# 05 Regression with time lags: Autoregressive Distributed Lag Models

Andrius Buteikis, andrius.buteikis@mif.vu.lt
http://web.vu.lt/mif/a.buteikis/

# Introduction

The goal of a researcher working with time series data does not differ too much from that of a researcher working with cross-sectional data: they both aim to develop a regression relating a dependent variable to some explanatory variable.

However, the analyst using time series data will face two problems that the analyst using cross-sectional data will not encounter:

1. One time series variable can influence another with a time lag;
2. If the variable is nonstantionary, a problem known as spurious regression may arise.

One should always keep in mind this general rule: **if you have nonstationary time series variables then you should not include them in a regression model**. The appropriate route is to **transform the variables** before running a regression in order to make them *stationary*. An exception to this rule, which will be presented in a later topic, occurs when the variables in a regression model are non-stationary and *cointegrated*.

**In this chapter we will assume all variables in the regression are stationary.**

# The Distributed Lag (DL) Model

We say that the value of the dependent variable, at a given point in time, should depend not only on the value of the explanatory variable at that time period, but also on the values of the explanatory variable in the past. A simple model to incorporate such **dynamic** effects has the form:

$$Y_t = \alpha + \beta_0 X_t + ... + \beta_q X_{t-q} + \epsilon_t$$

The individual coefficients $\beta_i$, $i = 0, ..., q$, called **lag weights**, define the pattern of how $X$ affects $Y$ **over time**. These coefficients collectively comprise the *lag distribution*.

Since the effect of the explanatory variable does not happen all at once but rather over several time periods. This model is sometimes referred to as a **distributed** (or weighted) **lag model**. Coefficients can be interpreted as measures of the influence of the explanatory variable on the dependent variable. In this case, we have to be careful with timing.

For instance, we interpret results as '$\beta_2$ measures the effect of the explanatory variable two periods ago on the dependent variable, *ceteris paribus*'.

# Selection of Lag Order

When working with distributed lag models, we rarely know *a priori* exactly how many lags we should include. Appropriately, the issue of lag length selection becomes a data-based one where we use statistical means to decide how many lags to include. There are many different approaches to lag length selection in econometrics literature. Here we outline a common one that does not require any new statistical techniques. This method uses *t-tests* for whether $\beta_q = 0$ to decide the length. A common strategy is to:

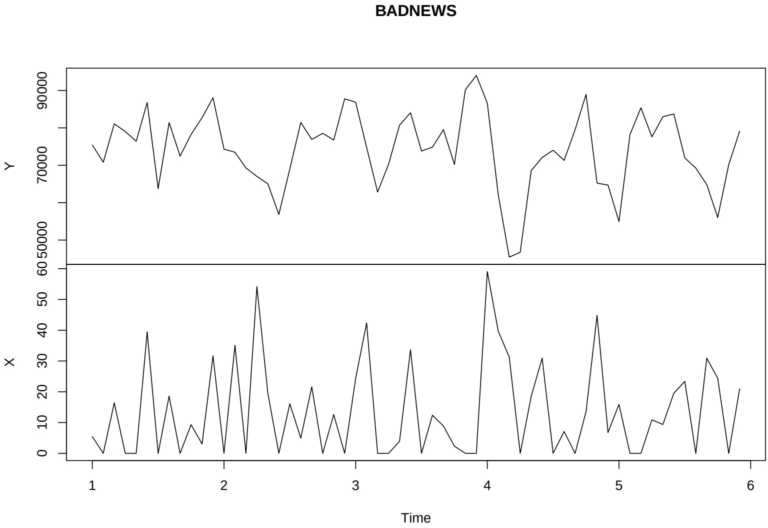
- Begin with a fairly large lag length, $q_{max}$, and test whether the coefficients on the maximum lag is equal to zero, i.e. test whether $\beta_{q_{max}} = 0$;
- If it is, drop the highest lag and re-estimate the model with the maximum lag equal to $q_{max} - 1$;
- If you find $\beta_{q_{max}-1} = 0$ in this new regression, then lower the lag order by one and re-estimate the mode;
- Keep on dropping the lag order by one and re-estimating the model until you reject the hypothesis that the coefficient on the longest lag is equal to zero.

## Exmaple: The Effect of Bad News on Market Capitalization

The share price of a company can be sensitive to bad news.
Suppose that Company B is in an industry which is particularly sensitive to the price of oil. If the price of oil goes up, then the profits of Company B will tend to go down and some investors, anticipating this, will sell their shares in Company B driving its price (and market capitalization) down.
However, this effect might not happen immediately. For instance, if Company B holds large inventories produced with cheap oil, it can sell these and maintain its profits for a while. But when new production is required, the higher oil price will lower profits.
Furthermore, the effect of the oil price jump might not last forever, since Company B also has some flexibility in its production process and can gradually adjust to higher oil prices. Hence, news about the oil price should affect the market capitalization of Company B, but the effect might not happen immediately and might not last too long.

