Basics of Time Series Analysis

Taeyong Park

Washington University in St. Louis

April 3, 2017
Overview

- **Time Series Data**: What is it? What is different?

- **Basic Concepts**: (Non-) Stationarity processes; Integrated variables; Unit root tests.

- **Univariate Models**: AR (Autoregressive) process; ARMA (Autoregressive moving average) process.

- **Dynamic Regression Models**: Autoregressive distributed lag model (ADL); Error correction model (ECM).

- **Brief Introduction to Some Advanced Topics**: Vector Autoregression (VAR) Models and Fractional Integration.
References

- Introduction to time series analysis

- Autoregressive distributed lag models and error correction models
  - *Time Series Symposium at Political Analysis* 24(1).

- Simultaneous-equation methods (vector autoregression models)
  - Patrick T. Brandt and John R. Freeman. 2009 “Modeling Macro-Political Dynamics.” *Political Analysis*. 
Examples in the social sciences

- Macroeconomic and financial data (daily, monthly, quarterly, yearly)
  - GDP growth, inflation, stock price indexes, etc.

- Political polls
  - Presidential approval ratings, congressional approval, macro-partisanship, consumer sentiment.

- Policy
  - Interest rates, exchange rates, welfare policies, foreign policy behavior (events data of cooperative and hostile actions between countries).

- Time Series Cross Section data and panel data
Quarterly Fluctuations of Presidential Approval and Income Growth: 1964–2015

Approval
Income growth


Time
Time Series Data: What is it? What is different?

- Time series analysis explains the temporal dependencies within and between social processes.

  - Within a social process (univariate models): The current value of a variable as a function of previous values of that same variable.
    
    \[ GDP_t = GDP_{t-1} + GDP_{t-2} + GDP_{t-3} + \varepsilon_t. \]

  - Between social processes (multivariate or dynamic regression models): The current value of a variable as a function of both previous values of that same variable and other variables.
    
    \[ \text{Pres. Approval}_t = \text{Pres. Approval}_{t-1} + \text{Pres. Approval}_{t-2} + \text{Inflation}_t + \text{Inflation}_{t-1} + \text{Inflation}_{t-2} + ... + \varepsilon_t. \]
Time Series Data: What is it? What is different?

- Primary objective of time series analysis: decomposing short and long-term processes of the temporal dependencies.
  - The effect of \( \text{Pres. Approval}_{t-1} \) on \( \text{Pres. Approval}_t \).
  - The effect of \( \text{Inflation}_t \) on \( \text{Pres. Approval}_t \).
  - The effect of \( \text{Inflation}_t + \text{Inflation}_{t-1} + \text{Inflation}_{t-2} \) on \( \text{Pres. Approval}_t \).

Main difference:
- In time series analysis data points taken over time may have an internal structure (such as autocorrelation, trend or seasonal variation) that should be accounted for.
- \( \star \) Losing important information and/or resulting in spurious inference (find an effect that in fact doesn’t exist).
Primary objective of time series analysis: decomposing short and long-term processes of the temporal dependencies.

- The effect of $\text{Pres.Approval}_{t-1}$ on $\text{Pres.Approval}_t$.
- The effect of $\text{Inflation}_t$ on $\text{Pres.Approval}_t$.
- The effect of $\text{Inflation}_t + \text{Inflation}_{t-1} + \text{Inflation}_{t-2}$ on $\text{Pres.Approval}_t$.

Main difference:

- In time series analysis data points taken over time may have an internal structure (such as autocorrelation, trend or seasonal variation) that should be accounted for.
  - Losing important information and/or resulting in spurious inference (find an effect that in fact doesn’t exist).
Stationary / Nonstationary Processes

- Spurious inference: A serious potential danger because of temporal dependencies among variables.
- It’s crucial to address possible sources of spurious inference.
  - Nonstationary time series.

Stationary stochastic processes
- A time series (e.g. GDP) is stochastic when an observed value at \( t \) (e.g. $3.5 billion) is a particular realization of all such possibilities.
  - A stochastic process is **stationary** when
    - Its mean and variance are constant over time.
    - The value of the covariance between the two time periods depends only on the distance or gap or lag between the two time periods and not the actual time at which the covariance is computed.
Consider \( Y_t \) as a stochastic time series with these properties:

\[
\text{Mean}: \quad E(Y_t) = \mu \tag{1}
\]

\[
\text{Variance}: \quad \text{Var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2 \tag{2}
\]

\[
\text{(Auto−)Covariance}: \quad \gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)] \tag{3}
\]

where \( \gamma_k \) is the covariance between the values of \( Y_k \) and \( Y_{t+k} \).

Suppose we shift the origin of \( Y \) from \( Y_t \) to \( Y_{t+m} \).

If \( Y_t \) is to be stationary,

\[ E(Y_t) = E(Y_{t+m}). \]
\[ \text{Var}(Y_t) = \text{Var}(Y_{t+m}). \]
\[ E[(Y_t - \mu)(Y_{t+k} - \mu)] = E[(Y_{t+m} - \mu)(Y_{t+m+k} - \mu)]. \]
A special case: a stochastic process is a white noise (purely random) process if

- it has zero mean,
- constant variance,
- and serially uncorrelated.

Example: The error term in the classical linear regression model.
- \( u_t \sim \text{IID Normal}(0, \sigma^2) \).
Stationary / Nonstationary Processes

Why are stationary time series important?

- A stationary time series tends to return to its mean and fluctuations around this mean has a broadly constant amplitude.
Why are stationary time series important?

- If a time series is nonstationary, we can study its behavior only for the time period under consideration.
- Each set of time series data is a particular episode.
- ⇒ Not generalizable.
Why are stationary time series important?

- If a time series is nonstationary, we can study its behavior only for the time period under consideration.
- Each set of time series data is a particular episode.
- Not generalizable.

- Spurious inference.
Stationary / Nonstationary Processes

- Spurious inference caused by the nonstationarity of time series.

