Lecture II.- Time Series with Stationary Variables:

ARDL models

Marcelo Villena, PhD
Santa María University

September 12, 2022

Outline

Example from the previous class... Detrending global temperature

- As we saw in the previous class, the evolution of the global temperature showed a linear trend, so we can assume that it can be written as:

$$x_t = \mu_t + y_t$$

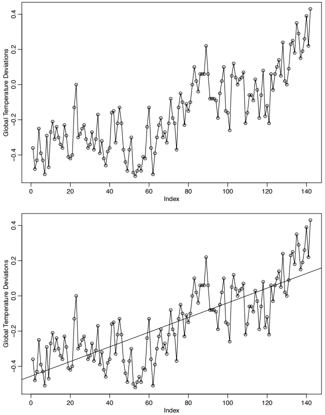
- We will see two ways of decomposing the series, "filtering" the trend.

Example from the previous class... Detrending global temperature

R Code

```
rm(list=ls())
mydata< −read.csv (“gtemp.csv”)
gtemp< −mydata$“gtem”
plot(gtemp, type=“o”, ylab= “Global Temperature Deviations”)
t< −1:142
summary(reg < − lm(gtemp˜ t))
plot(gtemp, type=“o”, ylab=“Global Temperature Deviations”)
abline(reg)
```

Example from the previous class... Detrending global temperature



Example from the previous class... Detrending global temperature

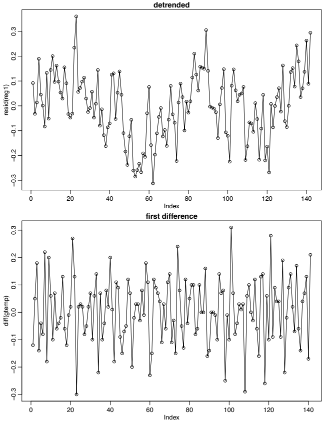
```
Call:
lm(formula = gtemp ~ t)

Residuals:
     Min       1Q   Median       3Q      Max
-0.31231 -0.08627  0.00681  0.09064  0.36023

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.4560863  0.0227675  -20.03   <2e-16 ***
t            0.0041677  0.0002762   15.09   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1349 on 140 degrees of freedom
Multiple R-squared:  0.6192,    Adjusted R-squared:  0.6164
F-statistic: 227.6 on 1 and 140 DF,  p-value: < 2.2e-16
```

Example from the previous class... Detrending global temperature

R Code

```
reg1= lm(gtemp~ time(gtemp), na.action=NULL)
par(mfrow=c(2,1))
plot(resid(reg1), type="o", main="detrended")
plot(diff(gtemp), type="o", main="first difference")
```

Example from the previous class... Detrending global temperature

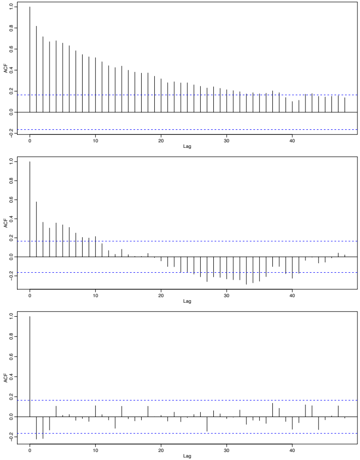


Example from the previous class... Detrending global temperature

R Code

```
par(mfrow=c(3,1))
acf(gtemp, 48, main="gtemp")
acf(resid(reg), 48, main="detrended")
acf(diff(gtemp), 48, main="first difference")
mean (diff (gtemp))
sd (diff (gtemp)) / sqrt (longitud (diff (gtemp)))
```

Example from the previous class... Detrending global temperature



On the decomposition of a series

- In the graphs we can appreciate that the first difference of the series produces different results than the trend removal by trend regression.
- In the case of the ACF graphs, the differenced process shows minimal autocorrelation, which may imply that the global temperature series is similar to a random walk with drift.
- It is interesting to note that this series could be seen as a random walk with drift.
- The mean of the differenced series, which is an estimate of the drift, is approximately ,0066, but with a large standard error.

