

MACROECONOMETRICS. Dynamic Regression.

Topic 4: Autoregressive Distributed Lag Model.

Antoni Espasa

Tor Vergata

3rd May 2016

3rd May 2016

- **Autoregressive Distributed Lag model.**
- The dynamics of the model.
- Relationship with transfer function models.
- Impulse response functions.
- Parsimonious long lags.
- Examples.

REFERENCES

- Espasa et al. “Forecasting with dynamic regression models”, unpublished typescript notes, Statistics Department, UC3M.

The single-equation dynamic econometric model

- The model that we study in this topic corresponds to an equation of a recursive VAR.
- First we will consider that the variables are stationary or that the model can be properly formulated taking first differences.

Different forms of representing the model

- There are two main ways of formulating a dynamic regression model: either
- as a usual regression model with lagged variables, in which case it is called an *Autoregressive Distributed Lags (ADL)* model, or
- with a structure of rational lag polynomials, in which case it is referred to as a *Transfer Function (TF)* model.

A simple example of an ADL model

- It includes one lagged value of the dependent variable, y_{t-1} , and one explanatory variable, x_t , and its first lag, x_{t-1} , giving rise to the following equation:
-
- $$y_t = c + a_1 y_{t-1} + b_0 x_t + b_1 x_{t-1} + \varepsilon_t$$

(4.5)
-
- where c , a_1 , b_0 and b_1 are parameters and ε_t is a disturbance term assumed to be white noise.
-

- For example, the value of the **exports** of one country could depend on an **index of the world trade** this quarter and the previous one (although this simple example should obviously include **more variables** like real exchange rates).

NUMERICAL EXAMPLE

- $y_t = 0.5y_{t-1} + 0.4x_t + 0.6x_{t-1} + a_t$
- $(1 - 0.5L)y_t = (0.4 + 0.6L)x_t + a_t$

Generalization to more lags and variables

- The simple model (4.5) may be generalized to include more explanatory variables and more lagged values of the dependent and the explanatory variables.
- For example, including **more lags of both y_t and x_t** in (4.5) leads to the following general formulation for the case of a single explanatory variable (**dynamic regression model**):

-

- $$y_t = c + a_1 y_{t-1} + \dots + a_r y_{t-r} + b_0 x_t + b_1 x_{t-1} + \dots + b_s x_{t-s} + \varepsilon_t.$$

(4.6)

-

This equation can also be written in terms of lag polynomials as:

- $a(L) y_t = c + b(L) x_t + \varepsilon_t \quad (4.7)$

-

- where $a(L) = (1 - \sum_{j=1}^r a_j L^j)$,

- $b(L) = (\sum_{j=0}^s b_j L^j)$ and

- ε_t is white noise

- The lag polynomial formulation is interesting because provides a representation of the dynamics which can be analysed on itself.

The generalization of (4.7) to include more explanatory variables, say x_{1t}, \dots, x_{kt} , gives the following equation:

- $$a(L) y_t = c + \sum_{i=1}^k b_i(L) x_{it} + \varepsilon_t \quad (4.8)$$
- where $a(L), b_1(L), \dots, b_k(L)$ are finite order lag polynomials of degrees r, s_1, \dots, s_k , respectively, and ε_t is assumed to be white noise.
- Since we are assuming that y_t is stationary all the roots of $a(L)$ are outside the unit circle.
- Model (4.8) is usually denoted as $ADL(r, s_1, \dots, s_k)$. Obviously, (4.5) and (4.7) are particular cases of it.

The models are not given, they must be discovered.

- It should also be remarked that specifications such as (4.5) are not usually given at the beginning of the building process of the empirical model.
- This is something that must be discovered by the analyst.

-

The steps to arrive to a final model could have been as follows.

In this case all the variables must be stationary.

- Marginalization with respect other variables to end up with say x_{1t}, \dots, x_{kt} , and y_t .

The **excluded variables** should not Granger caused the selected ones.

- The x 's variables must be **strongly exogenous**. (Not Granger caused by variable y).
- Formulation of a general model like (4.8) **with a sufficient number of lags**.
- **General unrestricted model (GUM)**. This could be (4.8).

The wrong specification of the number of lags.

- Take the following example:

- $y_t = b_1 x_{t-1} + a_t \quad (1)$

on which you are initially considering that a_t is white noise.

But if the dynamics relating x_t and y_t are more complex than what is stated in (1), then the residuals will not be white noise.

This should be tested, for instance, estimating an AR(p) model for a_t selecting p by AIC. If the selected value is different from one, model (1) is wrongly specified and should be reformulated.

Suppose we conclude that a_t follows an AR(1) model:

$$a_t = r_1 a_{t-1} + \varepsilon_t. \text{ Then, since according with (1)}$$

$$a_{t-1} = y_{t-1} - b_1 x_{t-2}, \text{ we end up with}$$

$$y_t = r_1 y_{t-1} + b_1 x_{t-1} + b_2 x_{t-2} + \varepsilon_t. \quad (2)$$

After the AR(p) hypothesis for a_t in (1) is not rejected we should enlarge the initial model with p additional lags in all the variables.

ADL MODELS WITH COMMON FACTORS.

- In the previous example we have seen how we should proceed to **ensure a proper dynamic specification for our model**.
- **Another question** is that having started from a GUM with p lags, we arrive to a simpler model, like (2).
- (2) can be formulated as:

$$\begin{aligned}(1-r_1 L) y_t &= (-b_1 L + b_2 L^2) x_t + \varepsilon_t, \\ (1-r_1 L) y_t &= (-b_1 + b_2 L) x_{t-1} + \varepsilon_t, \\ (1-r_1 L) y_t &= b_1(1 - \beta_2 L) x_{t-1} + \varepsilon_t, \quad (3)\end{aligned}$$

In (3) we can test if the one-root polynomials $(1-r_1 L)$ and $(1 - \beta_2 L)$ have the root in common. If this is the case,

$$\begin{aligned}(1-r_1 L) y_t &= b_1(1-r_1 L) x_{t-1} + \varepsilon_t. \text{ or} \\ y_t &= -b_1 x_{t-1} + \varepsilon_t / (1-r_1 L) \quad (4)\end{aligned}$$

We see that if in an ADL model, (3), we have **common roots** in the polynomials, **the model can be write as an ADL model of lower order but with an AR residual**.

But (4) can't be starting point model, because it includes a restriction that should be tested.

A GENERAL ADL MODEL WITH COMMON ROOTS.

From the GUM to a specific model.

- We should start from a general formulation (GUM) and then proceed to test down for zero restrictions in the corresponding polynomials in order to get a simpler model.
- Suppose that we have conclude that there is only one explanatory variable. Then the starting point could be a model like (4.6) that, for values of r and s sufficient large, can be taken as a general unrestricted model for testing the relevant hypothesis on its parameters.
-

- In particular, testing first for zero restrictions on the parameters of $a(L)$ and $b(L)$, we could find that the hypothesis $H_0: c=a_2=\dots=a_r=b_2=b_3=\dots=b_s=0$ is not rejected and so specify the model as
-
- $y_t = a_1 y_{t-1} + b_0 x_t + b_2 x_{t-1} + \varepsilon_t$
-
- or equivalently as
-
- $(1-a_1L) y_t = (b_0 + b_1L) x_t + \varepsilon_t.$

Problems in formulating a GUM for an ADL model.

- 1.- The selection of the **included variables**.

The set on included regressors must incorporate variables suggested by economic theory and sociological, environmental, meteorological... variables which could affect our variable of interest.

- 2.- Check that they are **strongly exogenous**.

The **formulation of the dynamics** is relatively simple: include a sufficient number of lags to ensure that the residuals are white noise.

Stability condition in the ADL model

$$a(L) y_t = c + \sum_{i=1}^k b_i(L) x_{it} + \varepsilon_t \quad (4.8)$$

It depends only on the roots of $a(L)$, which should be outside the unit circle. If this is the case, polynomial $a(L)$ can be inverted and, ignoring the intercept to keep notation simple, the ADL model can be written as

•

$$\bullet \quad y_t = \sum_{i=1}^k \beta_i(L) x_{it} + \frac{1}{a(L)} \varepsilon_t. \quad (4.12)$$

•

- where **$\beta_i(L) = b_i(L)/a(L)$** .
- This equation decompose the value of y_t in **two terms**: the first term captures the systematic dynamics due to the influence of the explanatory variables on the actual value of y_t
- while the second term reflects residual dynamics, which is the dynamics in y_t not explained by the k explanatory variables in the model.

This representation is a special case of *a Transfer Function model.*

- Within few slides we will go further into this kind of models.
- For the moment, it is enough to note that the residual term in (4.12), $\varepsilon_t/a(L)$, has its own lag structure and therefore, is somehow forecastable, while this is not the case in the ADL formulation, for instance equation (4.8), where the disturbance term, ε_t , is already white noise.