  - Two random walk models:

    \[ Y_t = Y_{t-1} + u_t \]  \hspace{1cm} (4)
    \[ X_t = X_{t-1} + v_t \]  \hspace{1cm} (5)

  - 500 observations of \( u_t \) from \( u_t \sim N(0, 1) \) and 500 observations of \( v_t \) from \( v_t \sim N(0, 1) \)
  - \( u_t \) and \( v_t \) are randomly generated: serially uncorrelated as well as mutually uncorrelated.
  - \( Y_t \) and \( X_t \) are nonstationary and mutually uncorrelated.
  - There should no relationship between \( X \) and \( Y \); \( R^2 \) should tend to be zero.
Stationary / Nonstationary Processes

n <- 500
y <- double(n)
u <- rnorm(n)
y[1] <- rnorm(1)
for(i in 2:n){
    y[i] <- y[i-1] + u[i]
}

x <- double(n)
v <- rnorm(n)
x[1] <- rnorm(1)
for(i in 2:n){
    x[i] <- x[i-1] + v[i]
}

summary(lm(y~x))
Stationary / Nonstationary Processes

|               | Estimate | Std. Error | t value | Pr(>|t|) |
|---------------|----------|------------|---------|----------|
| (Intercept)   | -8.22183 | 0.64145    | -12.82  | <2e-16   |
| x             | -0.65737 | 0.06087    | -10.80  | <2e-16   |

---

Multiple R-squared: 0.1898, Adjusted R-squared: 0.1881

|               | Estimate | Std. Error | t value | Pr(>|t|) |
|---------------|----------|------------|---------|----------|
| (Intercept)   | -25.23491| 0.49940    | -50.53  | <2e-16   |
| x             | 0.68065  | 0.06194    | 10.99   | <2e-16   |

---

Multiple R-squared: 0.1951, Adjusted R-squared: 0.1935
Stationary / Nonstationary Processes

Why are stationary time series important?

- If a time series nonstationary, we can study its behavior only for the time period under consideration.
- Each set of time series data is a particular episode.
- ⇒ Not generalizable.

- Spurious inference.
Why are stationary time series important?

- If a time series nonstationary, we can study its behavior only for the time period under consideration.
- Each set of time series data is a particular episode.
- \( \Rightarrow \) Not generalizable.

- Spurious inference.

- The traditional regression methodology (e.g., OLS for continuous response data) can be applied.
  - For example, the autoregressive distributed lag model and the error correction model are consistently estimated by ordinary least squares (OLS) if stationarity process, white noise error, and weakly exogenous regressors hold true.
Tests for Stationarity

Graphical analysis for an initial cue: Has the mean of the series been changing?

Stationary Time Series

Non-stationary Time Series
Tests for Stationarity

The unit root test: A widely used test of stationarity.

Consider the following stochastic process:

\[ Y_t = \rho Y_{t-1} + u_t, \quad \text{where} \quad -1 \leq \rho \leq 1, \quad u_t \overset{iid}{\sim} N(0, \sigma^2). \]

If \( \rho = 1 \), \( Y_t \) is a unit root process or a random walk process.

- \( Y_t = Y_0 + \sum u_t \) (initial value + the sum of random shocks)
- The impact of a particular shock does not die away.
- If \( Y_0 = 0 \), \( \text{Var}(Y_t) = t \times \sigma^2. \implies \text{As } t \text{ increases, } \text{Var}(Y_t) \text{ increases; Non constant.} \)
Tests for Stationarity

The unit root test: A widely used test of stationarity.

Consider the following stochastic process:

\[ Y_t = \rho Y_{t-1} + u_t, \quad \text{where } -1 \leq \rho \leq 1, \quad u_t \sim iid N(0, \sigma^2). \]

If \( \rho = 1 \), \( Y_t \) is a unit root process or a random walk process.

- \( Y_t = Y_0 + \Sigma u_t \) (initial value + the sum of random shocks)
- The impact of a particular shock does not die away.
- If \( Y_0 = 0 \), \( Var(Y_t) = t \times \sigma^2 \). \( \Rightarrow \) As \( t \) increases, \( Var(Y_t) \) increases; Non constant.

If \( |\rho| < 1 \), the stationary process: the impact of a particular shock dissipates eventually.
Tests for Stationarity

The general idea of the unit root test is check if $\rho = 1$.

$$Y_t = \rho Y_{t-1} + u_t,$$

where $-1 \leq \rho \leq 1$, $u_t \overset{iid}{\sim} N(0, \sigma^2)$.

$$Y_t - Y_{t-1} = \rho Y_{t-1} - Y_{t-1} + u_t$$

$$\Delta Y_t = (\rho - 1) Y_{t-1} + u_t$$

$$\Delta Y_t = \delta Y_{t-1} + u_t$$

Test the null hypothesis that $\delta = 0$ (i.e. $\rho = 1$). The alternative hypothesis is $\delta < 0$. 
Tests for Stationarity

\[ \Delta Y_t = \delta Y_{t-1} + u_t \]

where:

- \( H_0 : \delta = 0 \)
- \( H_1 : \delta < 0 \)

- The t test?
  - Dickey and Fuller: Under the null hypothesis, the t value of the estimated coefficient of \( Y_{t-1} \) does not follow the t distribution, but follow the \( \tau \) statistic.

- The Dickey-Fuller (DF) test: One-sided test using the critical values of the \( \tau \) statistic.

- The Augmented Dickey-Fuller (DF) test: Addresses the possibility that \( u_t \) are correlated by adding the lagged values of the dependent variable.

\[ \Delta Y_t = \delta Y_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta Y_{t-i} + u_t \]

- In R, \texttt{adf.test()}, \texttt{ur.df()}, and others are available.
Univariate Models

- The autoregressive (AR) process
- The autoregressive and moving average (ARMA) process
- Forecasting
  - \( GDP_t = \alpha GDP_{t-1} + \varepsilon_t \)
  - \( Stock.Prices_t = \alpha_1 Stock.Prices_{t-1} + \alpha_2 Stock.Prices_{t-2} + \varepsilon_t \)
- Understanding the stochastic processes for individual variables in multivariate models.
The Autoregressive (AR) Process

**AR(1)**

\[ Y_t = \alpha_1 Y_{t-1} + u_t, \]

where \( u_t \) is white noise.

\[ Y_0 = u_0 \]
\[ Y_1 = \alpha_1 Y_0 + u_1 = \alpha_1 u_0 + u_1 \]
\[ Y_2 = \alpha_1 + u_2 = \alpha_1 (\alpha_1 u_0 + u_1) + u_2 = \alpha_1^2 u_0 + \alpha_1 u_1 + u_2 \]
\[ \vdots \]
\[ Y_t = \alpha_1 Y_{t-1} + u_t = \alpha_1^t u_0 + \alpha_1^{t-1} u_1 + \cdots + u_t \]
The Autoregressive (AR) Process

AR(1)

\[ Y_t = \sum_{i=0}^{\infty} \alpha^i u_{t-i} \]

An infinite sum of exponential weighted random shocks.