On the decomposition of a series

- An advantage of differencing over the estimation of a trend, to eliminate trends, is that no parameters are estimated in the differencing operation. A disadvantage, however, is that differencing does not yield an estimate of the stationary process y_t.
- Thus, if an estimate of y_t is essential, then estimating a trend may be the most appropriate way to remove trends from the series. If the goal is to force the data to stationarity, then differencing may be more appropriate. Differencing is also a viable tool if the trend is fixed.
- In the U.S., the official decomposition and seasonal adjustment procedure is called ”seasonal adjustment. X-13-ARIMA
- **http://www.census.gov/srd/www/x13as/**

Non-stationary integrated processes and the unit root test

Recall that if a time series is stationary, its mean, variance and autocovariance (at different lags) remain the same regardless of the point in time at which they are measured, i.e. they are time invariant. On the other hand, we have seen that stationarity is a desirable characteristic, for example, in terms of the normality of the variables.

However, in practice we encounter:

1. **Non-stationary processes:** When a stochastic time series process is time-dependent.
2. **Integrated Processes:** a non-stationary process, which can be transformed to a stationary process by differentiating.

Integrated Processes

With respect to Integrated Processes, we start by defining:

- The sequence x_t is integrated of order d, $I(d)$, if it requires to be differentiated d times to become stationary.
- **All Integrated Processes are non-stationary, but not all non-stationary processes are integrated.**
- If the sequence x_t has a unit ration, then, it is an integrated process, and of that non-stationary.

Consequences of Integrated Processes (Unit Root)

- It is important to note that standard statistical tests are not appropriate when OLS is applied to integrated processes, see for example Granger 1974.
- If the sequence x_t is a unit ration process, then any shock has a permanent (non-decaying) effect. Hence, the time series is properly modeled by assuming a stochastic trend. The time series can then be defined as stationary differentiable, and the trend should be taken out by differentiating.
- In this context, **the terms non-stationarity, random walk, unit root and stochastic trend are considered synonymous.**

Is random walk nonstationarity?

$$y_t = y_{t-1} + \varepsilon_t$$

$$Var(y_t) = Var(y_{t-1} + \varepsilon_t)$$

$$Var(y_t) = Var(y_{t-1}) + \sigma_\varepsilon^2$$

$$Var(y_t) = Var(y_{t-2} + \varepsilon_{t-1}) + \sigma_\varepsilon^2$$

$$Var(y_t) = Var(y_{t-2}) + 2\sigma_\varepsilon^2$$

repeating this for t steps:

$$Var(y_t) = Var(y_0) + t\sigma_\varepsilon^2$$

If we assume that $y0$ is given:

$$Var(yt) = t\sigma_\varepsilon^2$$

The variance of the process increases with time, and therefore RW os not stationary.

Test of Unit Root

- Consider the following autoregressive process:

$$x_t = \alpha_1 x_{t-1} + \varepsilon_t \qquad (1)$$

- If $\alpha_1 = 1$, the sequence x_t is a unit root.
- The standard test to prove this hypothesis is to subtract x_{t-1} from the above equation such that:

$$\triangle xt = \gamma xt - 1 + \varepsilon t \qquad (2)$$

- where $\gamma = \alpha_1 - 1$, and $\triangle x_t = x_t - x_{t-1}$. In this context, proving the hypothesis that equation (1) has a unit ration, $\alpha_1 = 1$, is equivalent to proving the hypothesis of $\gamma = 0$ in equation (2).

Test of Unit Root

- This is basically **the Dickey-Fuller (DF) approach for unit roots**, see e.g., Dickey Fuller 1981.
- Additionally there is the **Augmented Dickey-Fuller test (ADF)**, and many other tests that are based on similar logic, which we will use during the course.

Example of Unit Root Test - Dickey-Fuller

R Code

```
install.packages("tseries")
library(tseries)
adf.test(gtemp)
adf.test(resid(reg1))
adf.test(diff(gtemp))
```

Example of Unit Root Test - Dickey-Fuller

Augmented Dickey-Fuller Test
data: gtemp
Dickey-Fuller = -2.0624, Lag order = 5, p-value = 0.5505
alternative hypothesis: stationary

Augmented Dickey-Fuller Test
data: resid(reg1)
Dickey-Fuller = -2.0624, Lag order = 5, p-value = 0.5505
alternative hypothesis: stationary

Augmented Dickey-Fuller Test
data: diff(gtemp)
Dickey-Fuller = -6.8179, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary

Distributed lagged model

- In regression analysis with time series data, when the regression model includes not only current values but also lagged (past) values of the explanatory variables (the X's), it is called a distributed lagged model.
- If the model includes one or more lagged values of the dependent variable among its explanatory variables, it is called an autorregressive model.