Equations (4.8) and (4.12) represent exactly the same model

$$a(L) y_t = c + \sum_{i=1}^k b_i(L) x_{it} + \varepsilon_t \quad (4.8)$$

- $y_t = \sum_{i=1}^k \beta_i(L) x_{it} + \frac{1}{a(L)} \varepsilon_t.$ (4.12)
- where $\beta_i(L) = b_i(L)/a(L).$
- Having estimated one immediately we can formulate the other.

With the roots of $a(L)$ outside the unit circle

- The residual term in (4.12) is stationary and
- The dynamic relationships with the exogenous variable converges to zero.
- The $b_i(L)$'s don't matter for the stability of the system. They just reflect the specificity of the dynamics of each regressor.
- But in all cases the convergence is governed by $a(L)$.

The importance of the $\beta(L)$ polynomials.

- They reflect the dynamic relationships of the dependent variable with each regressor.

EXAMPLE

$$y_t = 0.5y_{t-1} + 0.4x_t + 0.6x_{t-1} + a_t$$
$$(1 - 0.5L)y_t = (0.4 + 0.6L)x_t + a_t$$

$$y_t = \frac{0.4 + 0.6L}{1 - 0.5L} x_t + \frac{a_t}{1 - 0.5L}$$

- $y_t = (0.4 + 0.8L + 0.4L^2 + 0.2L^3 + \dots)x_t + w_t$
- $(1 - 0.5L)w_t = a_t[AR(1)]$

EXAMPLE

- **DIVIDENDS AND EARNINGS IN US**

- Both series are $I(1)$ with zero mean.

AN ADL MODEL FOR DIVIDENDS As function of EARNINGS

- In this example, the objective is to forecast an aggregate, economy-wide measure of dividend yield that is the dividend divided by share price.
- An obvious explanatory variable is aggregate earnings (profits) divided by share prices. The data used are
 -
 - D_t – Dividend yield, aggregate dividend/price ratio
 - E_t – Aggregate company earnings/price ratio
 -
- The series are observed quarterly, starting in 1978:1 and ending in 2006:4. The sample size is $n = 116$ and the source is Standard and Poors.
- This example can be representative for the dividend yield of a particular stock and the earnings of the corresponding company.

- For an individual investor it could be very **useful to have an accurate forecast** of the income which he could expect from this stock for, say, next quarter or next year.
- This could be forecast by an **ARIMA model**, using only past data on dividends, but since dividends depend on the **earnings of the company**, enlarging the univariate information set with earnings one can build a model which should provide more accurate forecasts.
- Besides, since in a given quarter, **earnings are announced before dividends**, that information of earnings could provide an additional improvement of the forecast of current quarter dividends.

EXAMPLE

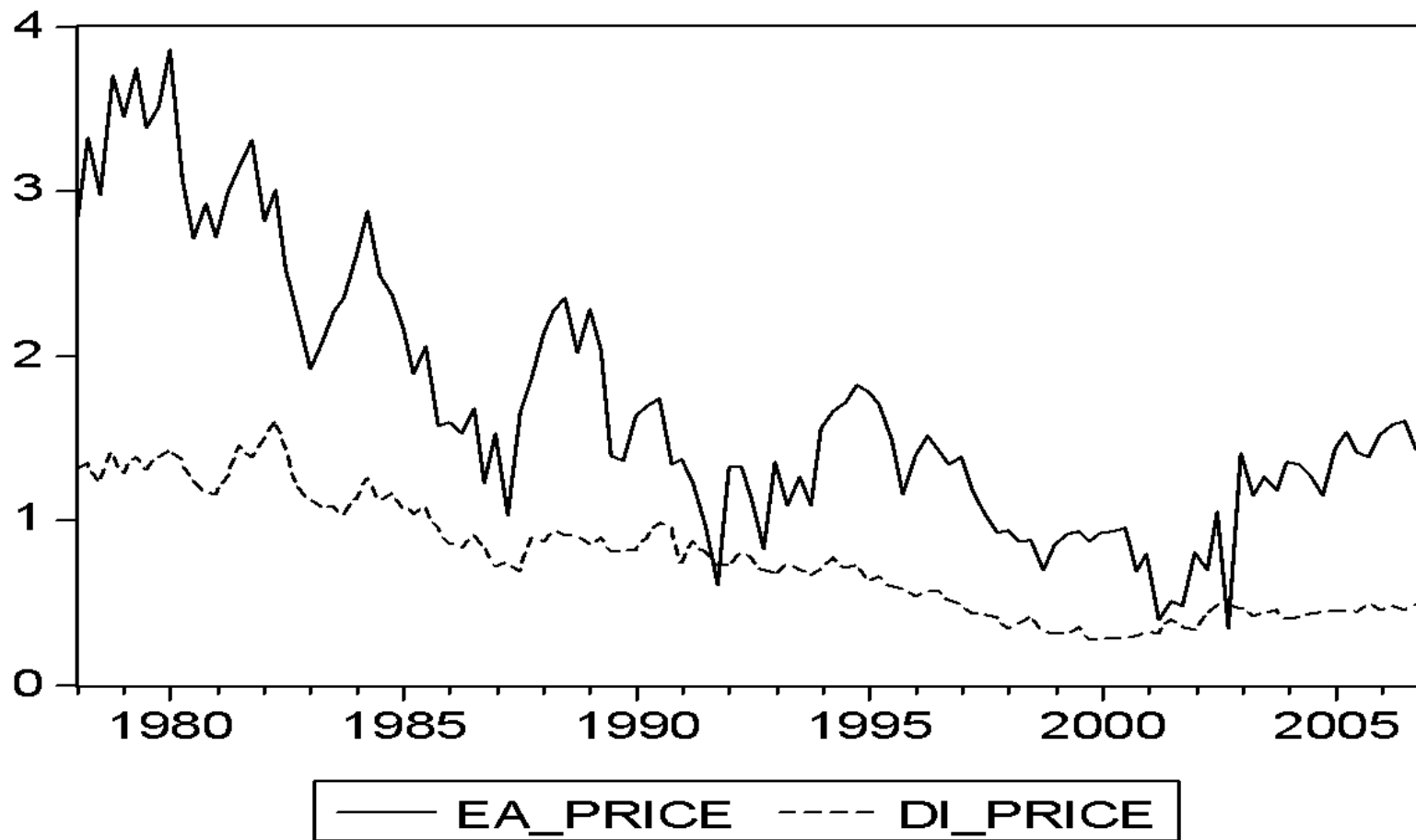
prepared by prof. Ana Pérez
Espartero

Example on dividends
and profits, after
stationary
transformations
of both variables.

Stationary transformation

- The data are plotted in Figure 4.2.
- Being non-stationary the **original series** seem to evolve rather smoothly with possible upward and downward local trends along time.
- Unlike, their **first differences**, plotted in Figure 4.3, move up and down in a sharply way around a constant zero mean, a typical pattern of the stationary series.
- In fact, both variables D_t and E_t are nonstationary and, for reasons to be explained later **their relationship will be formulated on their stationary transformations, ΔD_t and ΔE_t .**

