Causal Inference, Time Series and Economic History

2. Stationarity, Filtering and Seasonal Adjustment

Jason Lennard

Overview

- Non-stationarity
  - Definition, consequences, tests and solutions
- Filtering methods
  - OLS, Hodrick-Prescott, Hamilton and more
- Seasonality
  - OLS and beyond

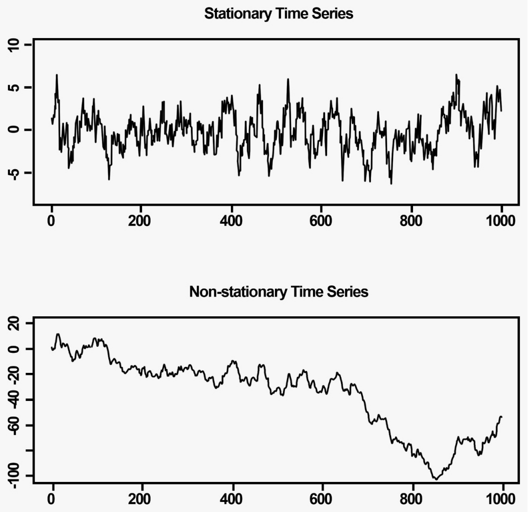
- *Class discussion paper:* Edvinsson, R., 'New annual estimates of Swedish GDP, 1800-2010', *Economic History Review*, 66 (2013), pp. 1101-26

Stationarity

A time series is (weakly) stationary if its mean, variance and autocovariance are constant over time

$$
\begin{aligned}
\mathit{Mean}: &\quad E(y_t) = \mu \\
\mathit{Variance}: &\quad Var(y_t) = \sigma^2 \\
\mathit{Autocovariance}: &\quad Cov(y_t, y_{t+h}) = \gamma_h
\end{aligned}
$$

Stationarity



Stationarity

Consequence

- The potential consequence of including non-stationary variables is spurious regression
- Spurious regression is "a problem that arises when regression analysis indicates a relationship between two or more unrelated time series processes simply because each has a trend, is an integrated time series (such as a random walk), or both" (Wooldridge, 2009)
- In practice, occurs when the regression residuals are non-stationary

Stationarity

Tests

Some common tests for stationarity are:

- The Dickey-Fuller (DF) test
- The Augmented Dickey-Fuller (ADF) test
- Many others

The Dickey-Fuller Test

- The starting point is the following model:

$$y_t = \rho y_{t-1} + u_t \tag{1}$$

- Non-stationarity implies that $\rho = 1$
- Why not just estimate equation (1) and test if $\rho = 1$?
- If y_t is non-stationary then the test is biased

The Dickey-Fuller Test

- To overcome this problem, subtract y_{t-1} from both sides of equation (1), $y_t = \rho y_{t-1} + u_t$:

$$y_t - y_{t-1} = \rho y_{t-1} - y_{t-1} + u_t \tag{2}$$

$$y_t - y_{t-1} = y_{t-1}(\rho - 1) + u_t \tag{3}$$

$$\Delta y_t = \delta y_{t-1} + u_t \tag{4}$$

- where Δ is the first difference operator and $\delta = \rho - 1$
- If $\hat{\delta} = 0$, we conclude that y_t is non-stationary, if $\hat{\delta} < 0$, we conclude that y_t is stationary
- The null hypothesis is $\delta = 0$ (y_t is non-stationary)
- To test this hypothesis, we use the τ test instead of the t test, which has different critical values

The Dickey-Fuller Test

- The model can take three forms, under three different null hypotheses:
  1. y_t is a random walk: $\Delta y_t = \delta y_{t-1} + u_t$
  2. y_t is a random walk with drift: $\Delta y_t = \beta_1 + \delta y_{t-1} + u_t$
  3. y_t is a random walk with drift around a deterministic trend: $\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + u_t$
- Each of the models has different critical values