Say we have data collected on a monthly basis over five years (i.e., 60 months) on the following variables:

- Y market capitalization of Company B ($000)
- X the price of oil (dollars per barrel) above the benchmark price



Since this is time series data and it is likely that previous months news about the oil price will affect current market capitalization, it is necessary to include lags of X in the regression. Below are present OLS estimates of the coefficients in a distributed lag model in which market capitalization is allowed to depend on present news about the oil price and news up to $q_{max} = 4$ months ago. That is:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + ... + \beta_4 X_{t-4} + \epsilon_t$$

```
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept) 91173.3150  1949.8502 46.7591   0.0000
## L(X, 0:4)0   -131.9943    47.4361 -2.7826   0.0076
## L(X, 0:4)1   -449.8597    47.5566 -9.4595   0.0000
## L(X, 0:4)2   -422.5183    46.7778 -9.0324   0.0000
## L(X, 0:4)3   -187.1041    47.6409 -3.9274   0.0003
## L(X, 0:4)4    -27.7710    47.6619 -0.5827   0.5627
```

Just looking at the coefficient values, what can we conclude about the effect of news about the oil price on Company B's market capitalization?

Increasing the oil price by one dollar per barrel in a given month is associated with:

1. An immediate reduction in market capitalization of $ 131'994, *ceteris paribus*.
2. A reduction in market capitalization of $ 449'860 on month later, *ceteris paribus*.

and so on. To provide some intuition about what the *ceteris paribus* condition implies in this context, note that, for example, we can also express the second statement as: 'Increasing the oil price by one dollar in a given month will tend to reduce the market capitalization in the following month by $ 449'860, **assuming that no other change in the oil price occurs**'.

Since the *p-value* corresponding to the explanatory variable $X_{t-4}$ is **greater** than 0.05, we cannot reject the null hypothesis that $\beta_4 = 0$ at the 5% level of significance. Accordingly we drop this variable from the model and re-estimate the lag length equal to 3, yielding the following results:

```
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept) 90402.2210  1643.1828 55.0165   0.0000
## L(X, 0:3)0   -125.9000    46.2405 -2.7227   0.0088
## L(X, 0:3)1   -443.4918    45.8816 -9.6660   0.0000
## L(X, 0:3)2   -417.6089    45.7332 -9.1314   0.0000
## L(X, 0:3)3   -179.9043    46.2520 -3.8896   0.0003
```

The *p-value* for testing $\beta_3 = 0$ is 0.0003, which is much less than 0.05. We therefore conclude that the variable $X_{t-3}$ does indeed belong in the distributed lag model. Hence $q = 3$ is the lag length we select for this model.

In a formal report, we would present this table of results. Since these results are similar to those discussed above, we will not repeat their interpretation.

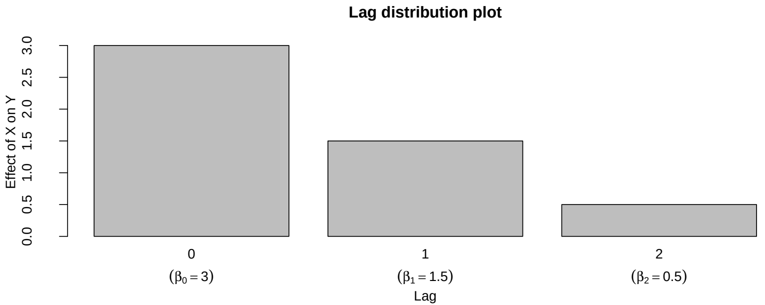
# Example: DL Model Coefficients and Effects based on Explanatory variable changes

Assume that our estimated model has the following coefficients:

$$Y_t = \alpha + 3 \cdot X_t + 1.5 \cdot X_{t-1} + 0.5 \cdot X_{t-2} + \epsilon_t$$

We can visualize the lag distribution (i.e. the lag weights) as follows:

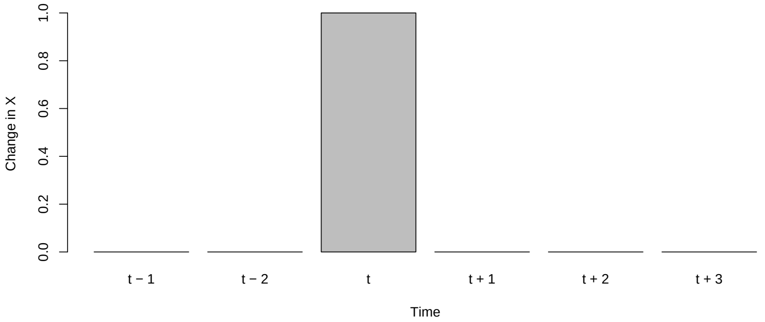
```
coefs <- c(3, 1.5, 0.5)
lbls <- NULL
for(i in 1:length(coefs)){
  lbls <- c(lbls, as.expression(bquote(atop(.(i-1),(beta[.(i-1)]==.(coefs[i]))))))
}
barplot(coefs, main = "Lag distribution plot", ylab = "Effect of X on Y",
        names.arg = lbls, xlab = "Lag", mgp = c(3, 2, 0)) # vs. mgp = c(3, 1, 0)
```