Distributed lagged model

Thus,

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + u_t$$

represents a distributed lagged model, while

$$Y_t = \alpha + \beta X_t + \gamma Y_{t-1} + u_t$$

is an example of an autoregressive model. The latter are also known as dynamic models, since they indicate the trajectory over time of the dependent variable relative to its past value(s).

Distributed lagged model

- More generally, we would write

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_k X_{t-k} + u_t$$

- which is the distributed lags model with a finite lag of k periods. The coefficient β_0 is known as the short-run or impact multiplier because it gives the change in the mean value of Y that follows a unit change in X in the same period.1
- Technically, β_0 is the partial derivative of Y with respect to X_t, β_1 with respect to X_{t-1}, β_2 with respect to X_{t-2}, and so on. Symbolically, $\frac{\partial Y_t}{\partial X_{t-k}} = \beta_k$.

Distributed lagged model

- If the change in X remains the same from the beginning, then $(\beta_0 + \beta_1)$ gives the change in (the mean value of) Y in the next period $(\beta_0 + \beta_1 + \beta_2)$ in the one that follows, and so on. These partial sums are denoted as interim, or intermediate, multipliers. Finally, after k periods we obtain:

$$\sum \beta i = \beta_0 + \beta_1 + \beta_2 + ... + \beta_k = \beta$$

- which is known as the long-run or total distributed lag multiplier, provided that the sum β exists (we will explain this later). If we define

$$\beta_i^* = \frac{\beta i}{\sum \beta i} = \frac{\beta i}{\beta}$$

- we obtain "standardized" β_i. The partial sums of the standardized β_i give the proportion of the long-run, or total, impact felt during a certain period.

Distributed lagged model

- **Autoregressive and distributed lag models are very common in economic analysis**.
- We will study them in detail in order to find out the following:

1. What is the role of lags in economics?
2. On what grounds are lags justified?
3. Is there any theoretical justification for the lagged models common in empirical econometrics?

Distributed lagged model

4. What is the relationship, if any, between autoregressive models and distributed lag models, and can they be derived from each other?
5. What are some statistical problems related to the estimation of such models?
6. Does the lagged-ahead relationship between variables imply causality? If so, how is it measured?

On the the nature of lagged phenomena

1. **Psychological reasons.** As a result of force of habit (inertia), people do not change their consumption habits immediately after a price reduction or an increase in income, perhaps because the process of change entails some immediate disadvantage.
2. **Technological reasons.** Suppose that the price of capital relative to labor is reduced, so that it is economically feasible to substitute labor for capital. Of course, the addition of capital takes time (gestation period). Moreover, if the price fall is expected to be temporary, firms may not rush to substitute labor for capital, especially if they expect that after the temporary fall the price of capital may rise beyond its previous level. Sometimes, imperfect knowledge also explains lags.

On the the nature of lagged phenomena

3. **Institutional reasons.** These reasons also contribute to lags. For example, contractual obligations may prevent companies from switching from one source of labor or raw materials to another. For example, those who placed funds in long-term savings accounts with fixed terms, such as one, three or seven years, are "locked in," even though money market conditions now allow higher returns elsewhere.

Estimation of distributed lag models

- We already established that distributed lag models play a very useful role in economics, but how do we estimate such models?
- Suppose we have the following model of distributed lags for one explanatory variable:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + u_t$$

- where we have not defined the lag length, i.e., how far back in the past we wish to go. Such a model is called an infinite lag model, while a model of the type shown above is called a finite distributed lag model because the lag length k is specified.

Estimation of distributed lag models

How do we estimate α and the β's of this equation? We can adopt two approaches:

1. **ad hoc estimation** and
2. **a priori constraints on the** $\beta's$, if we assume that (the $\beta's$) follow a systematic pattern.

Ad hoc estimation of distributed lag models.

- Since the explanatory variable X_t is assumed to be nonstochastic (or at least uncorrelated with the disturbance term u_t), equally nonstochastic are X_{t-1}, X_{t-2}, and so on. Therefore, in principle, the ordinary least squares (OLS) method is applicable.
- This is the approach of **Alt** and **Tinbergen** who suggest that to estimate a distributed lag models, we proceed sequentially, i.e., first regress Y_t on X_t, then regress Y_t on X_t and X_{t-1}, then regress Y_t on X_t , X_{t-1}, and X_{t-2},and so on. This sequential process stops when the regression coefficients of the lagged variables start to become statistically insignificant and/or the coefficient of at least one variable changes its sign from positive to negative, or vice versa.

Ad hoc disadventages.

