6: Stationarity

GECO 6281 Advanced Econometrics 1 (Lab)

Patrick Mokre

Fall 2020

Recapitulation

- OLS and GLS
- Static Panel Data: LSDV, FE and RE estimation
- IV Regression
- Dynamic Panel Data: Anderson-Hsiao and Arellano-Bond instruments

Stationarity Matters (in time series)

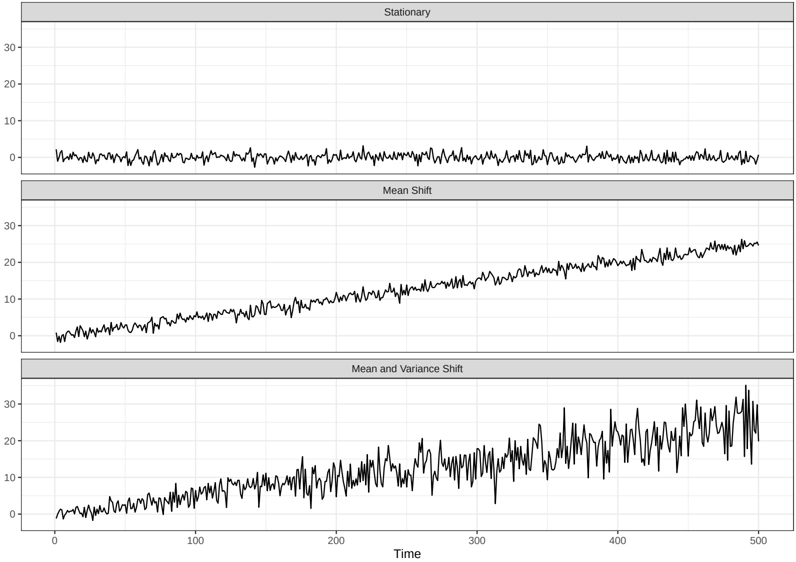
*A process is called **strictly stationary** if its properties are unaffected by a change of time origin.* (Verbeek, 2004, 258)

- Most econometrics investigation are concerned with **moments** of the joint distribution, i.e. **expected values** of some transformation of the data. This information is only meaningful if it is the same at any location within the data.
- We avoid **spurious regressions**, i.e. finding causality/correlation where there is none due to shared trends.
- **Wold's decomposition theorem** (every time series is made up by one deterministic and one stochastic time series) holds only for stationary data.

Consequences: Economists will often transform data until it looks stationary (first differencing, signal-noise filtering, ...). This makes interpretation challenging.

Stationarity Matters for us: Panel data has a time dimension, temporal interactions are exciting, but complex.

Stationary and Non-Stationary Processes



Stationarity is not the end of the story

- Stationarity really only matters because it **makes things easier**. There are interesting questions that go beyond stationary processes.
- Economic data is subject to a legal-political framework, which includes the possibility of structural breaks. It might be stationary for some time, but not for longer periods.
- Economic data is only observed for relatively short time periods. Stationarity is an **asymptotic property**.
- Only rely on stationarity if you absolutely have to. Many non-stationary methods of analysis do not require it and provide deeper insight!

How do we measure stationarity

- Engineers have it easier: "Stationarity can be defined in precise mathematical terms, but for our purpose we mean a flat looking series [...]" (NIST/SEMATECH e-Handbook of Statistical Methods, https://www.itl.nist.gov/div898/handbook/pmc/section4/pmc442.htm)
- Frequentist Econometricians have it harder: We pretend to attempt and find evidence **against** our hypothesis, then are relieved when we don't find it.
- Essentially stationarity/unit root testing is defining how a distribution could look if it was **not stationary**, then compare our data to this.

Defining Stationarity

Strict Stationarity: A process has the same statistical properties (moments) at any given point in time. The joint distribution of $(X_t, X_{t-1}, ..., X_{t-j})$ is the same as $(X_{t-k}, X_{t-k-1}, ..., X_{t-k-j})$ $\forall t, j, k$.