AR(2)

\[ Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + u_t. \]

AR(p): pth-order autoregressive process

\[ Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots + \alpha_p Y_{t-p} + u_t. \]
The Moving Average (MA) Process

MA(1)

\[ Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1}, \]

where \( u_t \) is white noise. \( Y \) at time \( t \) is equal to a constant plus a moving average of the current and past error terms.

\[ Y_t = u_t - \sum_{i=1}^{\infty} Y_{t-i} \]

An infinite sum of exponentially weighted past observations.

MA(p)

\[ Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \cdots + \beta_q u_{t-q}. \]
The Autoregressive and Moving Average (ARMA) Process

ARMA(1,1)

\[ Y_t = \theta + \alpha_1 Y_{t-1} + \beta_0 u_t + \beta_1 u_{t-1}. \]

ARMA(p,q)

\[ Y_t = \theta + \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + \beta_0 u_t + \beta_1 u_{t-1} + \cdots + \beta_q u_{t-q}. \]
Two workhorse models most frequently used in the social sciences:

- Autoregressive Distributed Lag Model
- Error Correction Model
Autoregressive Distributed Lag Model

- ADL or ARDL

- A general model ADL($p, q, k$):

\[
Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} \\
+ \beta_{10} X_{1t} + \beta_{11} X_{1t-1} + \cdots + \beta_{1q} X_{1t-q} \\
+ \beta_{20} X_{2t} + \beta_{21} X_{2t-1} + \cdots + \beta_{2q} X_{2t-q} + \cdots \\
+ \beta_{k0} X_{kt} + \beta_{k1} X_{kt-1} + \cdots + \beta_{kq} X_{kt-q} + \varepsilon_t.
\]

\[
Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{k} \sum_{i=0}^{q} \beta_{ji} X_{jt-i} + \varepsilon_t.
\]
Autoregressive Distributed Lag Model

- The ADL model is consistently estimated by ordinary least squares (OLS) if white noise error, stationarity process, and weakly exogenous regressors are assumed.
  - $\varepsilon_t \sim iid N(0, \sigma^2)$,
  - $|\sum_{i=1}^{p} \alpha_i| < 1$,
  - $E(\varepsilon_t, X_{js}) = 0, \forall t, s, j$.

- Autocorrelation may be a cause for concern – OLS estimator may be unbiased but does not have minimum variance.
  - Autocorrelation within a time series (e.g. between $X_{1t}, X_{1t-1}$ and $X_{1t-1}, X_{1t-2}$)
  - Autocorrelation (serial correlation) between different time series (e.g. between $X_{1t}, X_{1t-1}$ and $X_{2t-1}, X_{2t-2}$)

  - The Breusch-Godfrey Lagrange Multiplier (LM) test should be conducted to diagnose residual autocorrelation.
    - OLS vs. GLS (Generalized least squares).
Example

- \( \text{Pres. Approval} = f(\text{GDP.growth, income.growth, inflation, unemp.change}) \).

Example

- $\text{Pres. Approval} = f(\text{GDP.growth, income.growth, inflation, unemp.change})$.


- 1. Check for the stationarity of the individual variables.
Example

1. Check for the stationarity of the individual variables: The unit root test in R.

```r
library(tseries)
adf.test(APPROVAL)
```

Output:

```
Augmented Dickey-Fuller Test
data: APPROVAL
Dickey-Fuller = -4.4562, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary
```
Example

```r
adf.test(GDPLEVEL)

Augmented Dickey-Fuller Test
data:  GDPLEVEL
Dickey-Fuller = -1.9945, Lag order = 5,
p-value = 0.5784
alternative hypothesis: stationary

adf.test(GDPPCHT0)

Augmented Dickey-Fuller Test
data:  GDPPCHT0
Dickey-Fuller = -5.5462, Lag order = 5,
p-value = 0.01
alternative hypothesis: stationary
```
Example

- $Pres.\ Approval = f(GDP\ .\ growth, \ income\ .\ growth, \ inflation, \ unemp\ .\ change)$.


1. Check for the stationarity of the individual variables.
   - If yes, OLS is okay depending on the other assumptions.
   - If no, check for cointegration to see if the error correction model is appropriate or transform the variables.
Example

- \( \text{Pres. Approval} = f(\text{GDP.growth, income.growth, inflation, unemp.change}) \).


1. Check for the stationarity of the individual variables.
   - If yes, OLS is okay depending on the other assumptions.
   - If no, check for cointegration to see if the error correction model is appropriate or transform the variables.

2. Consider the causality issue to make sure that the explanatory variables are exogenous to the outcome.
   - Mostly, researchers rely on theoretical justifications. No formal test.
Example

- $\text{Pres. Approval} = f(\text{GDP.growth, income.growth, inflation, unemp.change})$.


1. Check for the stationarity of the individual variables.
   - If yes, OLS is okay depending on the other assumptions.
   - If no, check for cointegration to see if the error correction model is appropriate or transform the variables.

2. Consider the causality issue to make sure that the explanatory variables are exogenous to the outcome.
   - Mostly, researchers rely on theoretical justifications. No formal test.
   - $\text{Pres. Approval} \rightarrow \text{GDP.growth, income.growth, inflation, or unemp.change}$?
Example

- \( \text{Pres. Approval} = f(\text{GDP.growth}, \text{income.growth}, \text{inflation}, \text{unemp.change}) \).


1. Check for the stationarity of the individual variables.
   - If yes, OLS is okay depending on the other assumptions.
   - If no, check for cointegration to see if the error correction model is appropriate or transform the variables.

2. Consider the causality issue to make sure that the explanatory variables are exogenous to the outcome.
   - Mostly, researchers rely on theoretical justifications. No formal test.
   - \( \text{Pres. Approval} \rightarrow \text{GDP.growth}, \text{income.growth}, \text{inflation}, \text{or unemp.change} \)?

A general model ADL($p, q, k$):

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{k} \sum_{i=0}^{q} \beta_{ji} X_{jt-i} + \varepsilon_t.$$
ADL Modeling

A general model ADL($p$, $q$, $k$):

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{k} \sum_{i=0}^{q} \beta_{ji} X_{jt-i} + \varepsilon_t.$$ 

ADL modeling: How to select $p$, $q$, $k$.

$k$ is mostly guided by substantive theory.