Although ad hoc estimation seems straightforward and unobtrusive, it has many disadvantages, including the following:

1. There is no a priori guidance on the maximum length the lag should be.
2. As successive lags are estimated, fewer degrees of freedom remain, thus weakening statistical inference somewhat.
3. More importantly, in economic time series data, successive (lagged) values tend to be highly correlated; thus, multi-linearity comes to the fore.

Koyck's method for distributed lag models

- Koyck proposed a different method of estimating distributed lag models. Suppose we start with an infinite distributed lags model. If all β have the same sign, Koyck assumes that they reduce geometrically as follows.

$$\beta_k = \beta_0 \lambda^k \qquad k = 0, 1, ...$$

- where λ, such that $0 < \lambda < 1$ is known as the rate of decline, or decay, of the distributed lag, and where $1 - \lambda$ is known as the rate of adjustment.
- What the model postulates is that each successive β coefficient is numerically lower than each previous β (this statement is due to the fact that $\lambda < 1$), implying that, as one returns to the distant past, the effect of that lag on Y_t becomes progressively smaller, a very reasonable assumption.

Koyck's method for distributed lag models

Note these features of Koyck's scheme:

1. by assuming non-negative values for λ, Koyck eliminates the possibility that the β will change sign;
2. by assuming that $\lambda < 1$, he gives less weight to β in the distant past than today; and
3. It ensures that the sum of the β, which provides the long-run multiplier.

The proof of Koyck's model

The proof of Koyck's model is fairly straightforward. Given the following distributed lags model:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + ... + u_t$$

Applying the transformation $\beta_k = \beta_0 \lambda^k$ we have:

$$Y_t = \alpha + \beta_0 X_t + \beta_0 \lambda^1 X_{t-1} + \beta_0 \lambda^2 X_{t-2} + ... + u_t$$

Then we lag the equation by one period:

$$Y_{t-1} = \alpha + \beta_0 X_{t-1} + \beta_0 \lambda^1 X_{t-2} + \beta_0 \lambda^2 X_{t-3} + ... + u_t$$

The proof of Koyck's model

Then we multiply by λ:

$$\lambda Y_{t-1} = \lambda\alpha + \lambda\beta_0 X_{t-1} + \lambda\beta_0\lambda^1 X_{t-2} + \lambda\beta_0\lambda^2 X_{t-3} + ... + u_t$$

By subtracting both equations we have:

$$Y_t - \lambda Y_{t-1} = \alpha(1-\lambda) + \beta_0 X_t + (u_t - \lambda u_{t-1})$$

By reordering we have:

$$Y_t = \alpha(1-\lambda) + \beta_0 X_t + \lambda Y_{t-1} + v_t$$

where $v_t = (u_t - \lambda u_{t-1})$ is a moving average of u_t and u_{t-1}.

On the problems of Koyck's model

- The presence of autocorrelation can lead to misleading results as they violate the assumptions of the Gauss Markov Theorem.
- However, in the presence of correlated errors, we can still proceed to fit a model makes up for these violations.
- Considering this, the least squares estimator is no longer unbiased, but it does have the desirable large sample property of consistency, and if the errors are normally distributed, it is best in a large sample sense.
- It is also important to note that is assumed to be uncorrelated random errors with zero mean and constant variance.
- **In such cases, the time series assumption that the error term is independent of current, past, and future values of is no longer valid.**

Least Squares Estimation

In the presence of serially correlated errors, the consequences of least squares estimation are similar to the consequences of ignoring the presence of heteroskedasticity, namely

1. The least squares estimator is still a linear unbiased estimator, but is no longer best.
2. The formulas for the standard errors usually computed for the least square estimator are no longer correct. Although the usual least squares standard errors are not the correct ones, it is possible to compute correct standard errors for the least squares estimator when the errors are serially correlated. These standard errors are known as HAC (heteroskedasticity and autocorrelation consistent) standard errors, or Newey-West standard errors, and they are analogous to heteroskedasticity consistent, or **White, standard errors**.