Figure 4.2



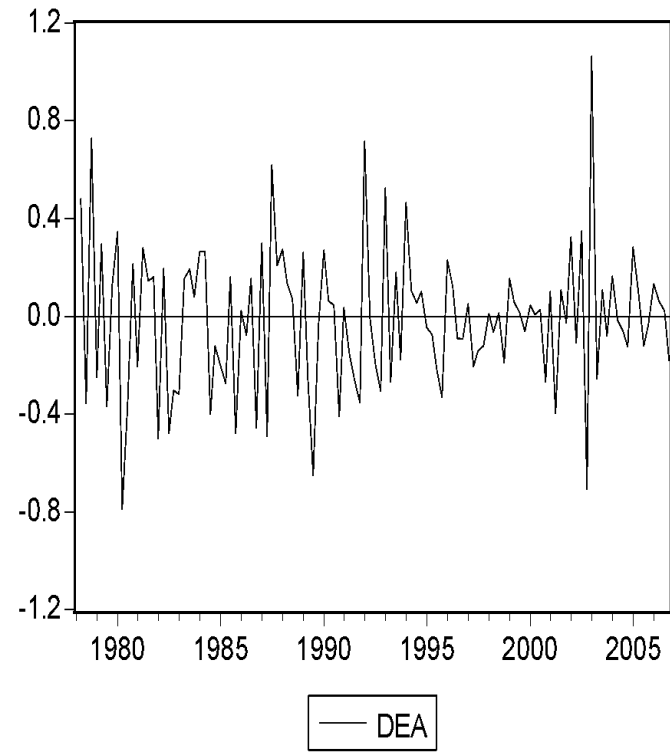
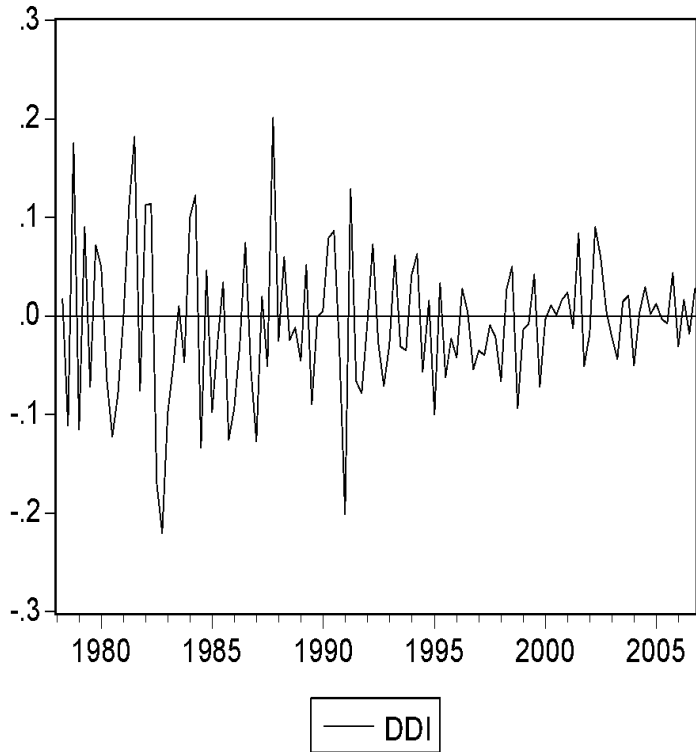
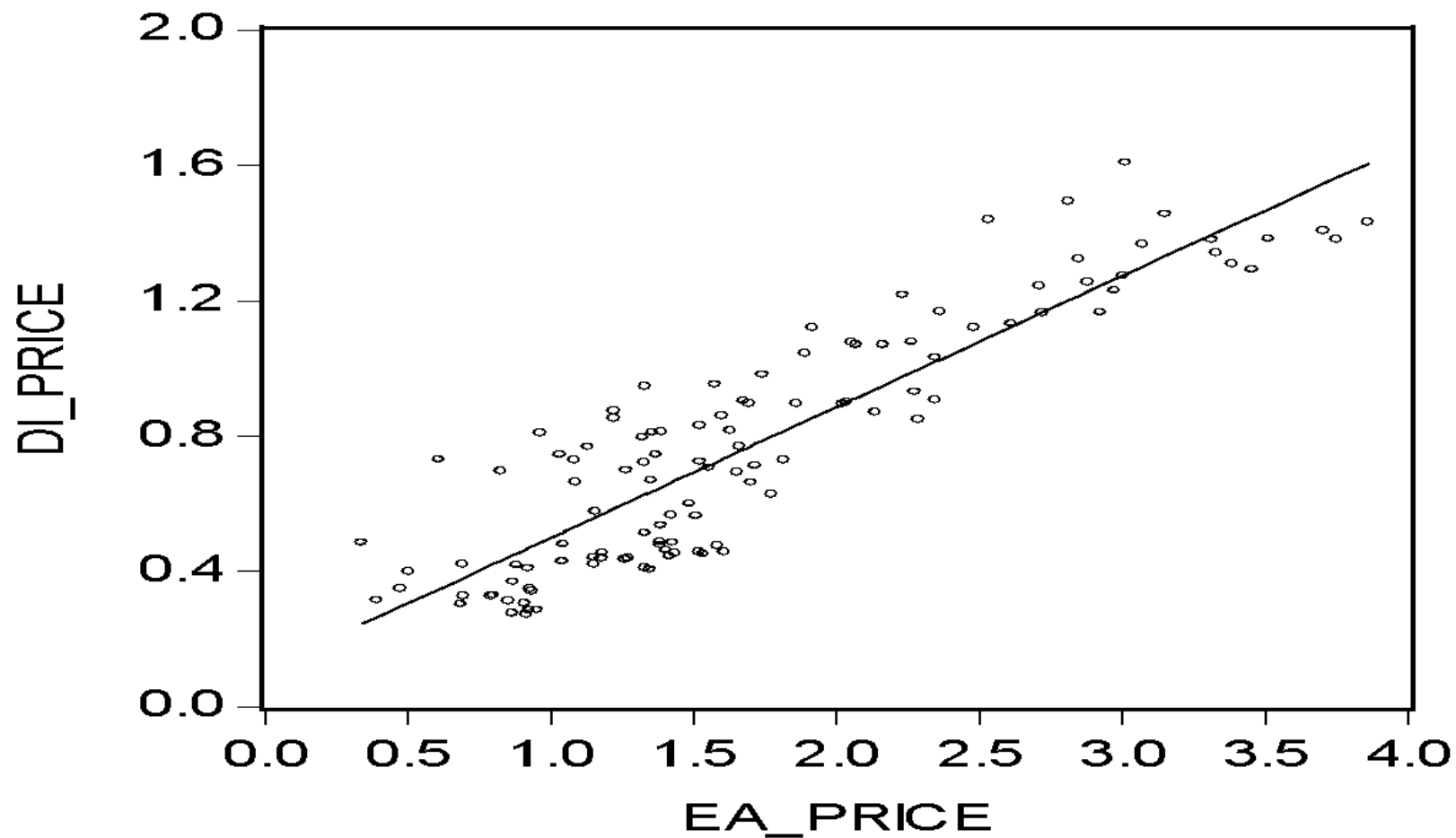


FIGURE 4. 3

DI_PRICE vs. EA_PRICE



Hypothesis Testing for DDI

Sample (adjusted): 1978Q2 2006Q4

Included observations: 115 after adjustments

Test of Hypothesis: Mean = 0

Sample Mean = -0.007294

Sample Std. Dev. = 0.073998

	<u>Value</u>	<u>Probability</u>
t-statistic	-1.056988	0.2928

$E(\Delta D_t)=0$ CAN NOT be rejected

Hypothesis Testing for DEA

Sample (adjusted): 1978Q2 2006Q4

Included observations: 115 after adjustments

Test of Hypothesis: Mean = 0

Sample Mean = -0.012349

Sample Std. Dev. = 0.296744

<u>Method</u>	<u>Value</u>	<u>Probability</u>
t-statistic	-0.446278	0.6562

$E(\Delta E_t)=0$ CAN NOT be rejected

DIVIDENDS

Correlogram

Sample: 1978:1 2006:4 Included observations: 116						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
. *****	. *****	1	0.965	0.965	110.83	0.000
. *****	. .	2	0.933	0.024	215.32	0.000
. *****	. .	3	0.906	0.063	314.77	0.000
. *****	. .	4	0.876	-0.051	408.63	0.000
. *****	. .	5	0.850	0.043	497.81	0.000
. *****	. .	6	0.825	-0.007	582.53	0.000
. *****	. *	7	0.810	0.140	664.83	0.000
. *****	. .	8	0.796	0.030	745.16	0.000
. *****	. *	9	0.771	-0.153	821.20	0.000
. *****	. *	10	0.743	-0.074	892.51	0.000
. *****	. .	11	0.717	-0.007	959.48	0.000
. *****	. .	12	0.691	0.011	1022.3	0.000
. *****	. .	13	0.665	-0.005	1081.1	0.000
. *****	. .	14	0.637	-0.041	1135.6	0.000
. *****	. **	15	0.601	-0.202	1184.6	0.000
. *****	. *	16	0.574	0.080	1229.7	0.000
. *****	. *	17	0.538	-0.137	1269.7	0.000
. *****	. *	18	0.493	-0.131	1303.6	0.000
. ****	. .	19	0.456	0.044	1332.9	0.000
. ****	. *	20	0.430	0.147	1359.2	0.000
. ****	. .	21	0.407	0.006	1383.1	0.000
. ****	. *	22	0.389	0.102	1405.1	0.000
. ****	. .	23	0.370	-0.015	1425.2	0.000
. ****	. .	24	0.353	0.001	1443.8	0.000
. ****	. *	25	0.329	0.094	1460.1	0.000

FROM A GENERAL GUM TO A SPECIFIC MODEL.

