Distributed Lag Models

Ass. Prof. Andriy Stavytskyy

References

- **Greene W. H.** (2011) Econometric Analysis, 7th ed., Prentice Hall.
- **Murray M. P.** (2005) Econometrics: A Modern Introduction. Prentice Hall.
- **Stock James H., Mark W. Watson** (2010) Introduction to Econometrics (3rd Edition)

Agenda

- Lagged variables
- ARDL Models
- Estimation with Exogenous Regressors
- Forms of lags
- Koyck transformation
- Estimation of Distributed Lag Models

LAGGED VARIABLES

Lagged Variables

- A possible source of any problem with the functional form is the ***lack of a lagged structure*** in the model.
- One way of overcoming autocorrelation is to ***add a lagged dependent variable*** to the model.
- However, although lagged variables can produce a better functional form, we need ***theoretical reasons*** for including them.

Why lags are useful

- **<u>Psychological reasons:</u>** behavior is habit-forming
  - *so things like labor market behavior and patterns of money holding can be captured using lags*

- **<u>Technological reasons:</u>** a firm's production pattern

- **<u>Institutional:</u>** unions

- **<u>Multipliers:</u>** short run and long run multipliers (how to read finite distributed lags in a model).

Inclusion of Lagged variables

- Inertia of the dependent variable, whereby a change in an explanatory variable ***does not immediately effect*** the dependent variable.
- The overreaction to 'news', particularly common in asset markets and often referred to as 'overshooting', where the asset 'overshoots' its long-run equilibrium position, before moving back towards equilibrium
- To allow the model to produce ***dynamic forecasts*****.**

Types of models

- If the regression model includes not only the current but also the lagged (past) values of the explanatory variables (the X's) it is called a **distributed-lag model**.
- If the model includes one or more lagged values of the dependent variable among its explanatory variables, it is called an **autoregressive model**. This model is known as a **dynamic model**.

Solution

?

Output Y(t) ← Dynamic Relationship ← Input X(t)

Possible solutions

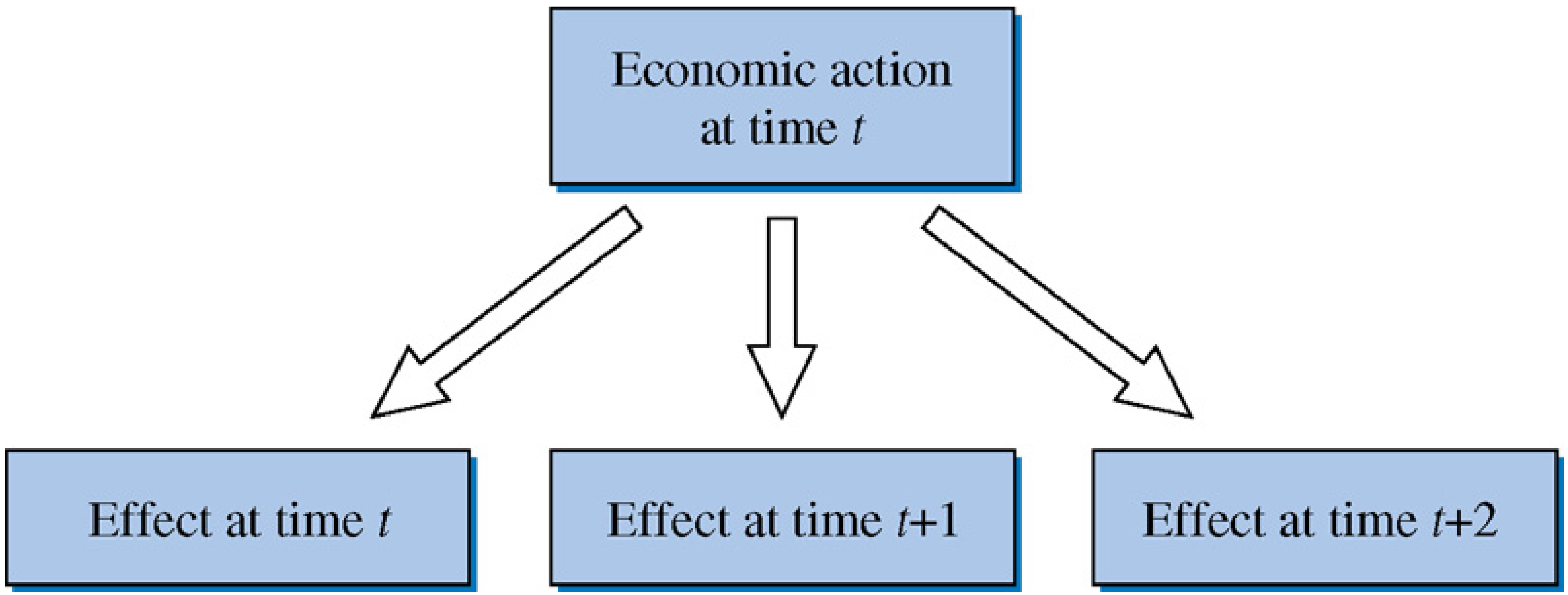
$$y_t = f(x_t, x_{t-1}, x_{t-2}, \ldots)$$