Adaptive expectations

- Suppose we postulate the following model:

$$Y_t = \beta_0 + \beta_1 X_t^* + u_t$$

- For example, let us assume that Y = demand for money (real cash balances) X^* = normal or expected long term or equilibrium interest rate, u optimal u = error term.
- Since the expectations variable X^* is not directly observable, we can propose the following hypothesis on how expectations are shaped:

$$X_t^* - X_{t-1}^* = \gamma(X_{t-1} - X_{t-1}^*)$$

- where γ, such that $0 < \gamma \leq 1$, is known as the expectations coefficient. This hypothesis is known as the adaptive expectations, progressive expectations or learning-by-error hypothesis, popularized by **Cagan and Friedman.**

Adaptive expectations

- Replacing:

$$Y_t = \beta_0 + \beta_1 \gamma X_t + \beta_1 (1-\gamma) X^*_{t-1} + u_t$$

- Now lagging our original equation one period, multiply it by $1-\gamma$, and subtracting the equation above, we obtain

$$Y_t = \gamma\beta_0 + \gamma\beta_1 X_t + (1-\gamma) Y_{t-1} + v_t$$

- where $v_t = u_t - (1-\gamma)u_{t-1}$
- Similar to the Koyck model!

$$Y_t = \alpha(1-\lambda) + \beta_0 X_t + \lambda Y_{t-1} + v_t$$

Adaptive expectations

- Until the rational expectations (RE) hypothesis, first put forward by J. Muth and later disseminated by Robert Lucas and Thomas Sargent, the adaptive expectations (AE) hypothesis was very popular in empirical economics.
- **Proponents of the RE hypothesis argue that the RE hypothesis is inadequate because the formulation of expectations is based only on past values of a variable, whereas the RE hypothesis assumes "that individual economic agents use currently available and relevant information in the formation of their expectations and do not rely solely on past experience...".**

Autoregressive Distributed Lag Models

- An autoregressive distributed lag model (ARDL) is a model that contains both independent variables and their lagged values as well as the lagged values of the dependent variable.
- In its more general form, with p lags of Y_t and q lags of X_t, an $ARDL(p,q)$ model can be written as:

$$Y_t = \delta + \theta_1 Y_{t-1} + ... + \theta_p Y_{t-p} + \delta_1 X_{t-1} + ... + \delta_q X_{t-q} + v_t$$

Autoregressive Distributed Lag Models

- The ARDL has several advantages.
  - It captures the dynamic effects from the lagged X's and the lagged $Y's$ by including a sufficient number of lags of Y and X,
  - We can eliminate serial correlation in the errors.
  - An ARDL model can be transfomed into one with only lagged X's,.

About Model Selection

- It may happen that several models describe the time series satisfactorily, making it necessary to select the most appropriate model.
- This selection process can be simple or a bit more complex, so it is necessary to use model selection criteria.
- The most common model selection criteria are the **AIC (Akaike Information Criterion)** and the **BIC (Bayesian Information Criterion)** which is a Bayesian extension of the first one.

Information Criteria

Definition

$$AIC = log\hat{\sigma}_k^2 + \frac{n+2k}{n}$$

where $\hat{\sigma}_k^2 = \frac{SSE_k}{n}$, and k is the number of model parameters, n the sample size, and SSE_k is equal to the sum of the squared residuals under the model k ($SSE_k = \sum_{t=1}^{n}(x_t - \bar{x})^2$).

The value of k that produces the minimum AIC represents the best model. The idea is that minimizing $\hat{\sigma}_k^2$ represents a reasonable objective, except that it decreases monotonically as k increases. Therefore, we should penalize the error variance by a term proportional to the number of parameters.

Information Criteria

Definitions

$$AICc = log\hat{\sigma}_k^2 + \frac{n+k}{n-k-2}$$

$$AICc = log\hat{\sigma}_k^2 + \frac{klogn}{n}$$

- BIC is also known as the **Schwarz Information Criterion (SIC)**. Several simulation studies have verified that BIC is adequate to obtain the correct order in large samples, while AICc tends to be superior in smaller samples where the relative number of parameters is large.

An Example: Okun's Law

- We will apply the finite distributed lag model to Okun's Law.
- Okun was an economist who posited that there was a relationship between the change in unemployment from one period to the next and the rate of growth of output in the economy.
- Mathetmatically, Okun's Law can be expressed as:

$$U_t - U_{t-1} = \gamma(G_t - G_N)$$

- where U_t is the unemployment rate in period t. G_t is the growth rate of output in period t, and G_N is the "normal" growth rate, which we assume constant over time.