- Since the variable to be forecast is D_t and **we consider E_t as exogenous** (after a Granger causality test), a single-equation model of the ADL type is proposed.
- the estimation is carried out only with data up to 2004:4. The **general ADL formulation** used to start the modelling process has been **an ADL(5,5)**.
- Testing for zero restrictions in the coefficients of that model, it has been found that the null hypothesis that the coefficients of lags 1, 2, 3 and 5 of ΔD_t and of lags 2, 3 and 5 of ΔE_t are equal zero, is not rejected. This leads to **the following model**
-
- $$\Delta D_t = 0.18 \Delta D_{t-4} + 0.13 \Delta E_t + 0.09 \Delta E_{t-1} - 0.06 \Delta E_{t-4} + \varepsilon_t.$$

(4.13)

ADL models IN FIRST DIFFERENCES: $\Delta DIVID = f(\Delta EARN)$

(CALCULOS HECHOS CON PCGETS)

Sample: 1978:1-2004:4

GUM Modelling DDi by GETS (using SP500_QUARTERLY_78) , 1979(3)-2004(4)

	Coeff	StdError	t-value	t-prob
Constant	-0.00447	0.00664	-0.673	0.5029
DDi_1	-0.22286	0.10492	-2.124	0.0364
DDi_2	-0.09403	0.10572	-0.889	0.3762
DDi_3	-0.02942	0.10455	-0.281	0.7790
DDi_4	0.17383	0.10378	1.675	0.0974
DDi_5	0.00229	0.09838	0.023	0.9815
DEa	0.12007	0.02454	4.892	0.0000
DEa_1	0.12709	0.02788	4.558	0.0000
DEa_2	0.06464	0.02928	2.208	0.0298
DEa_3	0.02381	0.02963	0.804	0.4238
DEa_4	-0.06337	0.02998	-2.114	0.0373
DEa_5	-0.00619	0.02766	-0.224	0.8235

RSS	0.35357	sigma	0.06268	R^2	0.35967	Radj^2	0.28141
LogLik	288.89690	AIC	-5.42935	HQ	-5.30430	SC	-5.12053
T	102	p	12	FpNull	0.00002	FpConst	0.00002

	value	prob	alpha
Chow(1992:2)	0.4202	0.9981	0.0100
Chow(2002:2)	0.8615	0.5721	0.0100
normality test	2.1432	0.3425	0.0100
AR 1-4 test	2.0182	0.0990	0.0100
ARCH 1-4 test	4.0920	0.0045	0.0000
hetero test	1.9443	0.0199	0.0050

Significance levels (alpha) set for subsequent tests; 1 test with alpha = 0 excluded.

Specific model of DDi, 1979 (3) - 2004 (4)

	Coeff	StdError	t-value	t-prob	Split1	Split2	reliable
DDi_4	0.18257	0.08888	2.054	0.0426	0.0067	0.0159	1.0000
DEa	0.12452	0.02293	5.432	0.0000	0.0000	0.0000	1.0000
DEa_1	0.09613	0.02223	4.325	0.0000	0.0000	0.0000	1.0000
DEa_4	-0.06811	0.02277	-2.991	0.0035	0.0005	0.0041	1.0000

$$\Delta D_t = 0.1826 \Delta D_{t-4} + 0.1245 \Delta E_t + 0.0961 \Delta E_{t-1} - 0.0681 \Delta E_{t-4} + \varepsilon_t$$

RSS	0.38090	sigma	0.06234	R^2	0.31018	Radj^2	0.28906
LogLik	285.09998	AIC	-5.51176	HQ	-5.47008	SC	-5.40882
T	102	p	4	FpNull	0.00000	FpGUM	0.54524

	value	prob
Chow(1992:2)	0.4023	0.9991
Chow(2002:2)	0.9175	0.5211
normality test	0.8803	0.6439
AR 1-4 test	2.0348	0.0958
hetero test	2.0493	0.0494

The model can also be written as:

-

$$\Delta D_t = 0.12 \Delta D_{t-4} + 0.13 \Delta E_t + 0.09 \Delta E_{t-1} + 0.06 (\Delta D_{t-4} - \Delta E_{t-4}) + \varepsilon_t.$$

-

- This representation indicates that not only increments in earnings matter in explaining increments in dividends, but also the past differences between both increments, in particular the difference in the previous year.

In terms of lag-polynomials, model (4.13) can be written as

-
- $(1 - 0.18L^4) \Delta D_t = (0.13 + 0.09L - 0.06L^4) \Delta E_t + \varepsilon_t. \quad (4.14)$
-
- Since both polynomials in (4.14) have not common roots, model (4.13) can not be simplified further and can be taken as a final model to forecast ΔD_t or D_t . In this last case, the model for D_t is derived from (4.13) by passing D_{t-1} to the right hand side as follows:
-
- $D_t = D_{t-1} + 0.18 \Delta D_{t-4} + 0.13 \Delta E_t + 0.09 \Delta E_{t-1} - 0.06 \Delta E_{t-4} + \varepsilon_t.$
-
- We will further develop this example in next sections.
-

Equation for earnings

- Only depends on its own
lags

.

Transfer Function models

-
- An alternative choice to ADL models for modelling dynamic relations are the *transfer functions* models.
- These models make use of rational distributed lag structures in modelling both the cross-variable dynamics and the residual dynamics. A general representation of this type would be given by the following equation:
-
- $$y_t = \sum_{i=1}^k \frac{\omega_i(L)}{\delta_i(L)} x_{it} + \eta_t \quad (4.15)$$
-
- where $\omega_i(L)$ and $\delta_i(L)$ are finite order lag polynomials, $\omega_i(L) = \omega_{i0} + \omega_{i1}L + \dots +$ and $\delta_i(L) = \delta_{i0} + \delta_{i1}L + \dots +$, and the residual term, η_t , follows an ARMA process given by:
-
- $$\eta_t = \theta(L) / \phi(L) \varepsilon_t, \quad (4.16)$$
-
- where $\theta(L)$ and $\phi(L)$ are finite order lag polynomials and ε_t is white noise.

STATIONARITY CONDITION

- Since we are assuming that the variables are stationary all the denominator polynomials $\delta_i(L)$ and $\phi(L)$ must have all their roots outside the unit circle.
- We also assume invertibility and then all the roots of $\theta(L)$ are also outside the unit circle.
-

EQUATION (4.15) CAN BE SPLIT IN TWO TERMS.

- The first one,

$$\sum_{i=1}^k \frac{\omega_i(L)}{\delta_i(L)} x_{it}$$

- represents the dynamic contribution of the explanatory variables to the contemporaneous value of the dependent variable.
- For each variable x_{it} , the filter $\omega_i(L)/\delta_i(L)$ embodies the dynamic relationship between this variable and y_t and its coefficients are called the *impulse response function* from x_{it} to y_t .