$$y_t = f(y_{t-1}, x_t)$$

$$y_t = f(x_t) + e_t$$

$$e_t = f(e_{t-1})$$

Why should we use lags?



Examples of dynamic causal effects

A **dynamic causal effect** is the effect on Y of a change in X over time.

- The effect of an increase in cigarette taxes on cigarette consumption this year, next year, in 5 years;
- The effect of a change in the Fed Funds rate on inflation, this month, in 6 months, and 1 year;
- The effect of a freeze in Florida on the price of orange juice concentrate in 1 month, 2 months, 3 months…

Distributed Lag – 1

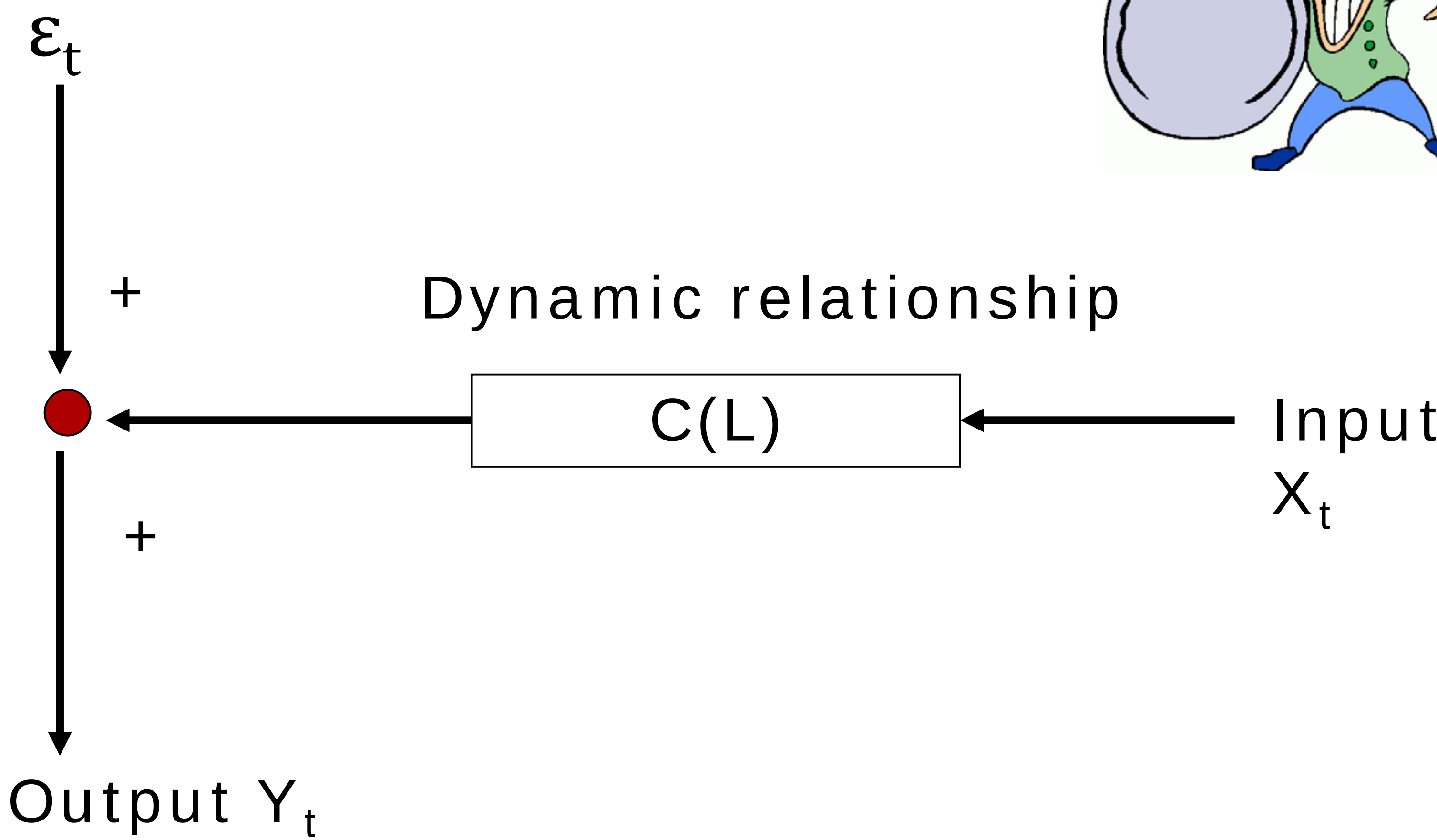
$$Y_t = c_0 x_t + c_1 x_{t-1} + c_2 x_{t-2} + \ldots + \varepsilon_t$$

