilon_t^2\right\} + 2\alpha E\{\varepsilon_t\varepsilon_{t-1}\} + \alpha^2 E\left\{\varepsilon_{t-1}^2\right\} \\ &= E\left\{\varepsilon_t^2\right\} + \alpha^2 E\left\{\varepsilon_{t-1}^2\right\} \\ &= (1+\alpha^2)\sigma^2. \end{aligned}$$

- Calculate the autocovariance between Y_t and Y_{t-1}

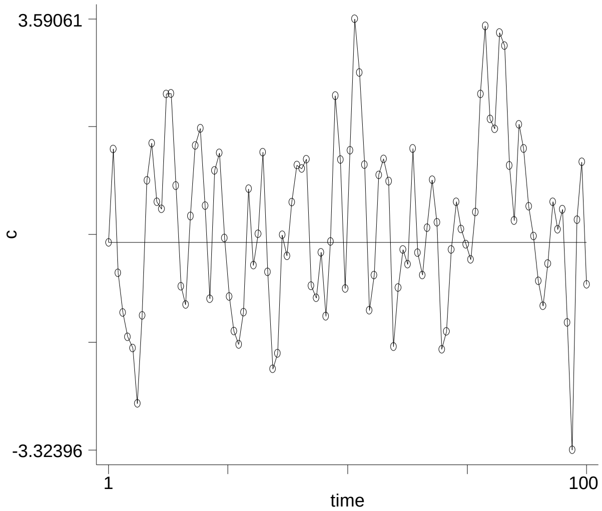
$$\begin{aligned} cov\{Y_t, Y_{t-1}\} &= E\{(Y_t - \mu)(Y_{t-1} - \mu)\} \\ &= E\{(\varepsilon_t + \alpha\varepsilon_{t-1})(\varepsilon_{t-1} + \alpha\varepsilon_{t-2})\} \\ &= \alpha E\left\{\varepsilon_{t-1}^2\right\} \\ &= \alpha\sigma^2. \end{aligned}$$

- Calculate the autocovariance between Y_t and Y_{t-2}

$$\begin{aligned} cov\{Y_t, Y_{t-2}\} &= E\{(Y_t - \mu)(Y_{t-2} - \mu)\} \\ &= E\{(\varepsilon_t + \alpha\varepsilon_{t-1})(\varepsilon_{t-2} + \alpha\varepsilon_{t-3})\} \\ &= 0. \end{aligned}$$

  In general, $cov\{Y_t, Y_{t-k}\} = 0$ for $k = 2, 3, 4, ...$

- What does the variance-covariance matrix Σ look like?



An MA(1) Process With $\alpha = 0.90$.

- If $|\alpha| < 1$, then we can transform (invert) the $MA(1)$ process into and $AR(\infty)$ process (more on this below).
- An example of a $MA(q)$ process

$$Y_t = \mu + \varepsilon_t + \alpha_1 \varepsilon_{t-1} + ... + \alpha_q \varepsilon_{t-q}$$

$$\varepsilon_t \sim IID(0, \sigma^2).$$

- The invertabity condition for general MA(q) processes is that the inverted roots of the lag polynomial lie inside the unit circle.

1.3 Autoregressive Processes

- An example of an $AR(1)$ process

$$Y_t = \mu + \theta Y_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim IID(0, \sigma^2).$$

- Calculate the mean of Y_t

$$\begin{aligned}
E\{Y_t\} &= E\{\mu + \theta Y_{t-1} + \varepsilon_t\} \\
&= \mu + E\{\theta Y_{t-1}\} + E\{\varepsilon_t\} \\
&= \mu + \theta E\{Y_{t-1}\} + E\{\varepsilon_t\} \\
&= \mu + \theta E(\mu + \theta Y_{t-2} + \varepsilon_{t-1}) \\
&= \mu + \theta\mu + \theta^2 E\{Y_{t-2}\} + \theta E\{\varepsilon_{t-1}\} \\
&= \mu + \theta\mu + \theta^2 E(\mu + \theta Y_{t-3} + \varepsilon_{t-2}) \\
&= etc... = \mu\left(1 + \theta + \theta^2 + ... + \theta^\infty\right) = \frac{\mu}{1-\theta} \text{ if } |\theta| < 1.
\end{aligned}$$

- The rule pertaining to the sum of a geometric series will be used a lot in this course.

Geometric Series Theorem: If $|r| < 1$, the geometric series $a + ar + ar^2 + ... + ar^{n-1} + ...$ converges to $a/(1-r)$ as $n \to \infty$. If $|r| \geq 1$, the series diverges unless $a = 0$. If $a = 0$, the series converges to 0.