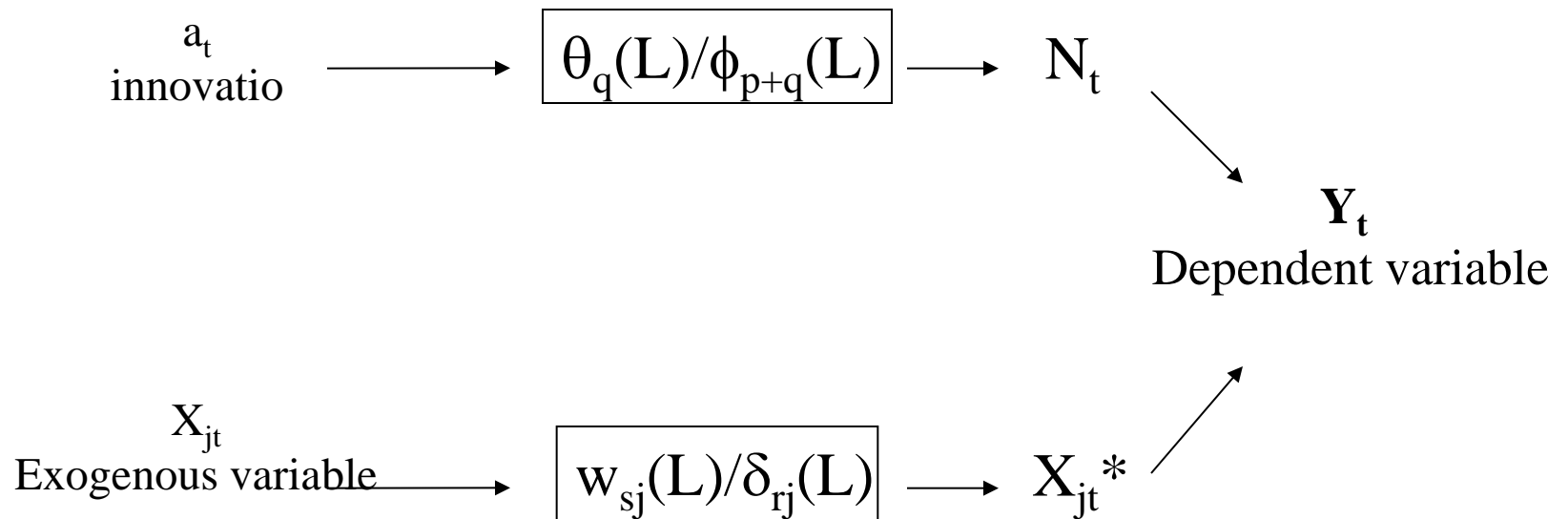
THE RESIDUAL TERM

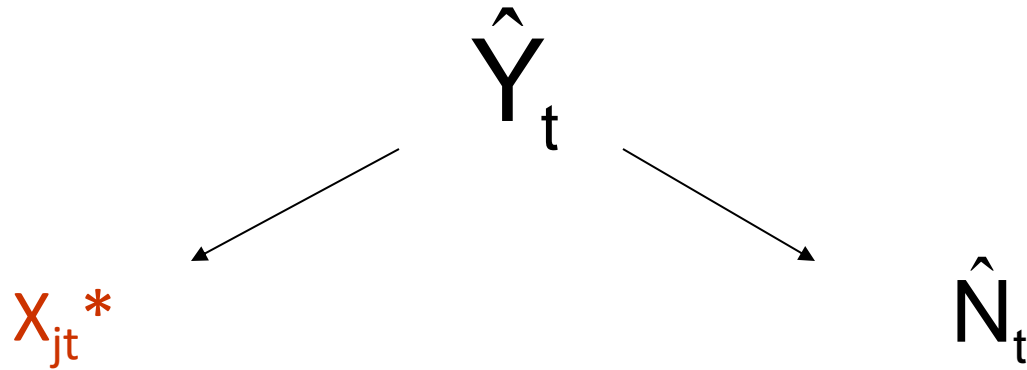
- The second term in the right hand side of equation (4.15), η_t , gathers the dynamics of y_t which are not explained by the explanatory variables.
- This term depends on its own past values and therefore is somehow forecastable.
- Hence, the forecastability of y_t is not completely accounted for by the explanatory variables, but also by the forecastable part of the residual term η_t .
-

- This can be better illustrated by noting that the invertible and stationary ARMA structure of η_t can be approximated by a high order AR(p) polynomial, say
-
- $$\theta(L)/\Phi(L) \cong 1/a(L), \quad (4.17)$$
-
- with $a(L) = (1 - a_1L - \dots - a_pL^p)$.

- Then the residual term η_t can be written down as
-
- $a(L)\eta_t = \varepsilon_t,$
-
- or equivalently as
-
- $\eta_t = (a_1L + \dots + a_pL^p)\eta_t + \varepsilon_t.$
-
- Using this decomposition of η_t , model (4.15) admits an equivalent representation as
-
- $y_t = \sum_{i=1}^k \frac{\omega_i(L)}{\delta_i(L)} x_{it} + \hat{\eta}_t + \varepsilon_t.$
-

FIGURE 4: The dynamic single equation model





- Contribution of the explanatory variables.
- The filters must be stationary, even when working with non-stationary variables.
- This is the forecastable part of the residual term.
- When Y_t is non-stationary the filter can have unit roots.

The dynamics behind model (4.15) can be summarized in the scheme in Figure 4.4.

- In this figure, we first find the dynamic filter determining the way in which each explanatory variable x_{it} affects the dependent variable y_t .
- Secondly, we have the residual structure, given by an ARMA process on the innovations ε_t .
- The joint effect of both elements makes up the observed value of the dependent variable y_t .
-

Balanced models

- The equality in equation (4.15) implies that all the properties of y_t must also hold in the right hand side of such equation.
- In Figure 4.4 we see that these properties come from the inputs, explanatory variables and innovations , and from the filters acting on them.
- If the model is well specified –all the relevant explanatory variables are included and the filters in the model are correct–, the residual term η_t can be represented in terms of a white noise error, like in (4.16), and we say that the model is balanced.

THE TRANSFER FUNCTION MODEL CAN BE FORMULATED AS AN ADL

- **TF:** $y_t = \sum_{i=1}^k \frac{w_i(L)}{\delta_i(L)} X_{it} + \frac{\theta(L)}{\delta(L)} e_t$.
- Since all the denominator polynomials of the regressors converge,
- the corresponding rational polynomials $\left(\frac{w_i(L)}{\delta_i(L)}\right)$ can be approximated by simple polynomial $(\beta_i(L))$.
- Also since the ARMA residual model is invertible, the arma polynomial $\left(\frac{\theta(L)}{\delta(L)}\right)$ can be approximated by an autoregressive polynomial $(\alpha(L))$.
- THEN
- $y_t = \sum_{i=1}^k \beta_i(L) X_{it} + \frac{e_t}{\alpha(L)}$
- multiplying by $\alpha(L)$ and denoting $b_i(L) = \beta_i(L)\alpha(L)$ we end up with:
-
- **ADL:** $\alpha(L)y_t = \sum_{i=1}^k b_i(L)X_{it} + e_t$

THE PROBLEM OF COMMON FACTORS IN RATIONAL LAGS

- The formulation of ADL models was simpler: include a sufficient number of lags which ensures that the model has a white noise error term.
- In TF models we face the problem that any rational polynomial can be modified adding a common factor in the numerator and the denominator without changing the model. Therefore the TF model is not identified unless it is restricted to non-common factor rational polynomials.
- In this context it is difficult to proceed from a general TF to a specific one, because the formulation of the former faced the common factor problem.
- For that reason the transfer function models are less popular.

A DEMAND MODEL FOR CURRENCY IN SPAIN

$$\Delta_4 \ln E_T = \frac{0,18}{1 - 0,84L} \Delta_4 \ln P_t + \frac{0,20}{1 - 0,60L} \Delta_4 \ln Y_t + \frac{1}{(1 - 0,99L + 0,40L^2)(1 + 0,65L^4)} a_t$$