Most of the time, we are satisified with the first two moments of the joint distribution (**mean** and **variance**).

Covariance Stationarity ("Weak Stationarity"): $\forall t, k, h$

$$E[X_t] = \mu < \infty \quad (1)$$

$$V[X_t] = \sigma^2_X = \gamma(0) <= \gamma_0 < \infty \quad (2)$$

$$cov(X_t, X_{t-k}) = cov(X_{t-h}, X_{t-k-h}) = \gamma(k) = \gamma_k \quad (3)$$

Autocovariance and Autocorrelation

Covariance stationary (or 'weakly stationary') processes are characterized by their time-constant autocovariance function $\gamma(.) : \mathbb{Z} \to \mathbb{R}$ and equivalently their autocorrelation function (ACF) $\rho(.) : \mathbb{Z} \to [-1, 1]$.

The autocorrelation function is **scaleable**.

$$\rho(k) = \rho_k = corr(X_t, X_{t-k}) = \frac{\gamma_k}{\gamma_0} \tag{4}$$

Note that $\rho(0) = 1$ and $\rho(-j) = \rho(j)$.

We will investigate covariance stationarity using the autocorrelation function.

Time Series Processes

Autoregressive Processes (AR): The past values of the dependent variable $y_{t-1}, ..., y_{t-p}$ explain the current realization y_t.

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \epsilon_t$$

Moving Average Processes (MA): Past values of the error term ($\equiv$ deviation from the mean) $\epsilon_{t-1}, ..., \epsilon_{t-q}$ explain the current realization y_t.

$$y_t = \epsilon_t + \alpha_1 \epsilon_{t-1} + \alpha_2 \epsilon_{t-2} + ... + \alpha_q \epsilon_{t-q}$$

Stationary Processes

- AR processes $X_t = \delta + \sum_{i=1}^{p} \phi_i X_{t-i} + \epsilon_t$ with $|\phi| \geq 1$ are not stationary.
- AR processes $X_t = \delta + \sum_{i=1}^{p} \phi_i X_{t-i} + \epsilon_t$ with $|\phi| < 1$ are not **generally stationary**. They become more stationary over time (**asymptotically stationary**) and are stationary if started in the **infinite past** or if the initial value is drawn from specific distributions.
- MA processes of finite order are **always stationary** except the few starting observations for which the lags do not (fully) exist.

In other words: AR processes are stationary if they have an MA representation via the **lag operator**:

$$\begin{aligned} LX_t &= X_{t-1} \\ X_t &= \delta + \phi_1 X_{t-1} + ...\phi_p X_{t-p} + \epsilon_t \\ (1 - \phi_1 L - ...\phi_p L_p)X_t &= \delta + \epsilon_t \\ \phi(L)X_t &= \epsilon_t \end{aligned} \tag{5}$$

Unit roots

Most tests for stationarity test for the existence (and seek to reject) a **unit root**.

A process is said to have a unit root it the **characteristic equation** of a process has one root m (i.e. one solution when it is set to zero) equal to one.

$$
\begin{aligned}
X_t &= \delta + \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \epsilon_t \\
0 &= 1 - \phi_1 m - ... - \phi_p m^p
\end{aligned} \tag{6}
$$

Testing for Stationarity: Dickey Fuller

The most popular way of stationarity-related testing is testing for a unit root (i.e., trying to reject a unit root hypothesis). Take the **Dicker-Fuller Method** in its μ and τ variants.

$$\begin{aligned} X_t &= \delta + \phi X_{t-1} + \epsilon_t \\ DF_\mu &= \frac{\hat{\phi} - 1}{s.e.(\hat{\phi})} \end{aligned} \tag{7}$$

$$\begin{aligned} X_t &= \delta + \gamma * t + \phi X_{t-1} + \epsilon_t \\ DF_\tau &= \frac{\hat{\phi} - 1}{s.e.(\hat{\phi})} \end{aligned} \tag{8}$$

The Dickey-Fuller test statistic is not standard normal distributed under the null hypothesis, critical values have to be calculated using Monte Carlo simulations.