An Example: Okun's Law

- We can rewrite the above equation in more familiar notation of the multiple regresion model by denoting the change in unemployment as: $\triangle U_t = U_{Ut} - U_{t-1}$. We then set $\gamma = \beta$ and $G_N = \alpha$. Including an error term to our equation yields:

$$\triangle U_t = \alpha + \beta_0 G_t + \mu_t$$

- Acknowledging the changes in output are likely to have a distributed-lag effect on unemployment - not all of the effect will take place instantenously. We can then further expand our equation to:

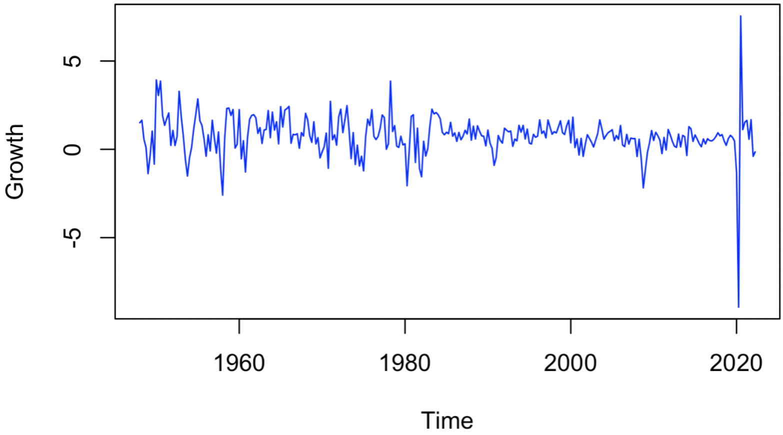
$$\triangle U_t = \alpha + \beta_0 G_t + \beta_1 G_{t-1} + \beta_2 G_{t-2} + ... + \beta_q G_{t-q} + \mu_t$$

An Example: Okun's Law

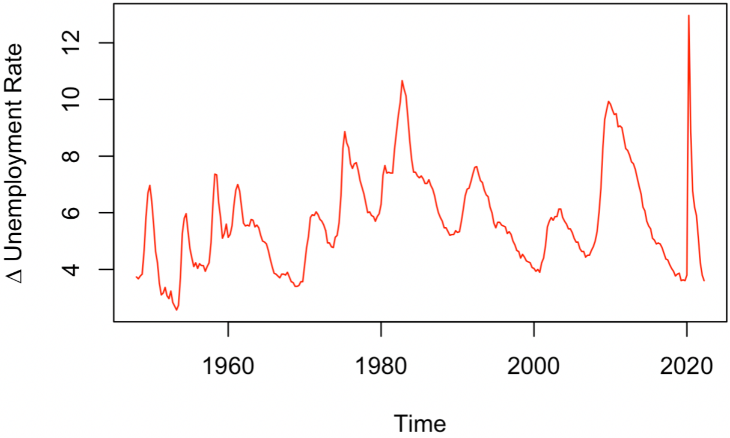
R Code

```
Okun<-read.csv("Okun.csv")
g <- ts(Okun$G, start=c(1948,1), frequency=4)
u <- ts(Okun$U, start=c(1948,1), frequency=4)
ts.plot(g, col='blue', ylab='Growth')
ts.plot(u, col='red', ylab='Δ Unemployment Rate')
```