## Temporary change in $X$

To see the interpretation of the lag weights we begin by considering a temporary change in $X$.

```r
lbls <- c(paste0("t - ", 1:2), "t", paste0("t + ", 1:3))
barplot(c(0, 0, 1, 0, 0, 0), ylab = "Change in X",
        names.arg = lbls, xlab = "Time")
```




Suppose that $X$ increases *temporary* by one unit in period $t$ and then returns to its original lower value for periods $t+1$, $t+2$,...

Taking the partial derivatives, we can derive that the immediate response is given by

$$\frac{\partial Y_t}{\partial X_t} = \beta_0 = 3$$

Sometimes this is referred to as the **impact** (or **short-run**) multiplier (or <...> effect of $X$ on $Y$).

After one period the equation becomes:

$$Y_{t+1} = \alpha + 3 \cdot X_{t+1} + 1.5 \cdot X_t + 0.5 \cdot X_{t-1} + \epsilon_t$$

So the change in period $t+1$ (i.e. the interim multiplier, or the **dynamic marginal effect of $X$ on $Y$ at one lag**) is:

$$\frac{\partial Y_{t+1}}{\partial X_t} = \beta_1 = 1.5$$

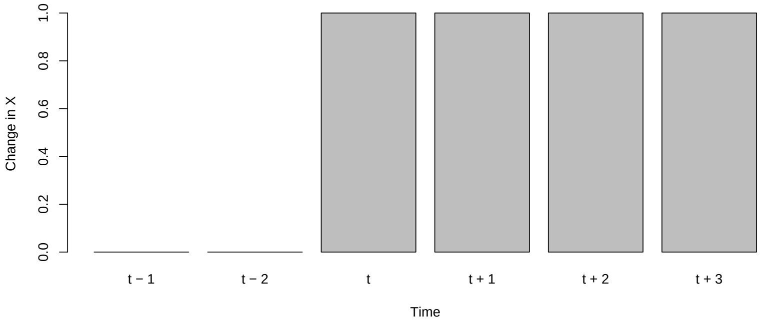
$$t+2 : \frac{\partial Y_{t+2}}{\partial X_t} = \beta_2 = 0.5, \quad t+k : \frac{\partial Y_{t+k}}{\partial X_t} = 0, \ k > 2$$

The effects of a temporary change in $X$ on $Y$ coincide with the lag distribution plot seen earlier.

# Permanent change in $X$

Now consider a **permanent** unit change in $X$.

```r
lbls <- c(paste0("t - ", 1:2), "t", paste0("t + ", 1:3))
barplot(c(0, 0, 1, 1, 1, 1), ylab = "Change in X",
        names.arg = lbls, xlab = "Time")
```




Suppose that $X$ increases by one unit in period $t$ and *remains higher* in all periods after $t$, than it was before $t$.

Now the dynamic marginal effects of $X$ on $Y$ are:

$$t : \frac{\partial Y_t}{\partial X_t} = \beta_0$$

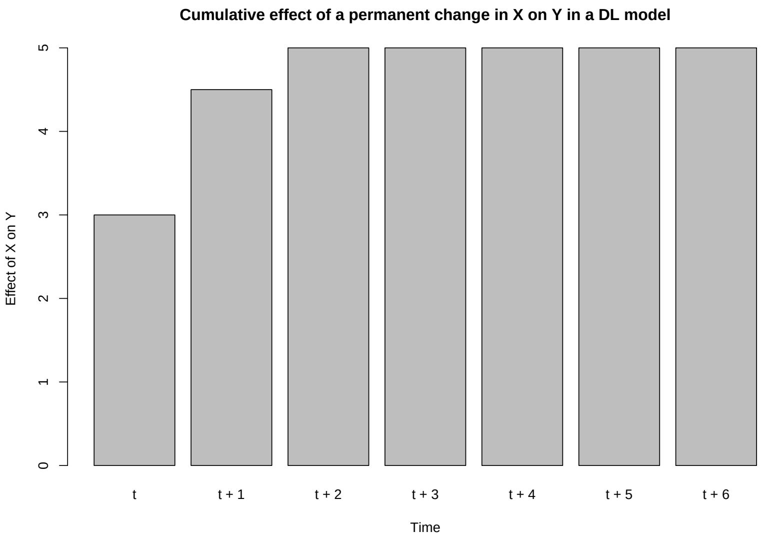
$$t+1 : \frac{\partial Y_{t+1}}{\partial X_t} + \frac{\partial Y_{t+1}}{\partial X_{t+1}} = \beta_0 + \beta_1$$

$$t+k : \sum_{j=0}^{k} \frac{\partial Y_{t+k}}{\partial X_{t+j}} = \beta_0 + \beta_1 + \beta_2,\ \ k > 1$$