$$Y_t = c_0 x_t + c_1 L x_t + c_2 L^2 x_t + \ldots + \varepsilon_t$$

$$Y_t = \left(c_0 + c_1 L + c_2 L^2 + \ldots\right) x_t + \varepsilon_t$$

$$Y_t = \left(\sum_{i=0}^{k} c_i L^i\right) x_t + \varepsilon_t$$

Distributed Lag – 2



Identification of C(L)

- How many lags?
- Which lags?
- What are the statistical estimation problems can appear?

Functional Form – 1

- An important assumption we make about the econometric model is that ***it has the correct functional form.***
- This requires the ***most appropriate variables*** in the model and that they are in the ***most suitable format***, i.e. logarithms etc.
- One of the most important considerations with financial data is that we need to model the dynamics appropriately, with the ***most appropriate lag structure.***

Functional Form – 2

- It is important we include all the relevant variables in the model, if we exclude an important explanatory variable, the regression has 'omitted variable bias'. This means the estimates are unreliable and the **t** and **F statistics** can not be relied on.

- Equally there can be a problem if we include variables that are not relevant, as this **can reduce the efficiency** of the regression, however this is not as serious as the omitted variable bias.

- The **Ramsey Reset test** can be used to determine if the functional form of a model is acceptable.

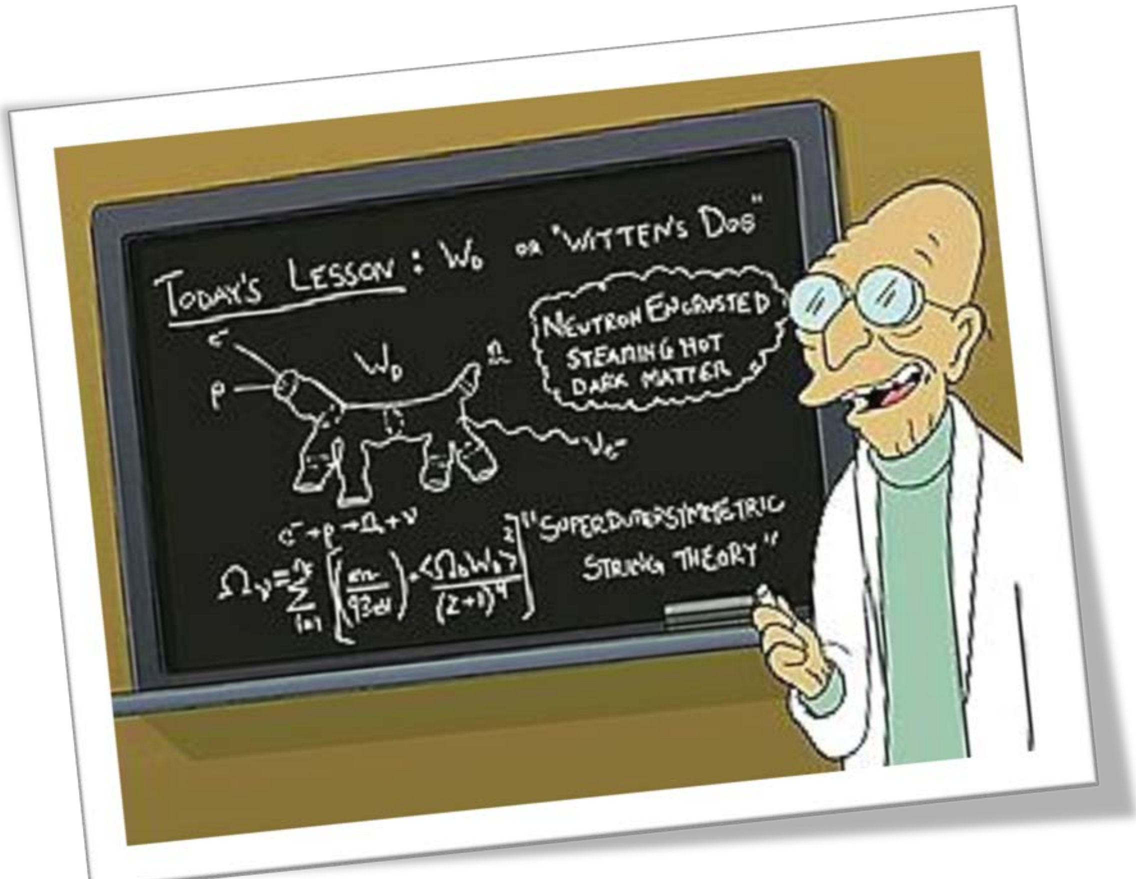
Ramsey Reset Test for Functional Form

- This test is based on **running the regression** and saving the residual as well as the fitted values.
- Then run a secondary regression of the residual on powers of these fitted values.

$$y_t = \alpha + \beta x_t + u_t$$

$$\hat{u}_t = \delta_0 + \delta_1 \hat{y}_t^2 + \delta_2 \hat{y}_t^3 + \ldots + \delta_{p-1} \hat{y}_t^p + v_t$$

ARDL MODELS


TODAY'S LESSON : W₀ OR "WITTEN'S DOG"
NEUTRON ENCRUSTED STEAMING HOT DARK MATTER
"SUPERDUPERSYMMETRIC STRING THEORY"

ARDL Models

- An **Autoregressive Distributed lag model** or ARDL model refers to a model with lags of both the dependent and explanatory variables.
- An ARDL(1,1) model would have 1 lag on both variables:

$$y_t = \alpha_0 + \alpha_1 x_t + \alpha_2 x_{t-1} + \alpha_3 y_{t-1} + u_t$$

Role of "Time" or "lag" in Economics

Distributed Lag Model

$$y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \ldots + \beta_K X_{t-K} + \mu_t$$

β_0 = the short-run or impact multiplier

$(\beta_0 + \beta_1)$ *or* $(\beta_0 + \beta_1 + \beta_2)$ are examples of interim or intermediate multipliers.

$$\sum_{i=0}^{K} \beta_i = \text{long-run or total distributed lag multiplier.}$$

$$\beta_i^* = \frac{\beta_i}{\sum_{i=0}^{K} \beta_i} = \text{standardized coefficient.}$$

Share of total impact.