An Example: Okun's Law



An Example: Okun's Law

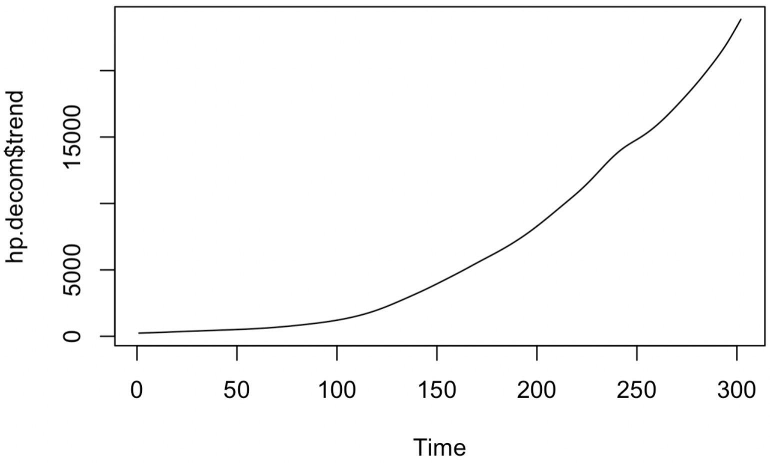


An Example: Okun's Law

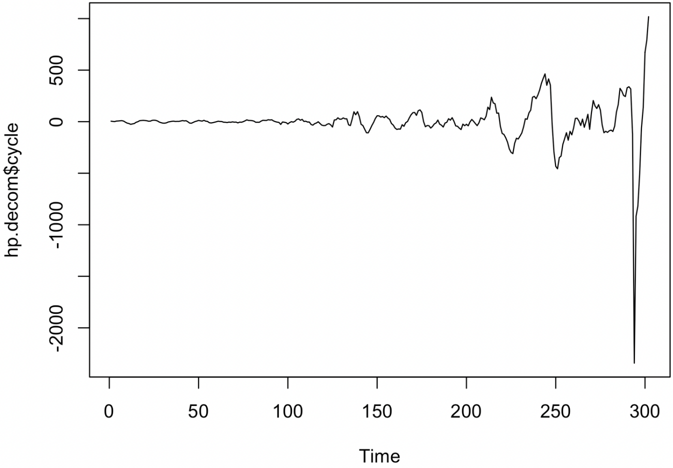
R Code

```
library(quantmod)
library(mFilter)
getSymbols('GDP',src='FRED') plot(GDP)
hp.decom <- hpfilter(GDP, freq = 1600, type = "lambda")
ts.plot(hp.decom$trend)
ts.plot(hp.decom$cycle)
g_Comp <- decompose(g) plot(g_Comp)
```

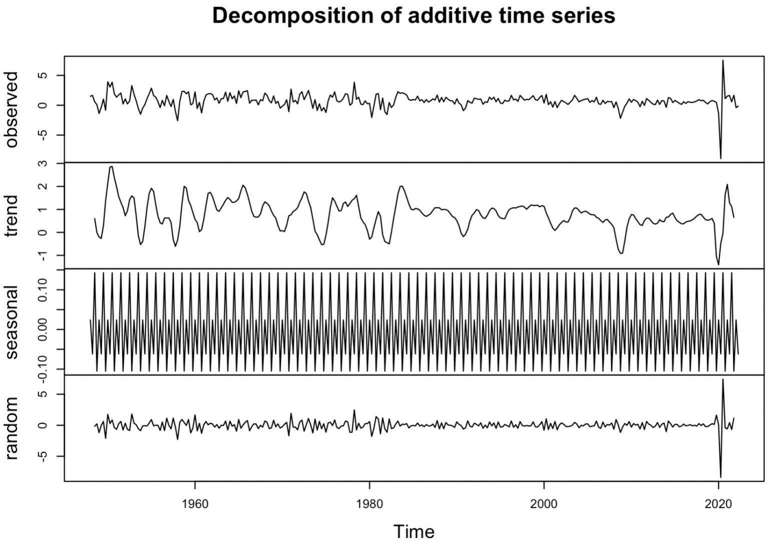
An Example: Okun's Law



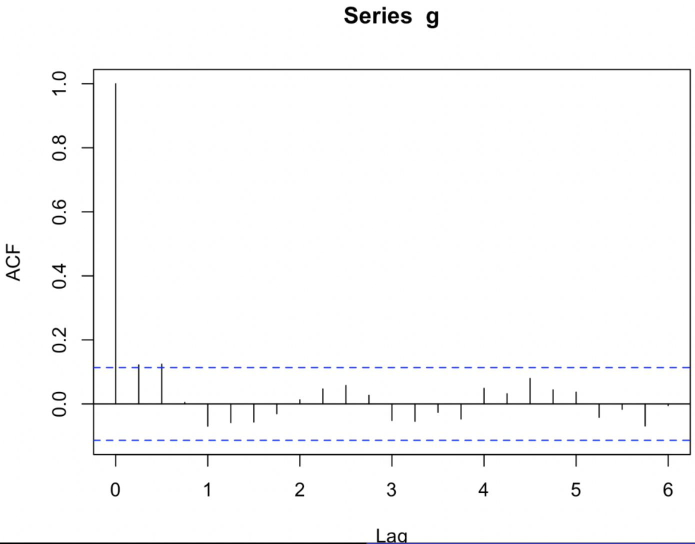
An Example: Okun's Law



An Example: Okun's Law



An Example: Okun's Law



An Example: Okun's Law

R Code

```
acf(g, type='correlation', plot=FALSE)$acf
acf(g, type='correlation')
okun.lag2 <- dynlm(d(u, 1) ~ L(g, 0:2))
okun.lag3 <- dynlm(d(u, 1) ~ L(g, 0:3))
summary(okun.lag2)
summary(okun.lag3)
```