$p$ and $q$?
ADL Modeling

- $p$ and $q$ in ADL($p$, $q$, $k$)?

$$\begin{align*}
Y_t &= \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{k} \sum_{i=0}^{q} \beta_{ji} X_{jt-i} + \varepsilon_t.
\end{align*}$$

$$\begin{align*}
\text{Appr}_t &= \alpha_0 + \alpha_1 \text{Appr}_{t-1} + \ldots + \alpha_p \text{Appr}_{t-p} \\
&+ \beta_{10} \text{GDP}_{1t} + \beta_{11} \text{GDP}_{1t-1} + \ldots + \beta_{1q} \text{GDP}_{1t-q} \\
&+ \beta_{20} \text{Income}_{2t} + \beta_{21} \text{Income}_{2t-1} + \ldots + \beta_{2q} \text{Income}_{2t-q} \\
&+ \beta_{30} \text{Inflation}_{3t} + \beta_{31} \text{Inflation}_{3t-1} + \ldots + \beta_{3q} \text{Inflation}_{3t-q} \\
&+ \beta_{40} \text{Unemp}_{4t} + \beta_{41} \text{Unemp}_{4t-1} + \ldots + \beta_{4q} \text{Unemp}_{4t-q} + \varepsilon_t.
\end{align*}$$
ADL Modeling

\[ ADL(1, 0, 4) : \]

\[ \text{Appr}_t = \alpha_0 + \alpha_1 \text{Appr}_{t-1} \]
\[ + \beta_1 \text{GDP}_1t + \beta_2 \text{Income}_2t + \beta_3 \text{Inflation}_3t + \beta_4 \text{Unemp}_4t + \varepsilon_t. \]

\[ ADL(1, 1, 4) : \]

\[ \text{Appr}_t = \alpha_0 + \alpha_1 \text{Appr}_{t-1} \]
\[ + \beta_{10} \text{GDP}_1t + \beta_{11} \text{GDP}_{1t-1} + \beta_{20} \text{Income}_2t + \beta_{21} \text{Income}_{2t-1} \]
\[ + \beta_{30} \text{Inflation}_3t + \beta_{31} \text{Inflation}_{3t-1} + \beta_{40} \text{Unemp}_4t + \beta_{41} \text{Unemp}_{4t-1} + \varepsilon_t. \]
ADL Modeling

Alternative models might change the substantive inference.
ADL Modeling

- Alternative models might change the substantive inference.

- The general-to-specific modeling strategy (De Boef and Keele 2008): Start with estimating ADL(8,8,4), for example, and search for the best structure using stepwise searching process such as comparing AIC.
  - ADL(8,8,4); ADL(7,7,4); ADL(6,6,4) ...
  - However,
ADL Modeling

- Alternative models might change the substantive inference.

- The general-to-specific modeling strategy (De Boef and Keele 2008): Start with estimating ADL(8,8,4), for example, and search for the best structure using stepwise searching process such as comparing AIC.
  - ADL(8,8,4); ADL(7,7,4); ADL(6,6,4) ...
  - However,

\[
\text{Appr}_t = \alpha_0 + \alpha_1 \text{Appr}_{t-1} + \alpha_1 \text{Appr}_{t-2} + \alpha_1 \text{Appr}_{t-3} \\
+ \beta_{10} \text{GDP}_{1t} + \beta_{11} \text{GDP}_{1t-1} + \beta_{12} \text{GDP}_{1t-2} \\
+ \beta_{20} \text{Income}_{2t} + \beta_{21} \text{Income}_{2t-1} + \beta_{22} \text{Income}_{2t-2} + \beta_{23} \text{Income}_{2t-3} \\
+ \beta_{30} \text{Inflation}_{3t} + \beta_{31} \text{Inflation}_{3t-1} \\
+ \beta_{40} \text{Unemp}_{4t} + \beta_{41} \text{Unemp}_{4t-1} + \varepsilon_t.
\]
ADL Modeling

- Alternative models might change the substantive inference.
- The general-to-specific modeling strategy (De Boef and Keele 2008): Start with estimating ADL(8,8,4), for example, and search for the best structure using stepwise searching process such as comparing AIC.
  - ADL(8,8,4); ADL(7,7,4); ADL(6,6,4) ...
  - However,

\[
\text{Appr}_t = \alpha_0 + \alpha_1 \text{Appr}_{t-1} + \alpha_1 \text{Appr}_{t-2} + \alpha_1 \text{Appr}_{t-3} + \beta_{10} \text{GDP}_{1t} + \beta_{11} \text{GDP}_{1t-1} + \beta_{12} \text{GDP}_{1t-2} + \beta_{20} \text{Income}_{2t} + \beta_{21} \text{Income}_{2t-1} + \beta_{22} \text{Income}_{2t-2} + \beta_{23} \text{Income}_{2t-3} + \beta_{30} \text{Inflation}_{3t} + \beta_{31} \text{Inflation}_{3t-1} + \beta_{40} \text{Unemp}_{4t} + \beta_{41} \text{Unemp}_{4t-1} + \varepsilon_t.
\]

Vast potential model space!
Then, why not the general model ADL(8,8,4)?
ADL Modeling

- Then, why not the general model ADL(8,8,4)?
- Overfitting is a real danger. ⇒ Spurious inference.
ADL Modeling

- Then, why not the general model ADL(8,8,4)?
- Overfitting is a real danger. ⇒ Spurious inference.

- A solution: use a machine learning-based penalization method to address the problem of overfitting instead of the stepwise search process.
- The Bayesian adaptive lasso algorithm for the ADL (the ADLBL method) penalizes “unimportant” lagged variables and mitigates overfitting.
ADL Modeling

\[ ADL(1, 0, 4) : \]

\[ Appr_t = \alpha_0 + \alpha_1 Appr_{t-1} \]
\[ + \beta_1 GDP_{1t} + \beta_2 Income_{2t} + \beta_3 Inflation_{3t} + \beta_4 Unemp_{4t} + \varepsilon_t. \]

In R, use the `lm` function. But before interpret the results, check for residual autocorrelation by conducting the Breusch-Godfrey Lagrange Multiplier (LM) test.

```r
library(lmtest)
bgtest(adl104, order = 1)
Breusch-Godfrey test for serial correlation of order up to 1
data: adl104
LM test = 0.43101, df = 1, p-value = 0.5115
```
ADL Modeling

ADL(1, 0, 4):

\[ \text{Appr}_t = \alpha_0 + \alpha_1 \text{Appr}_{t-1} + \beta_1 \text{GDP}_t + \beta_2 \text{Income}_t + \beta_3 \text{Inflation}_t + \beta_4 \text{Unemp}_t + \varepsilon_t. \]