The **Long-run cumulative effect of $X$ on $Y$** measures how much $Y$ will eventually change in response to a **permanent** change in $X$ on $Y$ as $t \longrightarrow \infty$:

$$\text{DL model: } \lim_{k \longrightarrow \infty} \sum_{j=0}^{k} \frac{\partial Y_{t+k}}{\partial X_{t+j}} = \sum_{k=0}^{p} \beta_k = \beta_0 + \beta_1 + \beta_2 = 5$$

```r
coefs <- c(3, 1.5, 0.5)
lbls <- c("t", paste0("t + ", 1:6))
barplot(c(cumsum(coefs), rep(sum(coefs), 4)),
        main = "Cumulative effect of a permanent change in X on Y in a DL model",
        ylab = "Effect of X on Y", names.arg = lbls, xlab = "Time")
```




Note: $Y_t$ contains lags of $X$, so the previous value changes in $Y_s$ have no effect on $Y_t$, $s \neq t$.

For some economic relationships it is possible that a **permanent change in $X$** leads to a **temporary change in $Y$**. This would be possible if the *positive* marginal effect at short lags would be offset by *negative* marginal effects at longer lags (or vice versa) so that the long-run cumulative effect is zero (i.e. $\sum_{i=0}^{p} \beta_i = 0$).

In such case a permanent change in $X$ would lead to a temporary change in $Y$ **over a finite number of periods** and $Y$ **would revert back to its original value at time** $t$.

For example, if $\beta_1 = 0.2$, $\beta_2 = 0.5$, $\beta_3 = -0.3$ and $\beta_4 = -0.4$. Then the long-run cumulative effect of $X$ on $Y$ is zero:

$$
\begin{aligned}
t &: \frac{\partial Y_t}{\partial X_t} = 0.2 \\
t+1 &: \frac{\partial Y_{t+1}}{\partial X_t} + \frac{\partial Y_{t+1}}{\partial X_{t+1}} = 0.2 + 0.5 = 0.7 \\
t+2 &: \frac{\partial Y_{t+2}}{\partial X_t} + \frac{\partial Y_{t+2}}{\partial X_{t+1}} + \frac{\partial Y_{t+2}}{\partial X_{t+2}} = 0.2 + 0.5 - 0.3 = 0.4 \\
t+k &: \sum_{j=0}^{k} \frac{\partial Y_{t+k}}{\partial X_{t+j}} = 0.2 + 0.5 - 0.3 - 0.4 = 0, \; k > 2
\end{aligned}
$$

# Dynamic Models with Stationary Variables

In regression analysis, researches are typically interested in measuring the effect of an explanatory variable (or variable**S**) on a dependent variable.

However, this goal is complicated when the researcher uses time series data since an explanatory variable may influence a dependent variable with a time lag.

This often necessitates the inclusion of lags of the explanatory variable in the regression. Furthermore, the dependent variable may be correlated with lags of itself, suggesting that lags of the dependent variable should also be included in the regression.

# Autoregressive Distributed Lag (ADL) Model

These considerations motive the commonly used **autoregressive distributed lag** (ADL) model:

$$Y_t = \alpha + \delta t + \phi_1 Y_{t-1} + ... + \phi_p Y_{t-p} + \beta_0 X_t + ... + \beta_q X_{t-q} + \epsilon_t$$

In this model:

- The dependent variable $Y$ depends on $p$ lags of itself;
- $Y$ also depends on the current value of an explanatory variable $X$ as well as $q$ lags of $X$;
- The model also allows for a deterministic trend $t$;
- Standard assumptions regarding residuals: $\mathbb{C}\mathrm{ov}(\epsilon_t, \epsilon_s) = 0$, for $t \neq s$ and $\mathbb{V}\mathrm{ar}(\epsilon_t) = \sigma^2$.

Since the model contains $p$ lags of $Y$ and $q$ lags of $X$, we denote it by $ADL(p, q)$.

In this chapter, we focus on the case where there is only *one* explanatory variable $X$. Note however, that we could equally allow for many explanatory variables in the analysis.

Let us consider two stationary variables $Y_t$ and $X_t$ and assume that it holds that:

$$Y_t = \alpha + \phi Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \epsilon_t, \quad 0 < \phi < 1$$

As an illustration, we can think of $Y_t$ as 'company sales' and $X_t$ as 'advertising', both in month $t$. If we assume that $\epsilon_t$ is a white noise process, independent of $X_t$, $Y_t$ and $X_{t-1}$ and $Y_{t-1}$, the above relation can be estimated by the use of ordinary least squares.

The interesting element in this equation is that it describes the dynamic effects of a change in $X_t$ upon current and future values of $Y_t$.