An Example: Okun's Law

```
Time series regression with "ts" data:
Start = 1948(3), End = 2022(2)

Call:
dynlm(formula = d(u, 1) ~ L(g, 0:2))

Residuals:
    Min      1Q  Median      3Q     Max
-1.7462 -0.2072 -0.0252  0.1732  4.5096

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.40578    0.03715  10.923  < 2e-16 ***
L(g, 0:2)0  -0.46333    0.02261 -20.488  < 2e-16 ***
L(g, 0:2)1  -0.08049    0.02260  -3.561 0.000431 ***
L(g, 0:2)2   0.01173    0.02264   0.518 0.604656
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4454 on 292 degrees of freedom
Multiple R-squared:  0.6115,    Adjusted R-squared:  0.6075
F-statistic: 153.2 on 3 and 292 DF,  p-value: < 2.2e-16
```

An Example: Okun's Law

```
Time series regression with "ts" data:
Start = 1948(4), End = 2022(2)

Call:
dynlm(formula = d(u, 1) ~ L(g, 0:3))

Residuals:
    Min      1Q  Median      3Q     Max
-1.7277 -0.2011 -0.0081  0.1762  4.5096

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.41604    0.03977  10.460  < 2e-16 ***
L(g, 0:3)0  -0.46364    0.02268 -20.444  < 2e-16 ***
L(g, 0:3)1  -0.07868    0.02283  -3.446 0.000652 ***
L(g, 0:3)2   0.01349    0.02285   0.590 0.555377
L(g, 0:3)3  -0.01687    0.02274  -0.742 0.458805
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4465 on 290 degrees of freedom
Multiple R-squared:  0.6122,	Adjusted R-squared:  0.6069
```

An Example: Okun's Law

R Code

```
# HAC (heteroskedasticity and autocorrelation consistent)
standard errors
library(lmtest)
library(sandwich)
coeftest(okun.lag2, vcov=vcovHAC(okun.lag2))
coeftest(okun.lag3, vcov=vcovHAC(okun.lag3))
```

An Example: Okun's Law

```
t test of coefficients:

             Estimate Std. Error t value  Pr(>|t|)    
(Intercept)  0.405777   0.078807  5.1490 4.822e-07 ***
L(g, 0:2)0  -0.463326   0.146190 -3.1693 0.0016900 ** 
L(g, 0:2)1  -0.080491   0.020914 -3.8486 0.0001459 ***
L(g, 0:2)2   0.011732   0.050205  0.2337 0.8153969    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

An Example: Okun's Law

```
t test of coefficients:

              Estimate Std. Error t value  Pr(>|t|)
(Intercept)   0.416043   0.079357  5.2427 3.054e-07 ***
L(g, 0:3)0   -0.463644   0.145910 -3.1776 0.0016454 **
L(g, 0:3)1   -0.078684   0.020524 -3.8338 0.0001548 ***
L(g, 0:3)2    0.013488   0.047525  0.2838 0.7767519
L(g, 0:3)3   -0.016865   0.019900 -0.8475 0.3974147
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

An Example: The Phillips Curve

- The Phillips Curve is an empirical model that describes the relationship between inflation and unemployment and was named after A.W. Phillips, the economist who discovered this relationship. Mathematically, the relationship between unemployment and inflation can be expressed as:

$$INF_t = \beta_1 - \beta_2 \triangle U_t + \mu_t$$

- where INF_t is the inflation rate during period, INF_{t-1}^{E} denotes inflationary expectations during period t, and $\triangle U_t = U_{Ut} - U_{t-1}$. We can rewrite the above equation using more familiar regression terms as:

$$INF_t = INF_{t-1}^{E} - \gamma(U_t - U_N)$$

Homework 2

- Test the econometric validity of one of these models, for at least three time periods.
- Are the series under analysis stationary?
- Do the results iof the model mprove using seasonally adjusted series?
- Compare the robustness of the Ad Hoc and Koyck models, for at least three time periods.
- How does perform the ARDL model?
- What is the role and interpretation of the results in these famous time series models?
- What are the economic conclussions that yopu can extract from your model?

References

- F.F. Alt, “Distributed Lags”, Econometrica, vol. 10, 1942, pp. 113-128.
- J. Tinbergen, “Long-Term Foreign Trade Elasticities”, Metroeconomica, vol. 1, 1949, pp. 174-185.
- L.M. Kx‘oyck, Distributed Lags and Investment Analysis, North Holland, Ámsterdam, 1954.
- Milton Friedman, A Theory of the Consumption Function, National Bureau of Economic Research, Princeton University Press, Princeton, Nueva Jersey, 1957.