E_t : Currency

P_t : CPI

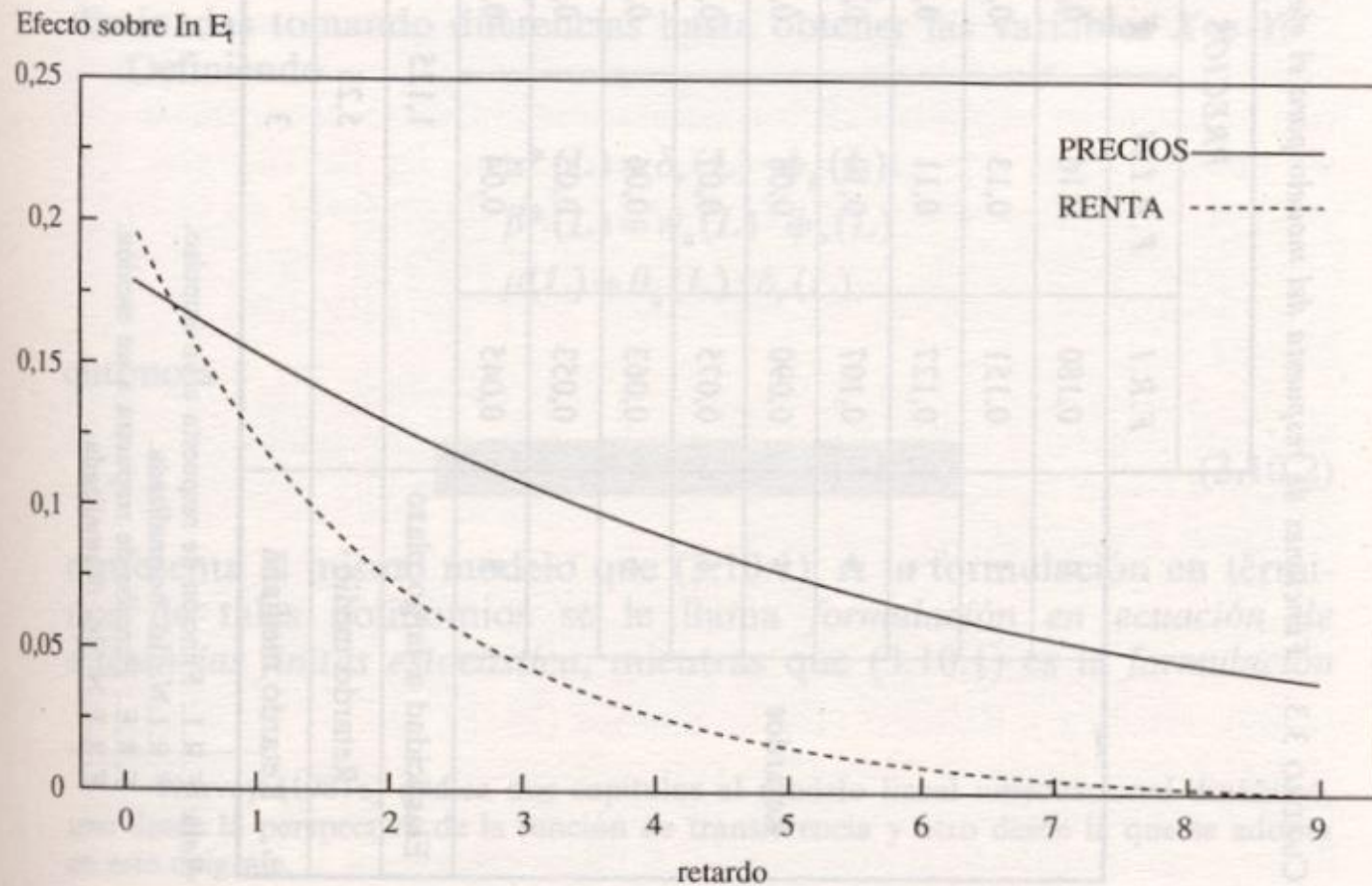
Y_t : GDP

CUADRO 3.3. Funciones de respuesta del modelo para el efectivo

		PRECIOS				PIB			
		F.R.I.	F.R.I.N.	F.R.E.	F.R.E.N.	F.R.I.	F.R.I.N.	F.R.E.	F.R.E.N.
Retardos	0	0,180	0,16	0,180	0,16	0,200	0,40	0,200	0,40
	1	0,151	0,13	0,331	0,29	0,120	0,24	0,320	0,64
	2	0,127	0,11	0,458	0,41	0,072	0,14	0,392	0,78
	3	0,107	0,10	0,565	0,50	0,043	0,09	0,435	0,87
	4	0,090	0,08	0,655	0,58	0,026	0,05	0,461	0,92
	5	0,075	0,07	0,730	0,65	0,016	0,03	0,477	0,95
	6	0,063	0,06	0,793	0,70	0,009	0,01	0,486	0,97
	7	0,053	0,05	0,846	0,75	0,006	0,01	0,492	0,98
	8	0,045	0,04	0,891	0,79	0,003	0,01	0,495	0,99
Elasticidad a largo plazo		1,125				0,5			
Retardo medio		5,25				1,5			
Retardo mediana		3				1			

Nota: F.R.I.: Función de respuesta a un impulso.
F.R.I.N.: Id. normalizada.
F.R.E.: Función de respuesta a un escalón.
F.R.E.N.: Id. normalizada.

GRÁFICO 3.2. Funciones de respuesta a un impulso para el modelo del efectivo.



- **Impulse
response
functions.**

$$\gamma_t = \sum_{i=1}^k v_i(L)x_{it} + \eta_i$$

The above model can proceed from an ADL or TF model.

η_t is an AR(p) process and
the lag polynomials $v_i(L)$ can be expanded as

$$v_i(L) = v_{i0} + v_{i1}L + v_{i2}L^2 + \dots$$

The coefficients of this lag polynomial, $\{v_{i0}, v_{i1}, \dots\}$, are called the ***impulse response function (IRF)*** and track the complete dynamic response of y to a explanatory variable x_i

THE COEFFICIENTS OF THE IRF

- In particular, u_{ij} represents the effect on the dependent variable of a transitory unit shock in the variable x_j occurred j periods before.

THE GAIN

Furthermore, the sum of all these coefficients, $\{u_{i0} + u_{i1} + \dots\}$, is usually called the *gain* and

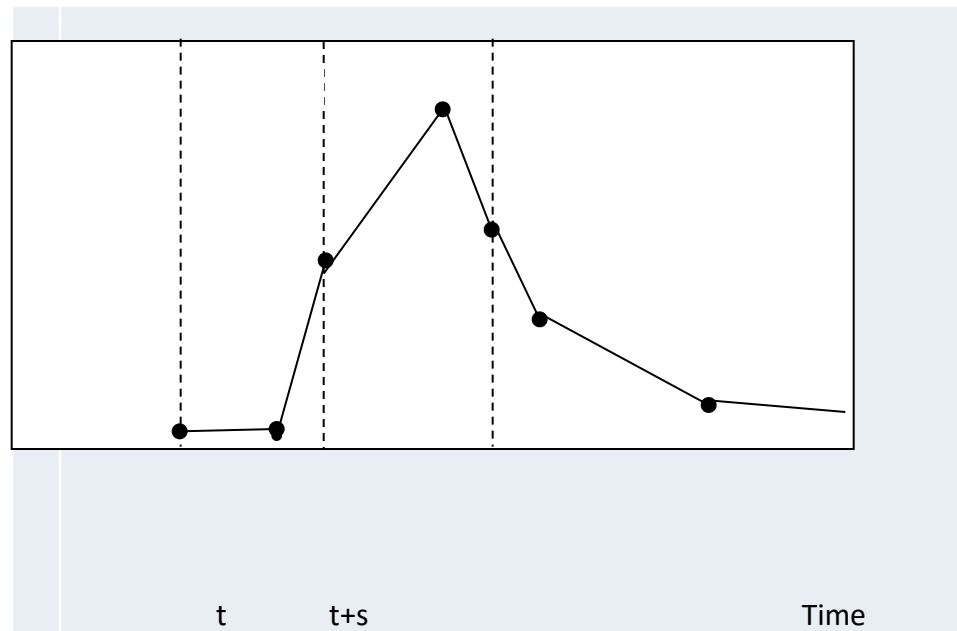
obviously represents the accumulation of all the impacts on y from a transitory unit shocks in x_i .

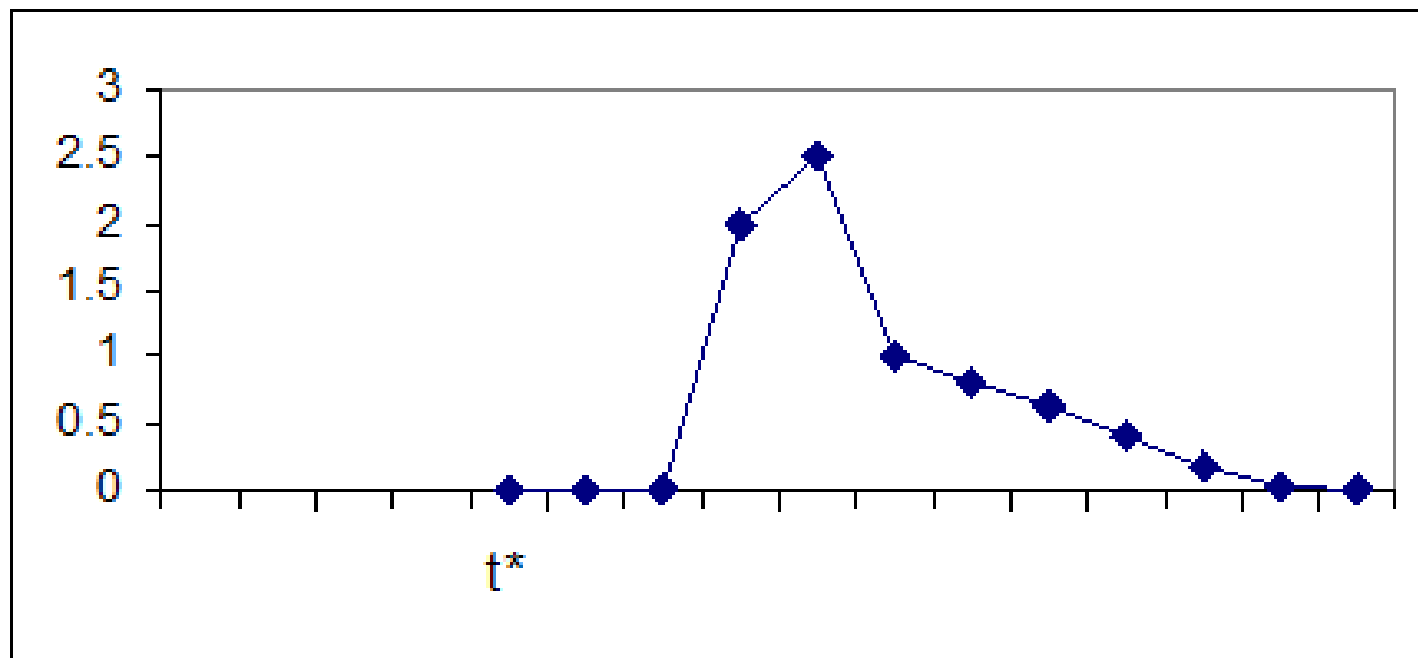
This value is obtained by taking $L=1$ in the corresponding lag polynomials, i.e.,

$$\text{gain} = u_i(1) = \omega_i(1) / \delta_i(1) = (\omega_{i0} + \dots + \omega_{is}) / (\delta_{i0} + \dots + \delta_{ir}).$$

Phases in the IRF

- In many cases the impulse response function can be characterized as having three phases.
- These are shown in Figure 4.5, which displays the effect on y_{t+j} , $j \geq 0$, of a transitory change in an explanatory variable x_t .
-
-
-





Phase 1

- First, there can be a delaying phase (**Phase I**) in which a transitory change in x_t does not affect the variable y . This phase lasts $s (\geq 0)$ periods if the reaction of y does not occur up to s periods after the change in x .
- For example, let us consider a very simple model like
-
- $$y_t = c + 0.8 x_{t-2} + 2 x_{t-3} + \varepsilon_t, \quad (4.19)$$
-
- where y_t and x_t could be the weekly sales of a certain product and the expenses on advertising it, respectively. A transitory change in the advertising expenditure made at time t^* does not affect weekly sales till two weeks later.