Testing for Stationarity: Augmented Dickey-Fuller

The Augmented Dickey-Fuller test statistic generalizes the DF logic to more lags and a more complex lag structure. The test statistic is a negative number, the smaller it is, the more strongly the Null hypothesis (unit root) is rejected against one of the possible alternative hypotheses (e.g. stationarity or trend stationarity).

$$\begin{aligned} \Delta X_t &= \delta + \pi X_{t-1} + \sum_{i=1}^{p} c_i \Delta X_{t-i} + \epsilon_t \\ ADF &= \frac{\hat{\pi}}{s.e.(\hat{\pi})} \end{aligned} \tag{9}$$

A time trend can easily be included into the regression equation. The optimal lag length is normally found using **information criteria**.

Testing Stationarity: KPSS Test

Rather than testing for non-stationarity, the test by Kwiatkowski, Perron, Shin and Smith test for a Null hypothesis of stationarity.

Every time series can be decomposed into (1) a deterministic time trend, (2) a random walk, and (3) a stationary error term (which is usually not white noise). KPSS tests if the random walk is stationary.

1 Run $X_t = \delta + \gamma * t + \epsilon_t$ and save the residuals

2 Calculate the partial error sums for each period t: $S_t = \sum_{i=1}^{t} e_i$

3 Calculate the error variance σ^2

$$KPSS = \sum_{i+1}^{T} \frac{S_t}{\sigma^2} \tag{10}$$

KPSS is test weaker than ADF, and produces many Type-1 errors, i.e. rejects too often. The test statistic is not standard normal distributed. The estimation of the long-run error variance in a time series is usually performed using Kernel estimation.

Stationarity in Panel Data: LLC (2002)

Levin-Lin-Chu (2002) test for balanced Panels. The test assumes equal AR(1) ρ coefficients in the panel, but allows for individual effects, time effects, and a time trend. It is carried in a three-step procedure.

H_0: Each time series contains a unit root. H_A: Each time series is stationary.

1 Run the ADF regression including appropriate deterministic variables (intercept, time trend, ...) and individual lag lengths p. $\Delta y_{i,t} = \delta y_{i,t-1} + \sum_{l=1}^{p} \theta_{i,l} \Delta y_{t-l} + \alpha_{m,i} d_{m,i}$

2 Create *orthogalized* residuals $\hat{e}_{i,t}$ and $\hat{v}_{i,t}$

2.1 $\hat{e}_{i,t} = \Delta y_{i,t} - \sum_{l}^{p} \hat{\theta}_l \Delta y_{i,t-l} - \hat{\alpha}_{m,i}$

2.2 $\hat{v}_{i,t-1} = y_{i,t-1} - \sum_{l}^{p} \hat{\theta}_l \Delta y_{i,t-l} - \hat{\alpha}_{m,i}$

3 Normalize the errors by the regression standard error from Step 1. $\tilde{e}_{i,t} = \hat{e}_{i,t} / \sigma^2_{\epsilon,i}$, $\tilde{v}_{i,t-1} = \hat{v}_{i,t-1} / \sigma^2_{\epsilon,i}$.

4 Estimate the ratio of short-run to long-run standard deviations, using a kernel (e.g. Bartlett kernel)

5 Pool the error terms by regressing $\tilde{e}_{i,t} = \delta \tilde{v}_{i,t-1} + u_{i,t}$

6 Calculate the t statistic for δ, adjust by the ratio from step 4

Stationarity in Panel Data: IPS (2003)

The **Im-Pesaran-Shin** test allows for individual AR(1) coefficients. It can thus be applied to unbalanced panels, as long as there are no time gaps in the data.

H_0: All time series contain a unit root H_A: Some time series do not contain a unit root