In R, use the `lm` function. But before interpret the results, check for residual autocorrelation by conducting the Breusch-Godfrey Lagrange Multiplier (LM) test.

```r
library(lmtest)
bgtest(adl104, order = 1)
Breusch-Godfrey test for serial correlation of order up to 1
data:  adl104
LM test = 0.43101, df = 1, p-value = 0.5115
```
ADL Modeling

\[ ADL(1, 1, 4) : \]

\[ Appr_t = \alpha_0 + \alpha_1 Appr_{t-1} + \beta_{10} GDP_{1t} + \beta_{11} GDP_{1t-1} + \beta_{20} Income_{2t} + \beta_{21} Income_{2t-1} + \beta_{30} Inflation_{3t} + \beta_{31} Inflation_{3t-1} + \beta_{40} Unemp_{4t} + \beta_{41} Unemp_{4t-1} \]

library(lmtest)
bgtest(adl114, order = 1)
Breusch–Godfrey test for serial correlation of order up to 1

data: adl114
LM test = 0.15958, df = 1, p-value = 0.6895
Interpretations: Short-Run and Long-Run Effects

ADL(1, 1, 4) :

\[
\text{Appr}_t = \alpha_0 + \alpha_1 \text{Appr}_{t-1} \\
+ \beta_{10} \text{GDP}_t + \beta_{11} \text{GDP}_{t-1} + \beta_{20} \text{Income}_{2t} + \beta_{21} \text{Income}_{2t-1} \\
+ \beta_{30} \text{Inflation}_{3t} + \beta_{31} \text{Inflation}_{3t-1} + \beta_{40} \text{Unemp}_{4t} + \beta_{41} \text{Unemp}_{4t-1} + \epsilon_t
\]

| INCPCHT0   | -3.241e-02 | 4.413e-02 | -0.735  | 0.463603 |
| INCPCHT1   | 8.615e-02  | 4.324e-02 | 1.992   | 0.047888 |

\( \beta_{20} = -0.03 \) \( \beta_{21} = 0.08 \).

The short-run effect: an effect of \( X_t \) on \( Y_t \) that occurs immediately.
The long-run effect, or long-run multiplier: the total causal effect of $X_t$ on $Y_t$. Given the general model ADL($p, q, k$),

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{k} \sum_{i=0}^{q} \beta_{ji} X_{jt-i} + \varepsilon_t,$$

$$LRE_{ADL} = \frac{\sum_{i=0}^{q} \beta_i}{1 - \sum_{i=1}^{p} \alpha_i}.$$
Interpretations: Short-Run and Long-Run Effects

\[ ADL(1, 1, 4) : \]
\[ Appr_t = \alpha_0 + \alpha_1 Appr_{t-1} \]
\[ + \beta_{10} GDP_{1t} + \beta_{11} GDP_{1t-1} + \beta_{20} Income_{2t} + \beta_{21} Income_{2t-1} \]
\[ + \beta_{30} Inflation_{3t} + \beta_{31} Inflation_{3t-1} + \beta_{40} Unemp_{4t} + \beta_{41} Unemp_{4t-1} + \epsilon_t \]

\[ LRE_{ADL} = \frac{\beta_{20} + \beta_{21}}{1 - \alpha_1} \]

YT1 \quad 7.455e-01 \quad 4.846e-02 \quad 15.386 \quad < 2e-16

INCPCHT0 \quad -3.241e-02 \quad 4.413e-02 \quad -0.735 \quad 0.463603

INCPCHT1 \quad 8.615e-02 \quad 4.324e-02 \quad 1.992 \quad 0.047888

\[ LRE_{ADL} = \frac{-0.03 + 0.08}{1 - 0.75} = 0.2 \]
Interpretations: Short-Run and Long-Run Effects

\( ADL(1, 1, 4) : \)

\[
Appr_t = \alpha_0 + \alpha_1 Appr_{t-1} \\
+ \beta_{10} GDP_{1t} + \beta_{11} GDP_{1t-1} + \beta_{20} Income_{2t} + \beta_{21} Income_{2t-1} \\
+ \beta_{30} Inflation_{3t} + \beta_{31} Inflation_{3t-1} + \beta_{40} Unemp_{4t} + \beta_{41} Unemp_{4t-1} + \varepsilon_t
\]

\[
LRE_{ADL} = \frac{\beta_{20} + \beta_{21}}{1 - \alpha_1}
\]

\begin{align*}
YT1 & \quad 7.455e-01 & 4.846e-02 & 15.386 & < 2e-16 \\
INCPCHT0 & \quad -3.241e-02 & 4.413e-02 & -0.735 & 0.463603 \\
INCPCHT1 & \quad 8.615e-02 & 4.324e-02 & 1.992 & 0.047888
\end{align*}

\[
LRE_{ADL} = \frac{-0.03 + 0.08}{1 - 0.75} = 0.2
\]
Interpretations: Short-Run and Long-Run Effects

\[ \text{LRE}_{ADL} = \frac{-0.03 + 0.08}{1 - 0.75} = 0.2 \]

Standard error?: The variance for the LRE is given by the formula for the approximation of the variance of a ratio of coefficients with known variances.

\[ \text{Var}(a/b) = \frac{1}{b^2} \text{Var}(a) + \frac{a^2}{b^4} \text{Var}(b) - 2\frac{a}{b^3} \text{Cov}(a, b) \]
Interpretations: Short-Run and Long-Run Effects

\[ \text{LRE}_{ADL} = \frac{-0.03 + 0.08}{1 - 0.75} = 0.2 \]

Standard error?: The variance for the LRE is given by the formula for the approximation of the variance of a ratio of coefficients with known variances.

\[ \text{Var}(a/b) = \left( \frac{1}{b^2} \right) \text{Var}(a) + \left( \frac{a^2}{b^4} \right) \text{Var}(b) - 2\left( \frac{a}{b^3} \right) \text{Cov}(a, b) \]

\[ \text{Var} \left( \frac{\beta_0 + \beta_1}{1 - \alpha} \right) = \left( \frac{1}{(\beta_0 + \beta_1)^2} \right) \text{Var}(1 - \alpha) + \frac{(1 - \alpha)^2}{(\beta_0 + \beta_1)^4} \text{Var}(\beta_0 + \beta_1) \]

\[ - 2 \left( \frac{1 - \alpha}{(\beta_0 + \beta_1)^3} \right) \text{Cov}(1 - \alpha, \beta_0 + \beta_1) \]
Interpretations: Short-Run and Long-Run Effects

\[
\text{Var} \left( \frac{\beta_0 + \beta_1}{1 - \alpha} \right) = \left( \frac{1}{(\beta_0 + \beta_1)^2} \right) \text{Var}(1 - \alpha) + \frac{(1 - \alpha)^2}{(\beta_0 + \beta_1)^4} \text{Var}(\beta_0 + \beta_1) - 2 \left( \frac{1 - \alpha}{(\beta_0 + \beta_1)^3} \right) \text{Cov}(1 - \alpha, \beta_0 + \beta_1)
\]

Well... might be doable.
Interpretations: Short-Run and Long-Run Effects

\[
\text{Var} \left( \frac{\beta_0 + \beta_1}{1 - \alpha} \right) = \left( \frac{1}{(\beta_0 + \beta_1)^2} \right) \text{Var}(1 - \alpha) + \frac{(1 - \alpha)^2}{(\beta_0 + \beta_1)^4} \text{Var}(\beta_0 + \beta_1) \\
- 2 \left( \frac{1 - \alpha}{(\beta_0 + \beta_1)^3} \right) \text{Cov}(1 - \alpha, \beta_0 + \beta_1)
\]

Well... might be doable.