- Calculate the variance of Y_t. First, define $y_t \equiv Y_t - \mu$. Then, calculate $v\{y_t\}$

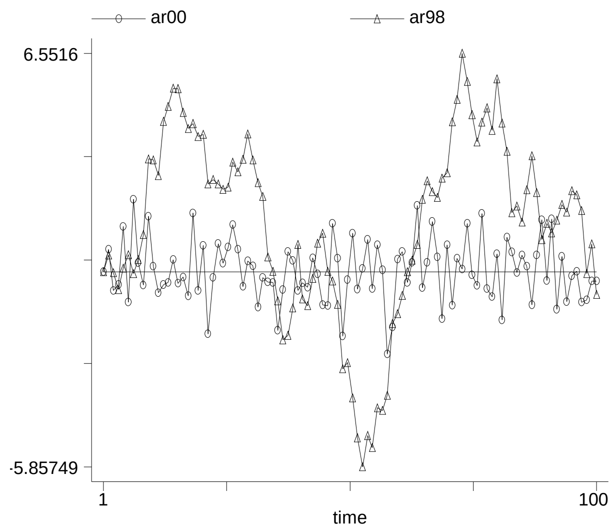
$$\begin{aligned}
V\{y_t\} &= E\{(y_t - E\{y_t\})^2\} = E\{(\theta y_{t-1} + \varepsilon_t)^2\} \\
&= E\left\{\varepsilon_t^2\right\} + \theta^2 E\left\{y_{t-1}^2\right\} = \sigma^2 + \theta^2 E\left\{y_{t-1}^2\right\} \\
&= \sigma^2 + \theta^2 E\{(\theta y_{t-2} + \varepsilon_{t-1})^2\} = etc... \\
&= \sigma^2\left(1 + \theta^2 + \theta^4 + ... + \theta^\infty\right) = \frac{\sigma^2}{1-\theta^2} \text{ if } |\theta| < 1.
\end{aligned}$$

- Calculate the autocovariance between Y_t and Y_{t-1}

$$\begin{aligned}
cov\{Y_t, Y_{t-1}\} &= E\{(Y_t - \mu)(Y_{t-1} - \mu)\} \\
&= E\{(\mu + \theta Y_{t-1} + \varepsilon_t - \mu)(Y_{t-1} - \mu)\} \\
&= E\{(\theta Y_{t-1} + \varepsilon_t)(Y_{t-1} - \mu)\} \\
&= E\left\{\theta Y_{t-1}^2 - \mu\theta Y_{t-1} + \varepsilon_t Y_{t-1} - \mu\varepsilon_t\right\} \\
&= E\left\{\theta Y_{t-1}^2 - \mu\theta Y_{t-1}\right\} = \theta E\left\{Y_{t-1}^2 - \mu Y_{t-1}\right\} \\
&= \theta E\{(\mu + \theta Y_{t-2} + \varepsilon_{t-1})(\mu + \theta Y_{t-2} + \varepsilon_{t-1}) - \mu(\mu + \theta Y_{t-2} + \varepsilon_{t-1})\} \\
&= \begin{array}{c} \theta E\{\mu^2 + 2\mu\theta Y_{t-2} + 2\mu\varepsilon_{t-1} + \theta^2 Y_{t-2}^2 + 2\theta Y_{t-2}\varepsilon_{t-1} + \\ ...\varepsilon_{t-1}^2 - \mu^2 - \mu\theta Y_{t-2} - \mu\varepsilon_{t-1}\} \end{array} \\
&= \theta E\left\{\mu\theta Y_{t-2} + \mu\varepsilon_{t-1} + \theta^2 Y_{t-2}^2 + 2\theta Y_{t-2}\varepsilon_{t-1} + \varepsilon_{t-1}^2\right\} \\
&= \theta E\left\{\mu\theta Y_{t-2} + \theta^2 Y_{t-2}^2 + \varepsilon_{t-1}^2\right\} = \theta^2 E\left\{\theta Y_{t-2}^2 + \mu Y_{t-2}\right\} + \theta\sigma^2 \\
&= +...+ = \theta^3 E\left\{\theta Y_{t-3}^2 + \mu Y_{t-3}\right\} + \theta^2\sigma^2 + \theta\sigma^2 \\
&= +...+ = \theta^\infty\sigma^2 + ... + \theta^2\sigma^2 + \theta\sigma^2 \\
&= \theta\sigma^2\left(1 + \theta + \theta^2 + ... + \theta^\infty\right) = \theta\frac{\sigma^2}{1-\theta^2} \text{ if } |\theta| < 1.
\end{aligned}$$

- In general, the autocovariance between Y_s and Y_k can be written as

$$cov\{Y_s, Y_t\} = \theta^{|s-t|} \frac{\sigma^2}{1-\theta^2} \text{ if } |\theta| < 1.$$



An AR(1) Process With $\theta = 0.98$ vs. White Noise.