- In cross-sectional models we often used econometric models to evaluate the **marginal effect** of some independent variable $X$ on a dependent variable $Y$, *ceteris paribus* (i.e. holding all other independent variables constant): $\partial Y / \partial X$;
- In time-series models we must consider not only **how much** a change in $X$ affects $Y$ but also **when** the effect occurs - is it immediate, does it take place over a period of time, is it permanent?

Taking the partial derivatives, we can derive that:

- The **immediate** response (of a unit change in $X_t$ ) is given by

$$\partial Y_t / \partial X_t = \beta_0$$

  Sometimes this is referred to as the **impact** (or **short-run**) multiplier. An increase in $X_t$ with one unit has an immediate impact on $Y$ of $\beta_0$ units.
- The effect (of a unit change in $X_t$ ) after one period is:

$$\partial Y_{t+1} / \partial X_t = \phi \partial Y_t / \partial X_t + \beta_1 = \phi\beta_0 + \beta_1$$

Note: this can also be derived in a more explicit way:

$$\begin{aligned} Y_{t+1} &= \alpha + \phi Y_t + \beta_0 X_{t+1} + \beta_1 X_t + \epsilon_{t+1} \\ &= \alpha + \phi(\alpha + \phi Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \epsilon_t) + \beta_0 X_{t+1} + \beta_1 X_t + \epsilon_{t+1} \\ &= \alpha(1+\phi) + \phi^2 Y_{t-1} + \beta_0 X_{t+1} + \boldsymbol{(\phi\beta_0 + \beta_1) X_t} + \phi\beta_1 X_{t-1} + \phi\epsilon_t + \epsilon_{t+1} \end{aligned}$$

- Similarly, the effect (of a unit change in $X_t$ ) after two periods:

$$\partial Y_{t+2} / \partial X_t = \phi \partial Y_{t+1} / \partial X_t = \phi(\phi\beta_0 + \beta_1)$$

  and so on. This shows that after the first period, the effect is *decreasing* if $|\phi| < 1$.

Imposing this so-called **stability** condition allows us to determine the long-run effect of a unit **temporary** change in $X_t$:

$$\lim_{k \longrightarrow \infty} \frac{\partial Y_{t+k}}{\partial X_t} = \lim_{k \longrightarrow \infty} \phi^k (\phi \beta_0 + \beta_1) = 0 \Longleftrightarrow |\phi| < 1$$

This says that for an $\mathrm{ADL}(1,1)$ model, a temporary unit increase in $X_t$ results in a change in $Y_t$, which decreases as $t$ increases and returns to the initial value of $Y_t$.

On the other hand, if the increase in $X_t$ is permanent (imposing $\overline{Y_{t-1} = Y_t = Y_{t+1} ... = Y, X_{t-1} = X_t = X_{t+1} ... = X}$), then the changes in $X_t$, $X_{t+1}, ...$ lead to the following cumulative marginal effects:

$$t: \quad \frac{\partial Y_t}{\partial X_t} = \beta_0$$

$$t+1: \quad \frac{\partial Y_{t+1}}{\partial X_t} + \frac{\partial Y_{t+1}}{\partial X_{t+2}} = \phi \frac{\partial Y_t}{\partial X_t} + \beta_0 = \phi\beta_0 + \beta_1 + \beta_0$$

$$= \frac{\partial Y_{t+1}}{\partial X_t} + \frac{\partial Y_t}{\partial X_t}$$

$$t+2: \quad \frac{\partial Y_{t+2}}{\partial X_t} + \frac{\partial Y_{t+2}}{\partial X_{t+1}} + \frac{\partial Y_{t+2}}{\partial X_{t+2}} = \phi \frac{\partial Y_{t+1}}{\partial X_t} + \phi \frac{\partial Y_{t+1}}{\partial X_{t+1}} + \beta_1 + \beta_0$$

$$= \phi(\phi\beta_0 + \beta_1) + \phi\beta_0 + \beta_1 + \beta_0$$

$$= \frac{\partial Y_{t+2}}{\partial X_t} + \frac{\partial Y_{t+1}}{\partial X_t} + \frac{\partial Y_t}{\partial X_t}$$

...

$$t+k: \quad \sum_{j=0}^{k} \frac{\partial Y_{t+k}}{\partial X_{t+j}} = \frac{\partial Y_t}{\partial X_t} + \frac{\partial Y_{t+1}}{\partial X_t} + \frac{\partial Y_{t+2}}{\partial X_t} + ... + \frac{\partial Y_{t+k}}{\partial X_t}$$

Imposing the stability condition allows us to determine the long-run effect of a **permanent** increase in $X$. It is given by the **long-run multiplier** (or equilibrium multiplier):

$$\beta_0+(\phi\beta_0+\beta_1)+\phi(\phi\beta_0+\beta_1)+\ldots=\beta_0+(1+\phi+\phi^2+\ldots)(\phi\beta_0+\beta_1)=\frac{\beta_0+\beta_1}{1-\phi}$$

This says that if the unit increase in $X_t$ (e.g. advertising) is permanent, the **expected** long-run permanent cumulative increase (or decrease) in $Y$ (e.g. sales) is given by $(\beta_0+\beta_1)/(1-\phi)$.