But what if ADL(2,2,4)?

\[
\text{LRE}_{\text{ADL}} = \frac{\beta_0 + \beta_1 + \beta_2}{1 - \alpha_1 - \alpha_2}
\]
Interpretations: Short-Run and Long-Run Effects

\[
\text{Var} \left( \frac{\beta_0 + \beta_1}{1 - \alpha} \right) = \left( \frac{1}{(\beta_0 + \beta_1)^2} \right) \text{Var}(1 - \alpha) + \frac{(1 - \alpha)^2}{(\beta_0 + \beta_1)^4} \text{Var}(\beta_0 + \beta_1) \\
- 2 \left( \frac{1 - \alpha}{(\beta_0 + \beta_1)^3} \right) \text{Cov}(1 - \alpha, \beta_0 + \beta_1)
\]

Well... might be doable.

But what if ADL(2,2,4)?

\[
\text{LRE}_{ADL} = \frac{\beta_0 + \beta_1 + \beta_2}{1 - \alpha_1 - \alpha_2}
\]

ADL(3,3,4)?

\[
\text{LRE}_{ADL} = \frac{\beta_0 + \beta_1 + \beta_2 + \beta_3}{1 - \alpha_1 - \alpha_2 - \alpha_3}
\]
Interpretations: Short-Run and Long-Run Effects

- An easy solution: Bayesian inference.
- `MCMCregress` in the `MCMCpack` package.

From the frequentist model,

|          | Estimate | Std. Error | t value | Pr(>|t|) |
|----------|----------|------------|---------|----------|
| YT1      | 7.455e-01| 4.846e-02  | 15.386  | < 2e-16  |
| INCPCHT0 | -3.241e-02| 4.413e-02  | -0.735  | 0.463603 |
| INCPCHT1 | 8.615e-02| 4.324e-02  | 1.992   | 0.047888 |

From the Bayesian model,

<table>
<thead>
<tr>
<th></th>
<th>2.5%</th>
<th>5%</th>
<th>50%</th>
<th>95%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>YT1</td>
<td>0.651086</td>
<td>-0.119684</td>
<td>0.000962</td>
<td>0.040652</td>
<td>0.054924</td>
</tr>
<tr>
<td>INCPCHT0</td>
<td>-0.105577</td>
<td>-0.105577</td>
<td>-0.032248</td>
<td>0.040652</td>
<td>0.054924</td>
</tr>
<tr>
<td>INCPCHT1</td>
<td>0.014812</td>
<td>0.014812</td>
<td>0.086471</td>
<td>0.157854</td>
<td>0.172117</td>
</tr>
</tbody>
</table>
Interpretations: Short-Run and Long-Run Effects

- An easy solution: Bayesian inference.
- MCMCregress in the MCMCpack package.

From the frequentist model,

\[
\begin{align*}
YT1 & \quad 7.455e-01 & \quad 4.846e-02 & \quad 15.386 & \quad < \quad 2e-16 \\
INCPCHT0 & \quad -3.241e-02 & \quad 4.413e-02 & \quad -0.735 & \quad 0.463603 \\
INCPCHT1 & \quad 8.615e-02 & \quad 4.324e-02 & \quad 1.992 & \quad 0.047888
\end{align*}
\]

From the Bayesian model,

<table>
<thead>
<tr>
<th></th>
<th>YT1</th>
<th>INCPCHT0</th>
<th>INCPCHT1</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5%</td>
<td>0.6510860</td>
<td>-0.11968481</td>
<td>0.0009618294</td>
</tr>
<tr>
<td>5%</td>
<td>0.6663087</td>
<td>-0.10557719</td>
<td>0.0148117736</td>
</tr>
<tr>
<td>50%</td>
<td>0.7458637</td>
<td>-0.03224817</td>
<td>0.0864708230</td>
</tr>
<tr>
<td>95%</td>
<td>0.8257843</td>
<td>0.04065239</td>
<td>0.1578539575</td>
</tr>
<tr>
<td>97.5%</td>
<td>0.8413791</td>
<td>0.05492385</td>
<td>0.1721170680</td>
</tr>
</tbody>
</table>
Interpretations: Short-Run and Long-Run Effects

\[ LRE_{ADL} = \frac{\beta_0 + \beta_1}{1 - \alpha} \]

- The MCMC output of the model: 50,000 values for each of the coefficient sampled from the posterior distributions.

- For each iteration, you can calculate \( \frac{\beta_0 + \beta_1}{1 - \alpha} \). \( \Rightarrow \) 50,000 values.

- Consider these values the posterior distribution of \( \frac{\beta_0 + \beta_1}{1 - \alpha} \).
Interpretations: Short-Run and Long-Run Effects

\[ \text{LRE}_{\text{ADL}} = \frac{\beta_0 + \beta_1}{1 - \alpha} \]

- The MCMC output of the model: 50,000 values for each of the coefficient sampled from the posterior distributions.

- For each iteration, you can calculate \( \frac{\beta_0 + \beta_1}{1 - \alpha} \). \( \Rightarrow \) 50,000 values.

- Consider these values the posterior distribution of \( \frac{\beta_0 + \beta_1}{1 - \alpha} \).

- Summary of the posterior distribution:

<table>
<thead>
<tr>
<th>2.5%</th>
<th>5%</th>
<th>10%</th>
<th>50%</th>
<th>90%</th>
<th>95%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.36</td>
<td>-0.26</td>
<td>-0.15</td>
<td>0.21</td>
<td>0.57</td>
<td>0.68</td>
<td>0.79</td>
</tr>
</tbody>
</table>
Interpretations: Short-Run and Long-Run Effects

The effect of income growth on presidential approval:

- In terms of the short-run effect, income growth lagged by one quarter matters for presidential approval.
- No contemporaneous effect.