- The stationarity condition for an $AR(1)$ process is $|\theta| < 1$. The import case of when $\theta = 1$ will be addressed shortly.
- An example of an $AR(p)$ process

$$Y_t = \mu + \theta_1 Y_{t-1} + ...\theta_p Y_{t-p} + \varepsilon_t$$

$$\varepsilon_t \sim IID(0, \sigma^2).$$

- The stationarity condition for general $AR(p)$ processes is that the inverted roots of the lag polynomial lie inside the unit circle.
- A stationary $AR(1)$ process can also be written as an $MA(\infty)$ process. This can be readily shown if we make use of the lag operator, L.

1.3.1 The Lag Operator, L

- See pages 38-41 in Enders.
- Some basic rules for using lag operators:

1. $L\alpha = \alpha$
2. $L(x_t) = x_{t-1}$
3. $L^2(x_t) = L[L(x_t)] = L(x_{t-1}) = x_{t-2}$
4. $L^s(x_t) = x_{t-s}$
5. $(1-L)\,x_t = x_t - L(x_t) = x_t - x_{t-1} = \Delta x_t$
6. $L(1-L)\,x_t = (1-L)x_{t-1} = x_{t-1} - Lx_{t-1} = x_{t-1} - x_{t-2} = \Delta x_{t-1}$

- Now, let's use these rules to show that an $AR(1)$ process can be written as an $MA(\infty)$ process (and vice versa). First, write down an $AR(1)$ process

$$x_t = \alpha x_{t-1} + \varepsilon_t$$

where $|\alpha| < 1$. Using rule #1 and #2 $\Longrightarrow$

$$\begin{aligned} x_t &= \alpha L x_t + \varepsilon_t \Longrightarrow \\ x_t - \alpha L x_t &= \varepsilon_t \Longrightarrow \\ (1-\alpha L)\,x_t &= \varepsilon_t \Longrightarrow \\ x_t &= \frac{\varepsilon_t}{(1-\alpha L)}. \end{aligned}$$

The last step is only possible when $(1-\alpha L)$ is invertible. It is here, since we are dealing with a stationary $AR(1)$ process with $|\alpha| < 1$. Now recall that the sum of the infinite geometric series $\sum_{s=0}^{\infty} \alpha^s L^s$ is equal to $\frac{1}{(1-\alpha L)}$, which allows us to re-write the above equation as

$$x_t = \sum_{s=0}^{\infty} \alpha^s L^s \varepsilon_t.$$

Using rule #4

$$x_t = \sum_{s=0}^{\infty} \alpha^s \varepsilon_{t-s}.$$

which, of course, is an MA(∞) process, $x_t = \varepsilon_t + \alpha\varepsilon_{t-1} + \alpha^2\varepsilon_{t-2} + ...\mathbb{R}$ makes use of this fact in order to estimate models with $MA(q)$ components.

1.3.2 The Random Walk Model (Without Drift)

- let us now take the $AR(1)$ model from above and set $\theta = 1$ and $\mu = 0$. This results in the well known random walk model

$$Y_t = Y_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim IID(0, \sigma^2).$$

- Is this process stationary? No. Why not?

$$
\begin{aligned}
E\{Y_T\} &= E\{Y_{T-1} + \varepsilon_T\} = \{Y_{T-2} + \varepsilon_{T-1} + \varepsilon_T\} \\
&= ... = E\left\{Y_0 + \Sigma_0^T \varepsilon_t\right\} = Y_0
\end{aligned}
$$

The mean is a finite constant. The variance, however, is not

$$
\begin{aligned}
V\{Y_t\} &= E\{(Y_t - E\{Y_t\})^2\} = E\{([Y_{t-1} + \varepsilon_t] - Y_0)^2\} = \\
&= E\{([(Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t] - Y_0)^2\} \\
&= E\{([([Y_{t-3} + \varepsilon_{t-2}] + \varepsilon_{t-1}) + \varepsilon_t] - Y_0)^2\} \\
&= +...+ = (E\{Y_0\} + \Sigma_0^T E\{\varepsilon_t\} - Y_0)^2 \\
&= (Y_0 + \Sigma_0^T E\{\varepsilon_t\} - Y_0)^2 = (\Sigma_0^T E\{\varepsilon_t\})^2 \\
&= \Sigma_0^T E\left\{\varepsilon_t^2\right\} = T\sigma^2.
\end{aligned}
$$

- Note the persistence of the shocks, ε_t, in this random walk model. The random walk model is said to have infinite memory.

Unit Root Setting $\theta = 1$ gives rise to the unit root problem. A unit root is synonymous with nonstationarity. The name unit root is due to the fact that the solution to the characteristic equation of an $AR(1)$ process has one root equal to unity when $\theta = 1$.