The long-run equilibrium relation between $Y$ and $X$ can be seen by taking the expectations of both sides of the $\mathrm{ADL}(1,1)$ model, which, under stationarity with $\mathbb{E}(Y_t)=\mathbb{E}(Y)$ and $\mathbb{E}(X_t)=\mathbb{E}(X)$, $\forall t\in\mathbb{Z}$, yield:

$$\mathbb{E}(Y)=\alpha+\phi\mathbb{E}(Y)+\beta_0\mathbb{E}(X)+\beta_1\mathbb{E}(X)$$

or

$$\mathbb{E}(Y)=\frac{\alpha}{1-\phi}+\frac{\beta_0+\beta_1}{1-\phi}\mathbb{E}(X)=\tilde{\alpha}+\tilde{\beta}\mathbb{E}(X)$$

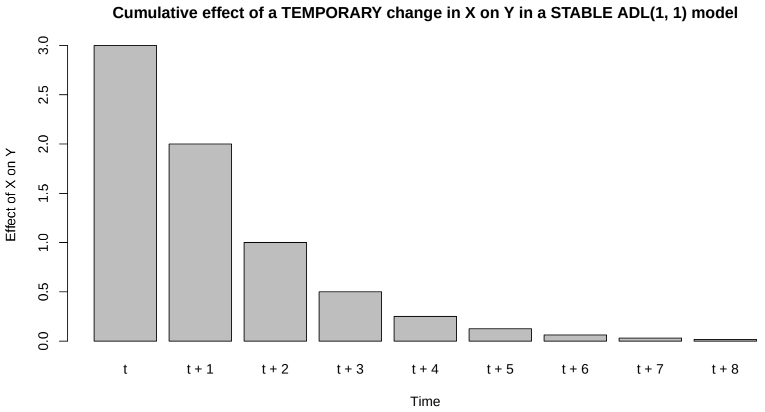
which presents an alternative derivation of the long-run multiplier - if $X$ changes to a new constant $\tilde{\tilde{X}}$, $Y$ will finally change to $\tilde{\tilde{Y}}=\tilde{\alpha}+\tilde{\beta}\tilde{\tilde{X}}$ (but it will take some time!).

# Example $\mathrm{ADL}(1, 1)$ model

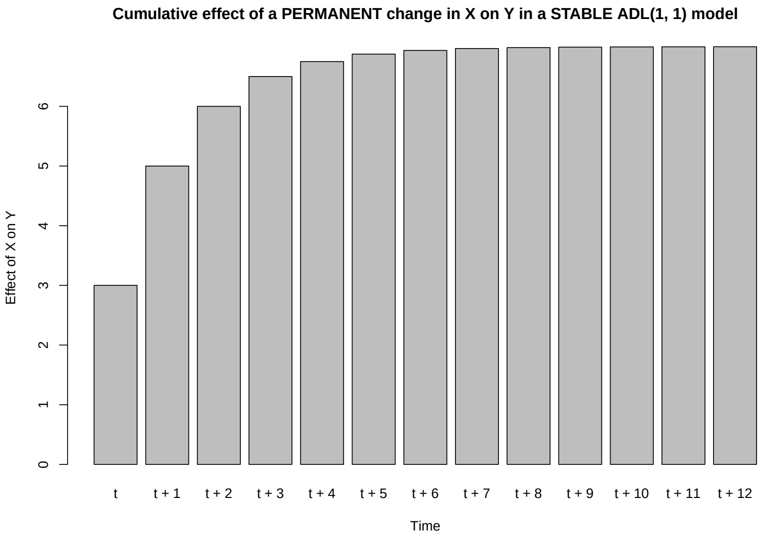
Assume that our $\mathrm{ADL}(1, 1)$ model is given by:

$$Y_t = \alpha + 0.5Y_{t-1} + 3X_t + 0.5X_{t-1} + \epsilon_t$$

```r
coefs <- c(3, 0.5)
lbls <- c("t", paste0("t + ", 1:8))
barplot(c(coefs[1], 0.5^(0:7) * (0.5 * coefs[1] + coefs[2])),
        main = "Cumulative effect of a TEMPORARY change in X on Y in a STABLE ADL(1, 1) model",
        ylab = "Effect of X on Y", names.arg = lbls, xlab = "Time")
```



```
coefs <- c(3, 0.5)
lbls <- c("t", paste0("t + ", 1:12))
barplot(cumsum(c(coefs[1], 0.5^(0:11) * (0.5 * coefs[1] + coefs[2]))),
        main = "Cumulative effect of a PERMANENT change in X on Y in a STABLE ADL(1, 1) model",
        ylab = "Effect of X on Y", names.arg = lbls, xlab = "Time")
```