The Augmented Dickey-Fuller Test

- An assumption of the DF test is that u_t is not serially correlated
- In case u_t is serially correlated, Dickey and Fuller developed the ADF test
- This involves adding lagged values of the dependent variable to the DF test equation

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta y_{t-i} + u_t \tag{5}$$

Solutions

- Cointegration
  - "A linear combination of two series, each of which is integrated of order one, is integrated of order zero" (Wooldridge, 2009)
- Transform the non-stationary variables so that they are stationary
  - E.g. first differences

Solutions

Best Practice

- Ramey (2016):

  A key question is how to specify a model when many of the variables may be trending. Sims et al.(1990) demonstrate that even when variables might have stochastic trends and might be cointegrated, the log levels specification will give consistent estimates. While one might be tempted to pretest the variables and impose the unit root and cointegration relationships to gain efficiency, Elliott (1998) shows that such a procedure can lead to large size distortions in theory. More recently, Gospodinov et al. (2013) have demonstrated how large the size distortions can be in practice. Perhaps the safest method is to estimate the [model] in log levels (perhaps also including some deterministic trends).

- In order to avoid spurious regression, the key issue is that the residuals are stationary

Filtering

- In various contexts in economic history it is useful to decompose a time series into different components
- For example, to measure:
  - Trends
  - Short-run cycles such as the business cycle (≈2-8 years)
  - Medium-run cycles such as financial cycles (≈8-30 years)
  - Long-run cycles such as Kondratieff cycles in economic growth (≥50 years)
  - Seasonality

OLS Filtering

In the case of a linear trend:

1. Estimate the following equation: $y_t = \alpha + \beta_1 t + u_t$
2. The trend is given by $\hat{\alpha} + \hat{\beta}_1 t$
3. The cycle is $\hat{u}_t$

In the case of a non-linear or quadratic trend:

1. Estimate the following equation: $y_t = \alpha + \beta_1 t + \beta_2 t^2 + u_t$
2. The trend is given by $\hat{\alpha} + \hat{\beta}_1 t + \hat{\beta}_2 t^2$
3. The cycle is $\hat{u}_t$

Hodrick-Prescott (HP) Filter

- The HP filter is useful to estimate the trend or cycle of a time series
- It produces a trend that looks like someone has "[run] a pen through the series to smooth out the wiggles" (Fregert, 2000)
- The filter was developed by Robert Hodrick and Edward Prescott
- It is widely used in central banks, international financial institutions and in academia
  - For example, a HP filter is used in the calculation of the countercyclical capital buffer as part of Basel III

Hodrick-Prescott Filter: Theory

- It begins with the idea that a time series, Y_t, is equal to a trend component, G_t, and a cyclical component, C_t, so that:

$$Y_t = G_t \times C_t$$

- Taking natural logarithms:

$$y_t = g_t + c_t$$

- This transformation eases the interpretation as the cycle then presents the *approximate percentage difference from the trend*
- Notice in this equation that there are two unobserved variables: g_t and c_t

Hodrick-Prescott Filter: Theory

- The HP filter picks values of g_t in each period that minimize:

$$\sum_{t=1}^{T}(y_t - g_t)^2 + \lambda \sum_{t=2}^{T-1}[(g_{t+1} - g_t) - (g_t - g_{t-1})]^2$$

- The first term minimizes the sum of squared cyclical fluctuations
- The second term minimizes the squared changes in the growth rate of the trend
- The parameter λ balances the trade-off between these objectives (smooth cycle versus smooth trend)
- Given y_t and g_t, we can now calculate the cyclical component, c_t, by rearranging the earlier equation so that $c_t = y_t - g_t$
- Let's look at an example in a spreadsheet

How to Pick λ? HP Rule

$$\lambda = 100 \times p^2$$

- where p is the number of periods per *year*
- Therefore:
  - $\lambda_{monthly} = 100 \times 12^2 = 14400$
  - $\lambda_{quarterly} = 100 \times 4^2 = 1600$
  - $\lambda_{yearly} = 100 \times 1^2 = 100$