For example, let us consider a very simple model like

$$y_t = c + 0.8 x_{t-2} + 2 x_{t-3} + \varepsilon_t, \quad (4.19)$$

where y_t and x_t could be the weekly sales of a certain product and the expenses on advertising it, respectively.

A transitory change in the advertising expenditure made at time t^* does not affect weekly sales till two weeks later.

UNDERSTANDING THE IRF

- In order to better understand the meaning of the IRF in this example we will assume that **the system is in equilibrium**, i.e. the error term is zero, and x_t is taking an equilibrium value, say x^e . Then we have
-
- $y_t = y^e = c + 0.8 x^e + 2 x^e$ (model in logs)
-
- However, **if a unit transitory shock in x occurs at time t^*** , so that x_{t^*} becomes $x^e + 1$, and afterwards comes back to x^e , we can write down the model equation for the periods at and after that change as follows
-
- $y_{t^*} = c + 0.8 x^e + 2 x^e = y^e$
- $y_{t^*+1} = c + 0.8 x^e + 2 x^e = y^e$
- $y_{t^*+2} = c + 0.8 (x^e + 1) + 2 x^e = (c + 0.8 x^e + 2 x^e) + \mathbf{0.8} = y^e + \mathbf{0.8}$
- $y_{t^*+3} = c + 0.8 x^e + 2 (x^e + 1) = (c + 0.8 x^e + 2 x^e) + \mathbf{2} = y^e + \mathbf{2}$
- $y_{t^*+4} = c + 0.8 x^e + 2 x^e = y^e$
-
- and so on.

Delaying phase

- In these equations above, we note that the values of y_{t^*} and y_{t^*+1} are unaffected by the change in x_{t^*} and **only at time t^*+2 changes start in y .**
- Thus, we can say that in this relationship there is a delaying phase of two periods, t^* and t^*+1 , in which there is no impact on sales due to the change in the advertising expenditures at t^* .
- Obviously, in a contemporaneous relationship between x_t and y_t , as **in the dividends and earnings example (4.13), this delaying phase does not exist.**

Free-response phase

$$y_t = c + 0.8 x_{t-2} + 2 x_{t-3} + \varepsilon_t, \quad (4.19)$$

- After the delaying phase, if it exists, or otherwise from the very beginning, y_t can show a free response phase (Phase II), i.e. a response with unconstrained parameters to the transitory change in x_{t^*} .
- In the sales-advertising example above, the response of y starts at t^*+2 and is given by the extra terms in the third and fourth equations above, namely 0.8 and 2, respectively.
- An increment of 1% in advertising at time t^* will increase sales by 0.8% at time t^*+2 and by 2% at time t^*+3 . From t^*+4 onwards the effect is zero.

Convergence phase

- Finally, in some cases, there is a **third phase**, which we call the convergence phase, in which **the effect on y of a transitory change in x declines smoothly to zero**.
- Obviously, this phase does not exist in the example above, where immediately after period t^*+3 , the change in x has no longer effects on y and the impact on y_{t^*+3+h} , $h>0$, becomes abruptly zero without a smooth transition.
- However, **the convergence phase does appear in models with lagged endogenous variables**. For instance, in a model like
-
- **$y_t = c + 0.8 x_{t-2} + 2 x_{t-3} + 0.5 y_{t-1} + \varepsilon_t$** , (4.20)
-
- the presence of the first lag of the endogenous variable with a coefficient 0.5 keeps 50% of the value of y_t in next period.

The autoregressive factor

- In general, if this coefficient is α , 100α % of the value of y_t is always kept in next period. In order to illustrate this property, let us write down model (4.20) as
-
- $(1-0.5L) y_t = c + (0.8 L^2 + 2L^3) x_t + \varepsilon_t$, (4.21)
-
- and pass the autoregressive polynomial multiplying y_t to the right hand side. Then the model becomes:
- $y_t = c' + \frac{0.8L^2 + 2L^3}{1-0.5L} x_t + \frac{\varepsilon_t}{1-0.5L}$,
-
- and the impulse response filter, $\nu(L)$, takes the form:
-
- $\nu(L)=b(L)/a(L)$
-
- where $a(L)=1-0.5L$ and $b(L)= 0.8L^2 + 2L^3$. Now the filter has a rational form and $a(L)$ is the polynomial that generates the convergence phase in the impulse response function

Calculation of the IRF in presence of endogenous lags.

- In general, if this coefficient is α , 100α % of the value of y_t is always kept in next period. In order to illustrate this property, let us write down model as

$$(1-0.5L) y_t = c + (0.8 L^2 + 2L^3) x_t + \varepsilon_t, \quad (4.21)$$

and pass the autoregressive polynomial multiplying y_t to the right hand side. Then the model becomes:

$$y_t = c + \frac{0.8L^2 + 2L^3}{1 - 0.5L} x_t + \frac{\varepsilon_t}{1 - 0.5L} ,$$

and the impulse response filter, $v(L)$, takes the form:

$$v(L) = b(L)/a(L)$$

where $a(L)=1-0.5L$ and $b(L)=0.8L^2+2L^3$.
Now the filter has a rational form and $a(L)$ is the polynomial that generates the convergence phase in the impulse response function. In this case, the coefficients of $v(L)$ are obtained by considering that

$$(1-0.5L) v(L) = (0.8L^2 + 2L^3)$$

Expanding $u(L)$ as $u(L)=u_0+u_1 L+u_2 L^2+\dots$ and working out the product of polynomials on the left hand side of the equation above, the coefficients $\{u_0, u_1, u_2, \dots\}$ are derived by equating powers of L on both sides of the equation. In particular, it turns out that:

$$u_0 = 0;$$

$$u_1 = 0;$$

$$u_2 = 0.8;$$

$$u_3 = 2 + 0.5 * 0.8 = 2.4;$$

$$u_4 = 0.5 (2 + 0.5 * 0.8) = 1.2;$$

$$u_5 = 0.5 [0.5 (2 + 0.5 * 0.8)] = 0.5^2 (2 + 0.5 * 0.8) = 0.6;$$

and so on

Summing up the results in this example

- **PHASE 1:** Thus, in this model, a change in x at time t^* has no effect on y at times t^* and t^*+1 ($\nu_0=\nu_1=0$) but
- afterwards the sequence of responses of y at times t^*+2, t^*+3, \dots are given by the coefficients ν_2, ν_3, \dots , respectively.
- **PHASE 2:** The first non-zero coefficient, ν_2 , tells us that an increase of 1% in advertising expenses in a given week will generate an increase in sales of 0.8% two weeks later, t^*+2 . The subsequent coefficient tell us that there will be an additional effect of 2.4% three weeks later, t^*+3 and
- **PHASE 3:** from there onwards the subsequent effects will be just the 50% of the previous one, i.e. 1.2% four weeks later, and so on.
- **The gain or cumulative effect** of the transitory unit shocks along all future time, which is given, in this case, by $\nu(1)=b(1)/a(1)=(0.8+2)/(1-0.5) = 5.6$. That is, an increase of 1% in current advertising expenses will amount to a global increase of 5.6% on future sales.

$$y_t = c + 0.8 x_{t-2} + 2 x_{t-3} + 0.5 y_{t-1} + \varepsilon_t, \quad (4.20)$$

the presence of the first lag of the endogenous variable with a coefficient 0.5 keeps 50% of the value of y_t in next period.

Therefore the IRF has not a breaking point with value zero.

The zero value is obtain in the infinity.

The IRF is declining to zero at the rate given by the coefficient of the endogenous variable.

Thus, in this model, a change in x at time t^* has no effect on y at times t^* and t^*+1 ($u_0=u_1=0$) but afterwards the sequence of responses of y at times t^*+2, t^*+3, \dots are given by the coefficients u_2, u_3, \dots , respectively.

The first non-zero coefficient, u_2 , tells us that an increase of 1% in advertising expenses in a given week will generate an increase in sales of 0.8% two weeks later.

- The subsequent coefficient tell us that there will be an additional effect of 2.4% three weeks later and
- from there onwards the subsequent effects will be just the 50% of the previous one, i.e. 1.2% four weeks later, and so on

Note that in this example, from t^*+4 onwards, the response of y to a transitory change in x_{t^*} is given by some coefficients which are constrained by the expression

$$u_j = 0.5 u_{j-1} = 0.5^{j-3} u_3, \quad \text{for } j \geq 4, \quad (4.22)$$

with $u_3 = 2.4$.

Note also that as the periods go further in the future, the impact decreases and it becomes zero in the very far apart lags ($u_j \rightarrow 0$ as $j \rightarrow \infty$).

This is so because the $a(L)$ polynomial in (4.21) has its roots outside the unit circle, in fact, it has only one root with a value of two.

When the IRF converges to zero,

- - either in a smooth way, if the model includes endogenous lags, or
 - in an abrupt one if it does not,

we see that the relationship between these two variables is such that an impulse –a transitory shock- in one variable (exogenous) **has not a permanent effect** on the other (endogenous).