The true long-run effect for this $\text{ADL}(1,1)$ model is $(3 + 0.5)/(1 - 0.5) = 7$. We could also examine at which point would we reach the eventual "long-run" for this process:

```
paste0("Cumulative effect at t + 10: ",
 sum(c(coefs[1], 0.5^(0:9) * (0.5 * coefs[1] + coefs[2]))))
```

```
## [1] "Cumulative effect at t + 10: 6.99609375"
```

```
paste0("Cumulative effect at t + 20: ",
 sum(c(coefs[1], 0.5^(0:19) * (0.5 * coefs[1] + coefs[2]))))
```

```
## [1] "Cumulative effect at t + 20: 6.99999618530273"
```

```
paste0("Cumulative effect at t + 50: ",
 sum(c(coefs[1], 0.5^(0:49) * (0.5 * coefs[1] + coefs[2]))))
```

```
## [1] "Cumulative effect at t + 50: 7"
```

# Setting up the Error Correction Model Form

## Another look at the long-run relationship

Assume that the long-run equilibrium relationship can be described as

$$Y = \widetilde{\alpha} + \widetilde{\beta} X$$

Alternatively, it could the logarithms of a proportional long-run equilibrium relationship $\tilde{Y} = k\tilde{X}$, where $\tilde{Y}$ can be though of as inventory and $\tilde{X}$ as sales; or $\tilde{Y}$ as consumption and $\tilde{X}$ as income, etc.
In general the equilibrium relationship may include more variables and it need not be directly proportional.
For example - the **Cobb-Douglas production function**:

$$Y = AL^{\beta} K^{\alpha},$$

where

- $Y$ - total production;
- $L$ - labor;
- $K$ - capital;
- $A$ - total productivity factor;
- $\alpha$ - capital output elasticity; $\beta$ - labor output elasticity;
- $\alpha, \beta$ are constants, determined by available technology.

Looking at our specified general dynamic relationship between $Y_t$ and $X_t$ as an $\mathrm{ADL}(1,1)$ model:

$$Y_t = \alpha + \phi Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \epsilon_t$$

We want to ask: under what conditions is the general dynamic relationship consistent with the long-run equilibrium relationship ?

To examine the long-run relationship, we the factors, which could cause divergence from the equilibrium will either "zero out" (i.e. disappear), or be equal to their expected value. This includes:

- The random component, $\epsilon_t$;
- Stochastic fluctuations, namely $X_{t-1}$ and $Y_{t-1}$

In the long-run, as $t \longrightarrow \infty$, the relationship of $(Y_t, X_t)$ will approach equilibrium and there will be no more shocks. The equilibrium implies that we will be talking about the expected values $\mathbb{E}(Y_t) = \mathbb{E}(Y_{t-1}) = Y^*$, $\mathbb{E}(X_t) = \mathbb{E}(X_{t-1}) = X^*$ and $\mathbb{E}(\epsilon_t) = 0$, $\forall t \in \mathbb{Z}$. This gives us:

$$Y^* = \alpha + \phi Y^* + \beta_0 X^* + \beta_1 X^*$$

$$Y^* = \frac{\alpha}{1-\phi} + \frac{\beta_0 + \beta_1}{1-\phi} X^*$$

Equating it with our assumed long-run relationship $Y = \widetilde{\alpha} + \widetilde{\beta} X$ yields the following parameter relationships:

$$\frac{\alpha}{1-\phi} = \widetilde{\alpha} \iff 1 - \phi = \frac{\alpha}{\widetilde{\alpha}} \iff \phi = 1 - \frac{\alpha}{\widetilde{\alpha}}$$

$$\frac{\beta_0 + \beta_1}{1-\phi} = \widetilde{\beta} \iff \beta_1 = (1-\phi)\widetilde{\beta} - \beta_0$$

This allows us to re-write the general dynamic relationship as:

$$\begin{aligned} Y_t &= \alpha + \phi Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \epsilon_t \\ &= \widetilde{\alpha}(1-\phi) + \left(1 - \frac{\alpha}{\widetilde{\alpha}}\right) Y_{t-1} + \beta_0 X_t + \left((1-\phi)\widetilde{\beta} - \beta_0\right) X_{t-1} + \epsilon_t \end{aligned}$$

or by applying the coefficient parametrizations between the long-run equilibrium and the dynamic equations:

$$\begin{aligned} Y_t - Y_{t-1} &= \widetilde{\alpha}(1-\phi) - \frac{\alpha}{\widetilde{\alpha}} Y_{t-1} + \beta_0 (X_t - X_{t-1}) + (1-\phi)\widetilde{\beta} X_{t-1} + \epsilon_t \\ \Delta Y_t &= \widetilde{\alpha}(1-\phi) - (1-\phi) Y_{t-1} + \beta_0 \Delta X_t + (\beta_0 + \beta_1) X_{t-1} + \epsilon_t \\ \Delta Y_t &= \beta_0 \Delta X_t - (1-\phi)\left[Y_{t-1} - \widetilde{\alpha} - \frac{\beta_0 + \beta_1}{1-\phi} X_{t-1}\right] + \epsilon_t \end{aligned}$$

gives

$$\Delta Y_t = \beta_0 \Delta X_t - (1-\phi)\left[Y_{t-1} - \widetilde{\alpha} - \widetilde{\beta} X_{t-1}\right] + \epsilon_t$$