Difference Stationary Process The random walk model can be made stationary by differencing the time series Y_t

$$
\begin{aligned}
Y_t - Y_{t-1} &= Y_{t-1} - Y_{t-1} + \varepsilon_t \rightarrow \\
\Delta Y_t &= \varepsilon_t.
\end{aligned}
$$

We will make use of this fact quite often in this course. Processes which can be made stationary by differencing are called difference stationary processes.

Integrated Stochastic Process A stochastic process which can be made stationary through differencing is also known as an integrated stochastic process. If a stochastic process can be made stationary by taking first differences, it is said to be integrated of order one, I(1). If a stochastic process can be made stationary by taking d differences, it is said to be integrated of order d, I(d). Processes which are already stationary are said to be integrated of order zero, I(0).

1.3.3 The Random Walk Model With Drift

- let us now take the $AR(1)$ model from above and, once again, set $\theta = 1$. This time, however, we will allow $\mu \neq 0$. This results in a random walk model with drift

$$
\begin{aligned}
Y_t &= \mu + Y_{t-1} + \varepsilon_t \\
\varepsilon_t &\sim IID(0, \sigma^2).
\end{aligned}
$$

Stochastic Trend Y_t will tend to drift upwards or downwards, depending on the sign of the drift parameter, μ. This drift is called a stochastic trend.

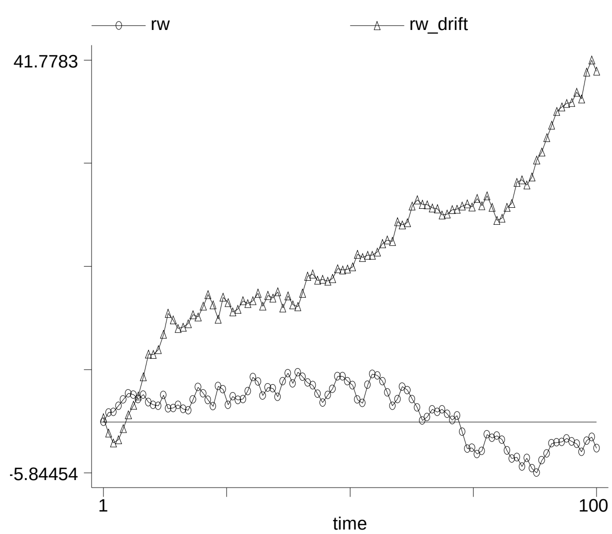
- Neither the mean nor the variance are constant

$$
\begin{aligned}
E\{Y_T\} &= \mu + Y_{T-1} + \varepsilon_T \\
&= \mu + \mu + Y_{T-2} + \varepsilon_{T-1} + \varepsilon_T \\
&= ... = Y_0 + \Sigma_0^T \mu + \Sigma_0^T \varepsilon_t \\
&= E\left\{Y_0 + \Sigma_0^T \mu + \Sigma_0^T \varepsilon_t\right\} = Y_0 + T\mu.
\end{aligned}
$$

$$
\begin{aligned}
V\{Y_t\} &= E\{(Y_t - E\{Y_T\})^2\} = E\{(Y_t - E\left\{Y_0 + \Sigma_0^T \mu + \Sigma_0^T \varepsilon_t\right\})^2\} \\
&= E\{(Y_t - Y_0 - T\mu + E\left\{\Sigma_0^T \varepsilon_t\right\})(Y_t - Y_0 - T\mu + E\left\{\Sigma_0^T \varepsilon_t\right\})\} \\
&= E\left\{\Sigma_0^T \varepsilon_t^2\right\} = \Sigma_0^T E\left\{\varepsilon_t^2\right\} = T\sigma^2.
\end{aligned}
$$

- The random walk model with drift can also be made stationary by differencing the time series Y_t

$$
\begin{aligned}
Y_t - Y_{t-1} &= \mu + Y_{t-1} - Y_{t-1} + \varepsilon_t \rightarrow \\
\Delta Y_t &= \mu + \varepsilon_t \\
\text{where } \varepsilon_t &\sim IID(0, \sigma^2).
\end{aligned}
$$



Random Walk vs. Random Walk With Drift.

1.3.4 Trend Stationary Processes

- Now let us add a time trend, t, to an otherwise stationary $AR(1)$ process

$$Y_t = \mu + t + \theta Y_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim IID(0, \sigma^2).$$

- This is a nonstationary process with mean

$$E\{Y_t\} = \frac{\mu + t}{1-\theta}$$

  and variance

$$V\{Y_t\} = \frac{\sigma^2}{1-\theta^2}.$$

Trend Stationary Processs This model can be made stationary by removing, or, subtracting the deterministic time trend from the original time series.[1] Processes which can be made stationary by removing a deterministic trend are called trend stationary processes (TSP).

[1] H-P trends and quadratic, cubic, etc. trends are also deterministic even though they are not linear.