How to Pick λ? Ravn-Uhlig Rule

$$\lambda = 1600 \times p^4$$

- where p is the number of periods per *quarter* (see Ravn and Uhlig (2002) for details)
- Therefore:
  - $\lambda_{monthly} = 1600 \times 3^4 = 129600$
  - $\lambda_{quarterly} = 1600 \times 1^4 = 1600$
  - $\lambda_{yearly} = 1600 \times 0.25^4 = 6.25$
- Stata follows this rule by default *if* the time variable is set to daily, weekly, monthly, quarterly, half-yearly or yearly. If not, the default value is 1600
- Ultimately, choice of λ is up to the researcher, but beware that a higher (lower) λ results in a more (less) volatile cycle

Hodrick-Prescott Filter: Pros and Cons

Strengths

- Simple to calculate
- Trend can change over time
- Reasonable to the naked eye

Weaknesses

- Sensitive to choice of λ
- The end-point problem (Mise et al., 2005)
  - A solution is to discard observations at beginning and end of series
  - Problematic for analysis in real time
- "Produces spurious dynamic relations that are an artefact of the filter and have no basis in the true data-generating process" (Hamilton, 2018)

Hamilton Filter

- The intuition of the Hamilton (2018) model is the following
- Let's try and forecast a time series, such as GDP in 2023
- In order to form a forecast of GDP for 2023, we might use the most recent data available to us (such as the GDP outturn in 2021, 2020 and so on)
- If in 2023, GDP is lower than we forecasted, something unexpected must have happened to move GDP away from its historic path, such as a financial crisis, for example
- The difference between the observed value and the forecast is regarded as the cyclical component of the time series

Hamilton Filter

- For quarterly data:

$$y_{t+8} = \beta_0 + \beta_1 y_t + \beta_2 y_{t-1} + \beta_3 y_{t-2} + \beta_4 y_{t-3} + u_{t+8}$$

- The trend is given by $\hat{\beta}_0 + \hat{\beta}_1 y_t + \hat{\beta}_2 y_{t-1} + \hat{\beta}_3 y_{t-2} + \hat{\beta}_4 y_{t-3}$
- The cycle is $\hat{u}_{t+8}$

Hamilton Filter

- For monthly data:

$$\begin{aligned}
y_{t+24} = \beta_0 + \beta_1 y_t + \beta_2 y_{t-1} + \beta_3 y_{t-2} + \beta_4 y_{t-3} + \beta_5 y_{t-4} \\
+ \beta_6 y_{t-5} + \beta_7 y_{t-6} + \beta_8 y_{t-7} + \beta_9 y_{t-8} + \beta_{10} y_{t-9} \\
+ \beta_{11} y_{t-10} + \beta_{12} y_{t-11} + u_{t+24} \quad (6)
\end{aligned}$$

- For yearly data:

$$y_{t+2} = \beta_0 + \beta_1 y_t + u_{t+2}$$

Other Filters

- There are many other filters, each with different properties
- For example:
    - Band-pass filter
    - Wavelet filter
    - Segmented trend model
    - Unobserved components model

Seasonality

- In addition to a trend and cycle, a monthly or quarterly time series may also contain a seasonal component
- A number of economic time series are "seasonally adjusted"
- A key example is the unemployment rate
- Unemployment is affected by seasonal factors such as weather, holidays, term dates in schools etc.
- This regular seasonal pattern is often removed so that we can meaningfully interpret changes in the time series from one period to the next

OLS Seasonal Adjustment

- Step 1: For a quarterly time series, estimate the following equation with a dummy variable for each quarter and no intercept (to avoid the dummy variable trap!):

$$y_t = \beta_1 D_{1t} + \beta_2 D_{2t} + \beta_3 D_{3t} + \beta_4 D_{4t} + u_t$$