- No statistically reliable long-run effect. (Note: alternative model specification may change the result.)
Multivariate Models (Dynamic Regression Models)

Two workhorse models most frequently used in the social sciences:

- Autoregressive Distributed Lag Model
- Error Correction Model
Multivariate Models (Dynamic Regression Models)

- Two workhorse models most frequently used in the social sciences:
  - Autoregressive Distributed Lag Model
  - Error Correction Model
The error correction model or ECM

- Equivalent to the ADL.
- Appropriate for use with stationary data and cointegrated time series. (Note: There is a counter-argument that the ECM should generally not be used with political data. See the debate between Grant and Lebo (2016) and Keele, Linn, and Webb (2016) appeared in the time series symposium at *Political Analysis* 24(1)).
Error Correction Model

Equivalent to the ADL? Consider ADL(1,1,1):

\[ Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \]

\[ Y_t - Y_{t-1} = \alpha_0 + (\alpha_1 - 1) Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \]

\[ \Delta Y_t = \alpha_0 + (\alpha_1 - 1) Y_{t-1} + \beta_0 \Delta X_t + \beta_1 X_{t-1} + \varepsilon_t \]

\[ \Delta Y_t = \alpha_0 + (\alpha_1 - 1) Y_{t-1} + \beta_0 \Delta X_t + (\beta_1 + \beta_0) X_{t-1} + \varepsilon_t \]

\[ \Delta Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 \Delta X_t + \beta_1 X_{t-1} + \varepsilon_t \]
Error Correction Model

Equivalent to the ADL? Consider ADL(1,1,1):

\[ Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \]

\[ Y_t - Y_{t-1} = \alpha_0 + (\alpha_1 - 1) Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \]

\[ \Delta Y_t = \alpha_0 + (\alpha_1 - 1) Y_{t-1} + \beta_0 X_t - \beta_0 X_{t-1} + \beta_1 X_{t-1} + \varepsilon_t \]

\[ \Delta Y_t = \alpha_0 + (\alpha_1 - 1) Y_{t-1} + \beta_0 \Delta X_t + (\beta_1 + \beta_0) X_{t-1} + \varepsilon_t \]

\[ \Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_0^* \Delta X_t + \beta_1^* X_{t-1} + \varepsilon_t \]

General ECM\((p, q, k)\):

\[ \Delta Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{k} \beta_j \Delta X_{jt} + \sum_{j=1}^{k} \sum_{i=1}^{q} \gamma_{ji} X_{jt-i} + \varepsilon_t. \]
The ECM is consistently estimated by ordinary least squares (OLS) for stationary or cointegrated data if white noise error and weakly exogenous regressors are assumed.

\[ \varepsilon_t \overset{iid}{\sim} N(0, \sigma^2), \]
\[ E(\varepsilon_t, X_{js}) = 0, \forall t, s, j. \]
The ECM is consistently estimated by ordinary least squares (OLS) for stationary or cointegrated data if white noise error and weakly exogenous regressors are assumed.

- \( \varepsilon_t \overset{iid}{\sim} N(0, \sigma^2) \),
- \( E(\varepsilon_t, X_{js}) = 0, \forall t, s, j \).

Cointegrated data?: Individually integrated and jointly cointegrated.

Two or more series are cointegrated if each component series is integrated and some linear combination of these series is stationary.
Example

- Monthly events data from Brandt, Colaresi, and Freeman ("The Dynamic of Reciprocity, Accountability, and Credibility," JCR 2008.)
- The foreign policy behaviors of the U.S., the Israelis, and the Palestinians.
- Measured based on Agence France Presse (AFP) and Reuters news stories.
- Monthly survey for Jewish people's support for peace.

\[ A2P = f(I2P, P2I, JPI) \]
Example

1. Check for the stationarity of the individual variables.

library(tseries)
adf.test(A2P)

Augmented Dickey-Fuller Test
data: A2P
Dickey-Fuller = -3.244, Lag order = 4, p-value = 0.08413
alternative hypothesis: stationary

adf.test(I2P)

Augmented Dickey-Fuller Test
data: I2P
Dickey-Fuller = -2.4997, Lag order = 4, p-value = 0.3693
alternative hypothesis: stationary
Example

- 1. Check for the stationarity of the individual variables.

```r
library(tseries)
adf.test(P2I)

Augmented Dickey-Fuller Test
data: P2I
Dickey-Fuller = -2.331, Lag order = 4, p-value = 0.4393
alternative hypothesis: stationary

adf.test(JPI)

Augmented Dickey-Fuller Test
data: JPI
Dickey-Fuller = -2.5401, Lag order = 4, p-value = 0.3525
alternative hypothesis: stationary
```
Example

1. Check for the stationarity of the individual variables.
   - None of the variables rejected the null hypothesis of a unit root. ⇒ May be individually integrated.
   - First differences of these variables \((Y_t - T_{t-1})\) might be stationary.
   - However, this type of transformation removes any long-run relationships in the data.
Example

1. Check for the stationarity of the individual variables.
   - None of the variables rejected the null hypothesis of a unit root. ⇒ May be individually integrated.
   - First differences of these variables \((Y_t - T_{t-1})\) might be stationary.
   - However, this type of transformation removes any long-run relationships in the data.
   - Cointegration methods ⇒ describe stationary equilibrium relationships, preserving long-run information.
     * The error correction model.
Example

1. Check for the stationarity of the individual variables.
   ▶ None of the variables rejected the null hypothesis of a unit root. ⇒
     May be individually integrated.
   ▶ First differences of these variables ($Y_t - T_{t-1}$) might be stationary.
   ▶ However, this type of transformation removes any long-run relationships in the data.
   ▶ Cointegration methods ⇒ describe stationary equilibrium relationships, preserving long-run information.
     ★ The error correction model.