Alternatively, we can get the same expression by adding and subtracting $Y_{t-1}$ and adding and subtracting $\beta_0 X_{t-1}$ (though there is no direct economic explanation).

# Error Correction Model (ECM): Introduction

The derived formulation:

$$\Delta Y_t = \beta_0 \Delta X_t - (1 - \phi) \left[ Y_{t-1} - \tilde{\alpha} - \tilde{\beta} X_{t-1} \right] + \epsilon_t$$

is an example of an **error-correction model** (ECM).

It says that the change in $Y_t$ is due to the current change in $X_t$ plus an error-correction term: if $Y_{t-1}$ is above the equilibrium value corresponding to $X_{t-1}$, that is, if the 'disequilibrium error' in the square brackets is positive, then a 'go to equilibrium' mechanism generates additional negative adjustment in $Y_t$. The speed of adjustment is determined by $1 - \phi$, which is the adjustment parameter. Note that stability assumption ensures that $0 < 1 - \phi < 1$. Therefore only a part of any disequilibrium is made up for in the current period.

If there were no adjustment to be made on account of a previous disequilibrium, and no random disturbance, then $Y_{t-1} = \tilde{\alpha} - \tilde{\beta} X_{t-1}$ and the equation becomes

$$\Delta Y_t = \beta_0 \Delta X_t$$

which implies that if equilibrium is to be maintained, the change in $Y_t$ should be equal to a **proportional** change in $X_t$. The reason for a *proportional change* is that if $\beta_1 \neq 1$, then changes in $X_t$ are themselves a source of disequilibrium (which will in turn call for adjustment of $Y$ in subsequent periods), and not just the random term $\epsilon_t$.

Notice that without prior knowledge of the long-run parameters, we cannot estimate the above ECM in its current form. This is because without knowing $\tilde{\alpha}$ and $\tilde{\beta}$, we cannot construct the disequilibrium error $Y_{t-1} - \tilde{\alpha} - \tilde{\beta} X_{t-1}$.

In the absence of such knowledge, in order to directly estimate the ECM, we must first multiply out the term in parenthesis to obtain (note that we already obtained this expression when deriving the $\mathrm{ECM}$):

$$\Delta Y_t = (1-\phi)\tilde{\alpha} + \beta_0 \Delta X_t - (1-\phi) Y_{t-1} + (1-\phi)\tilde{\beta} X_{t-1} + \epsilon_t$$

and $\Delta Y_t$ can now be OLS-regressed on $\Delta X_t$, $Y_{t-1}$ and $X_{t-1}$, obtaining estimates of all short-run and long-run parameters.

We can further generalize. For example, if:

$$Y_t = \alpha + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \epsilon_t$$

then, the ECM is:

$$\Delta Y_t = -\phi_2 \Delta Y_{t-1} + \beta_0 \Delta X_t - \beta_2 \Delta X_{t-1} - (1 - \phi_1 - \phi_2) \left[ Y_{t-1} - \tilde{\alpha} - \tilde{\beta} X_{t-1} \right] + \epsilon_t$$

Note that the original model must be rewritten in differences plus a disequilibrium error. To estimate this model, it is again necessary to express it by multiplying out the term in parenthesis.

## ECM: Considerations For the Next Lecture

It is possible for more than two variables to enter into an equilibrium relationship. For example:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \gamma_0 Z_t + \gamma_1 Z_{t-1} + \phi Y_{t-1} + \epsilon_t$$

This equation then can be transformed to:

$$\Delta Y_t = \beta_0 \Delta X_t + \gamma_0 \Delta Z_t - (1-\phi)\left[Y_{t-1} - \tilde{\alpha} - \tilde{\beta} X_{t-1} - \tilde{\gamma} Z_{t-1}\right] + \epsilon_t$$

All the ECM's may be consistently estimated via OLS **provided all the predictors are stationary**.

As long as it can be assumed that the error term $\epsilon_t$ is a *white noise* process, or - more generally - *is stationary* and independent of $X_t, X_{t-1}, ...$ and $Y_{t-1}, Y_{t-2}, ...$, the ADL models can be estimated *consistently* by ordinary least squares (OLS). Problems may arise, however, if, along with $Y_t$ and $X_t$, the implied $\epsilon_t$ is also non-stationary.

This will be discussed in the next topic.