- The β coefficients represent the mean unemployment rate for each quarter
- Step 2: Calculate the residuals from the equation, $\hat{u}_t$
- Step 3: Add the mean of the series to the residual ($\tilde{y}_t = \bar{y} + \widehat{u}_t$) for the seasonally adjusted series

OLS Seasonal Adjustment: Pros and Cons

Strengths

- Simple to calculate

Weaknesses

- Assumes that the seasonality is constant, but in reality could change over time
  - For example, as graduation rates increase over time, the degree of seasonal unemployment in the summer may increase relative to other times of the year

Other Approaches

- U.S. Census Bureau's X-13
- TRAMO-SEATS
- Harvey's (1989) unobserved components model

Class Discussion Paper: Edvinsson (2013)

Research Question

Class Discussion Paper: Edvinsson (2013)

Data

- Log real GDP per capita at constant prices

- HP filter with $\lambda = 100$ and $\lambda = 10000$

Class Discussion Paper: Edvinsson (2013)

Results

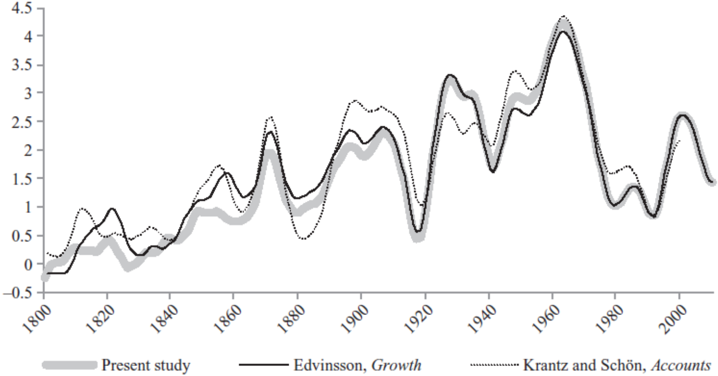


Figure 5. *Annual growth of the trend (HP-filter, λ = 100) in Swedish GDP per capita according to various studies, 1800–2010*

Note: The trend is calculated using an HP-filter (setting λ = 100) on the logarithms of real GDP per capita.
Sources: Online app. S1; Edvinsson, *Growth*, pp. 307–11; Krantz and Schön, *Accounts*, pp. 35–9.

Class Discussion Paper: Edvinsson (2013)

Results

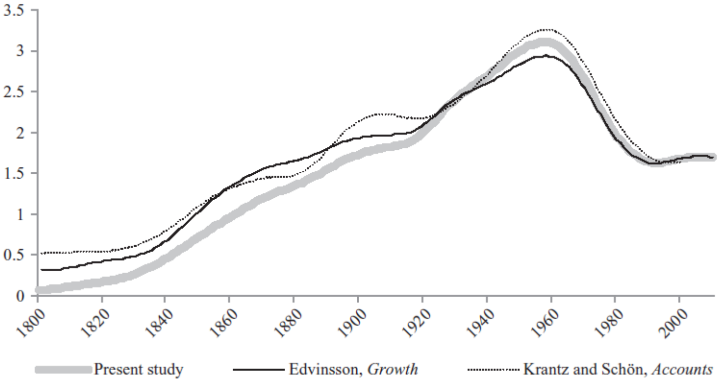


Figure 6. *Annual growth of the trend (HP-filter, λ = 10,000) in Swedish GDP per capita according to various studies, 1800–2010*

Note: The trend is calculated using an HP-filter (setting λ = 10,000) on the logarithms of real GDP per capita.
Sources: See fig. 5.

Next Class

- *Class discussion paper:* Kelly, M. and Ó Gráda, C., 'Numerare est errare: Agricultural output and food supply in England before and during the Industrial Revolution', *Journal of Economic History*, 73 (2013), pp. 1132-63

Further Reading

- Wooldridge, *Introductory Econometrics*, chapters 10 and 18