2. Check for cointegration.
Example

2. Check for cointegration.

\[ A2P_t = \beta_0 + \beta_1 I2P_t + u_t. \]

\[ u_t = A2P_t - \beta_0 - \beta_1 I2P_t. \]

- Suppose we now subject \( u_t \) to unit root analysis and find that it is stationary.
- \( A2P_t \) and \( I2P_t \) are individually integrated, but their linear combination is stationary.
2. Check for cointegration.

\[ A2P_t = \beta_0 + \beta_1 I2P_t + u_t. \]
\[ u_t = A2P_t - \beta_0 - \beta_1 I2P_t. \]

- Suppose we now subject \( u_t \) to unit root analysis and find that it is stationary.
- \( A2P_t \) and \( I2P_t \) are individually integrated, but their linear combination is stationary.
- Then, \( A2P_t \) and \( I2P_t \) are cointegrated.
2. Check for cointegration: Find out if the regression residuals are stationary.

```r
# Engle-Granger two steps
library(urca)
step1 = lm(A2PT0 ~ I2PT0 + P2IT0 + JPIT0, data=data)
etaT0 = step1$residuals
step2 = ur.df(etaT0, type="none", selectlags = "BIC")
```
Example

Coefficients:

|                | Estimate | Std. Error | t value | Pr(>|t|) |
|----------------|----------|------------|---------|----------|
| z.lag.1        | -1.1161  | 0.1381     | -8.08   | 1.17e-12 |
| z.diff.lag     | 0.1049   | 0.0972     | 1.08    | 0.283    |

---

Value of test-statistic is: -8.0803

Critical values for test statistics:

<table>
<thead>
<tr>
<th>tau1</th>
<th>1pct</th>
<th>5pct</th>
<th>10pct</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.58</td>
<td>-1.95</td>
<td>-1.62</td>
<td></td>
</tr>
</tbody>
</table>
Example

1. Check for stationarity.

2. Check for cointegration. \( \Rightarrow \) ECM is appropriate.

3. Exogenous regressors?

Example

General ECM\((p, q, k)\):

\[
\Delta Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{k} \beta_j \Delta X_{jt} + \sum_{j=1}^{k} \sum_{i=1}^{q} \gamma_{ji} X_{jt-i} + \varepsilon_t.
\]

The same issue with specifying \(p, q\) arises as in the ADL modeling. The ECMBL method (Estimating ECM via the Bayesian adaptive lasso) is a solution.
Example

General ECM($p$, $q$, $k$):

$$\Delta Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{k} \beta_j \Delta X_{jt} + \sum_{j=1}^{k} \sum_{i=1}^{q} \gamma_{ji} X_{jt-i} + \varepsilon_t.$$  

The same issue with specifying $p$, $q$ arises as in the ADL modeling. The ECMBL method (Estimating ECM via the Bayesian adaptive lasso) is a solution.

ECM(1,1,3)

$$\Delta A2P_t = \alpha_0 + \alpha_1 A2P_{t-1} + \beta_{10} \Delta I2P_{1t} + \beta_{11} I2P_{1t-1} + \beta_{20} \Delta P2l_{2t} + \beta_{21} P2l_{2t-1} + \beta_{30} \Delta JPI_{3t} + \beta_{31} JPI_{3t-1} + \varepsilon_t$$
5. Check for residual autocorrelation by conducting the Breusch-Godfrey Lagrange Multiplier (LM) test.
   ▶ bgtest in R.

\[
\Delta A_2P_t = \alpha_0 + \alpha_1 A_2P_{t-1} + \beta_{10} \Delta I_2P_1t + \beta_{11} I_2P_1t_{t-1} + \beta_{20} \Delta P_2I_2t + \beta_{21} P_2I_2t_{t-1} + \beta_{30} \Delta JPI_3t + \beta_{31} JPI_3t_{t-1} + \epsilon_t
\]
5. Check for residual autocorrelation by conducting the Breusch-Godfrey Lagrange Multiplier (LM) test.

▶ bgtest in R.

$\text{ECM}(1,1,3)$

\[
\Delta A2P_t = \alpha_0 + \alpha_1 A2P_{t-1} + \beta_{10}\Delta I2P_{1t} + \beta_{11} I2P_{1t-1} + \beta_{20}\Delta P2I_{2t} + \beta_{21} P2I_{2t-1} + \beta_{30}\Delta JPI_{3t} + \beta_{31} JPI_{3t-1} + \varepsilon_t
\]

Estimated short-run effect:

<table>
<thead>
<tr>
<th></th>
<th>A2PT1</th>
<th>I2PT1</th>
<th>I2PCHT0</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5%</td>
<td>-1.245</td>
<td>0.166</td>
<td>0.041</td>
</tr>
<tr>
<td>50%</td>
<td>-1.037</td>
<td>0.491</td>
<td>0.238</td>
</tr>
<tr>
<td>97.5%</td>
<td>-0.827</td>
<td>0.816</td>
<td>0.434</td>
</tr>
</tbody>
</table>
Interpretations: Short-Run and Long-Run Effects

\[
\Delta Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{k} \beta_j \Delta X_{jt} + \sum_{j=1}^{k} \sum_{i=1}^{q} \gamma_{ji} X_{jt-i} + \varepsilon_t,
\]

\[
LRE_{ECM} = -\frac{\sum_{i=0}^{q} \gamma_i}{\sum_{i=1}^{p} \alpha_i}.
\]

- The MCMC output of the model: 50,000 values for each of the coefficient sampled from the posterior distributions.
- For each iteration, you can calculate LRE. ⇒ 50,000 values.
- Consider these values the posterior distribution of LRE.
- Summary of the posterior distribution:

<table>
<thead>
<tr>
<th></th>
<th>2.5%</th>
<th>5%</th>
<th>50%</th>
<th>95%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.17</td>
<td>0.22</td>
<td>0.47</td>
<td>0.72</td>
<td>0.78</td>
</tr>
</tbody>
</table>
Some Advanced Topics

Vector Autoregression (VAR)

\[A2P_t = \alpha_1 + \sum_{j=1}^{k} \beta_{1j} A2P_{t-j} + \sum \gamma_{1j} I2P_{t-j} + \sum \delta_{1j} P2I_{t-j} + u_{1t}\]

\[I2P_t = \alpha_2 + \sum_{j=1}^{k} \beta_{2j} A2P_{t-j} + \sum \gamma_{2j} I2P_{t-j} + \sum \delta_{2j} P2I_{t-j} + u_{2t}\]

\[P2I_t = \alpha_3 + \sum_{j=1}^{k} \beta_{3j} A2P_{t-j} + \sum \gamma_{3j} I2P_{t-j} + \sum \delta_{3j} P2I_{t-j} + u_{3t}\]
Some Advanced Topics

Fractional Integration: Somewhere between stationary and a unit root (integrated).

\[ Y_t = \rho Y_{t-1} + u_t, \quad |\rho| = 1 - c \]

- \( c = 0 \Rightarrow \text{Integrated.} \)
- \( c = 1 \Rightarrow \text{Stationary.} \)
- Between \( \Rightarrow \) Fractionally integrated. (\( c \) is very small \( \Rightarrow \) Near-integrated).
- Asymptotically stationary but behave as integrated series in finite samples.

- One unit root test rejects the null of a unit root; but another does not.
- Many dynamic political data are fractionally integrated.
- Estimating the degree of integration is controversial.