To summarize, in this second example there is:

- a **delaying phase** of no response which lasts two periods (t^* and t^*+1),
- following by a second **phase of free response** in periods t^*+2 and t^*+3 , and
- finally a **convergence phase** starting at t^*+4

- . Thus, this example shows that the presence of **the first lag of the endogenous variable** (this can be generalized to the presence of more lags or lags of any order) **extends the dynamic** relationship between y_t and x_t beyond the horizon given by the largest lag of x_t included in the model.

THE LONG RUN EFFECT: GAIN.

$$y_t = c + 0.8 x_{t-2} + 2 x_{t-3} + 0.5 y_{t-1} + \varepsilon_t,$$
$$(1-0.5L) y_t = c + (0.8 L^2 + 2L^3) x_t + \varepsilon_t$$

GAIN

- Summing up all these coefficients we get the gain or cumulative effect of the transitory unit shocks along all future time,
- which is given, in this case, by
$$v(1)=b(1)/a(1)=(0.8+2)/(1-0.5) = 5.6.$$
- That is, an increase of 1% in current advertising expenses will amount to a global increase of 5.6% on future sales.

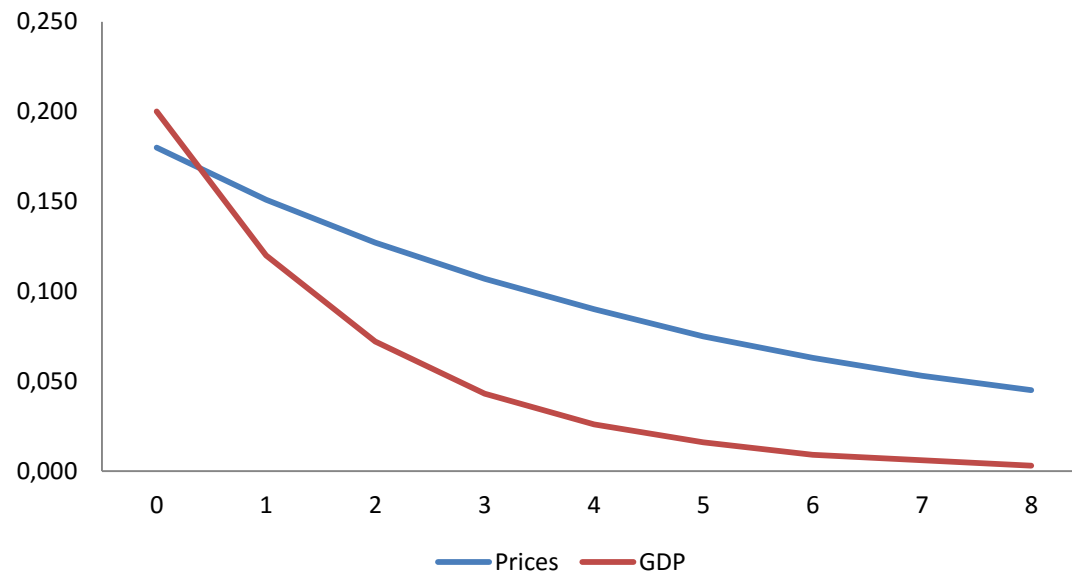
- **DEMAND FOR
CURRENCY IN SPAIN**

THE IMPULSE RESPOND FUNCTION WITH RESPECT PRICES

$$\begin{aligned}\frac{0,18}{1 - 0,84L} &= 0,18(1 + 0,84L + 0,84^2L^2 + 0,84^3L^3 + \dots) = \\ &= 0,18 + 0,151L + 0,127L^2 + 0,107L^3 + \dots\end{aligned}$$

IMPULSE RESPOND FUNCTION WITH RESPECT GDP

$$\frac{0,20}{1 - 0,60L} = 0,20(1 + 0,60L + 0,60^2L^2 + 0,60^3L^3 + \dots) =$$
$$0,20 + 0,12L + 0,072L^2 + 0,043L^3 + \dots$$



		PRICES				GDP			
		F.R.I	F.R.I.N.	F.R.E.	F.R.E.N.	F.R.I	F.R.I.N.	F.R.E.	F.R.E.N.
Lags	0	0.180	0.16	0.180	0.16	0.200	0.40	0.200	0.40
	1	0.151	0.13	0.331	0.29	0.120	0.24	0.320	0.64
	2	0.127	0.11	0.458	0.41	0.072	0.14	0.392	0.78
	3	0.107	0.10	0.565	0.50	0.043	0.09	0.435	0.87
	4	0.090	0.08	0.655	0.58	0.026	0.05	0.461	0.92
	5	0.075	0.07	0.730	0.65	0.016	0.03	0.477	0.95
	6	0.063	0.06	0.793	0.70	0.009	0.01	0.486	0.97
	7	0.053	0.05	0.846	0.75	0.006	0.01	0.492	0.98
	8	0.045	0.04	0.891	0.79	0.003	0.01	0.495	0.99
Gain		1.125				0.5			
Mean lag		5.25				1.5			
median lag		3				1			

- **Parsimonious
long lags and
moving average
lags.**

PARSIMONIUS LONG LAGS

Aaron and Muelbauer

- Suppose w_t is ΔX_t .
- The restrictions in the lags are obtained making use of the fact that:
- $\Delta_m = (1-L_m) = (1-L)(1+L+L^2+L^3+\dots+L^{m-1})$.
- So $\alpha \Delta_m X_{t-1} = \alpha(1+L+L^2+\dots+L^{m-1})(1-L) X_{t-1} = \alpha[w_{t-1} + w_{t-2} + \dots + w_{t-m}]$.
Therefore with just one coefficient α we capture the lags 1 to m , restricted to have the same coefficient.

- Then **the general restricted AR structure** from which it could be of interest to select a more parsimonious AR model will be one with the following lags:

1,2,3,12 on ΔX_t [generating the regressors ΔX_{t-1} , ΔX_{t-2} , ΔX_{t-3} and ΔX_{t-12} .

All these lags have free coefficients

4-6 on $\Delta_3 X_{t-4}$ [generating the regressors ΔX_{t-4} , ΔX_{t-5} and ΔX_{t-6}].

These three lags are restricted to have the same coefficient.

7-11 on $\Delta_5 X_{t-7}$ [generating the regressors ΔX_{t-7} to ΔX_{t-11}]

These five lags are restricted to have the same coefficient.

In cases could also be interested to include

the lag 13 on $\Delta_{12} X_{t-13}$.

Try all possible combinations and select the model with less AIC .

At least try with lags 1,2,3,12 on ΔX_t .

PLL when the stationary variable is

$$\Delta\Delta_{12}X_t$$

- The same procedure as before applies, but **the restricted general model includes lags on**
 $\Delta\Delta_{12}X_t$ (1,2,3, and 12),
- $\Delta_3\Delta_{12}X_{t-4}$ (4-6),
- $\Delta_5\Delta_{12}X_{t-7}$ (7-11).

MOVING AVERAGE LAGS

- In an invertible MA(1) model its AR representation has coefficients declining to zero as an exponential function of the structural MA parameter.
- Very frequently the MA parameter, say b , is between minus 0.8 and minus 0.4 and between 0.4 and 0.8. Then the first four lags in the AR formulation for the first case will have the values b , b^2 , b^3 and b^4 , b^l , where l is the order of the lag. In the second case the values of the lags will be given by $(-1)^l b^l$.

MOVING AVERAGE LAGS RESTRICTIONS

- We can apply the restrictions $\sum_{i=1}^l b_i = 1$ and $(-1)^i b_i$ from lag one to lag l with values for b as: -0.8, -0.7, -0.6, -0.5, -0.4, 0.4, 0.5, 0.6, 0.7 and 0.8, and make the selection by Autometrics.
- With monthly or quarterly data it could also be interesting to try restrictions for lags s and $2s$, where s is the number of observations per year.

FROM GENERAL TO SPECIFIC

- The starting point should be a very general model that may be reduced by a sequence of inference procedures until a more specific, reasonably parsimonious and readily interpretable formulation is obtained.
- This reduction process includes:
 - testing on the significance of the parameters of the model,
 - selecting among competing models and
 - carrying out careful diagnostic to ensure that the error term is white noise.

- Automatic model selection from a general-to-specific methodology can be implemented through a computer program called Autometrics.

Four types of statistics to help evaluate the quality of the estimated model.

- First, those concerning **the estimation results**, such as the estimates of the parameters themselves, their standard deviations and the t -values associated with them.
- Second, some summary statistics for comparison between **alternative fitted models**, such as the R^2 coefficient, the residual variance and some information criteria like the AIC and the SIC.
- Third, some diagnostic tests for possible **mis-specification**, such as those concerning the hypothesis of white noise, homoscedasticity and normality of the residuals.
- Finally, some statistics to test if the model is **stable along time**.