Example: dynamic causal effects

The **distributed lag model** is:

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 X_{t-2} + u_t$$

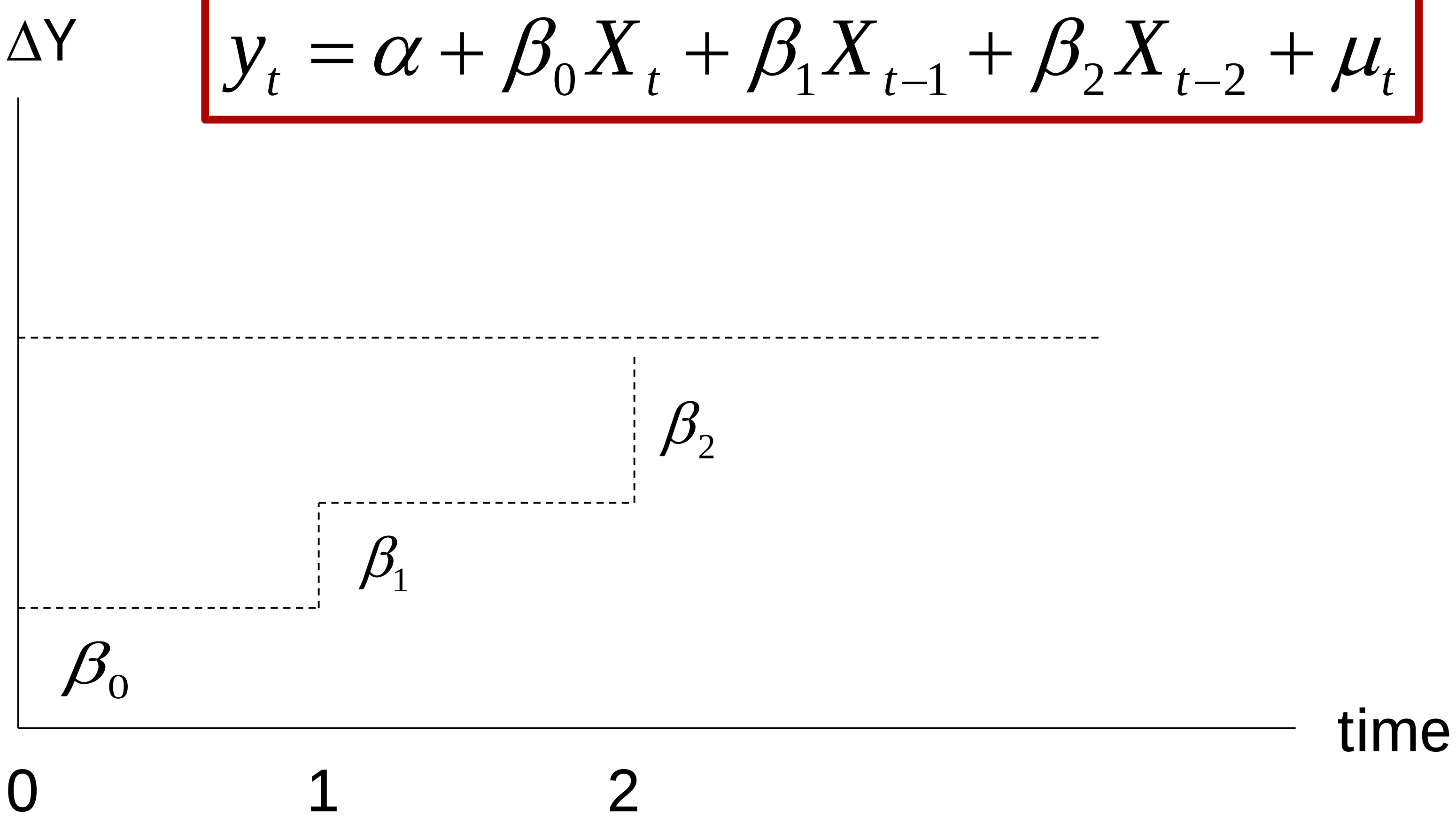
- β_1 = ***impact effect of change in X*** = effect of change in X_t on Y_t, holding past X_t constant
- β_2 = ***1-period dynamic multiplier*** = effect of change in X_{t-1} on Y_t, holding constant X_t, X_{t-2}, X_{t-3},…
- β_3 = ***2-period dynamic multiplier*** (etc.)= effect of change in X_{t-2} on Y_t, holding constant X_t, X_{t-1}, X_{t-3},…
- ***Cumulative dynamic multipliers:***

the 2-period cumulative dynamic multiplier $= \beta_1 + \beta_2 + \beta_3$

Demonstration of distributed Lag

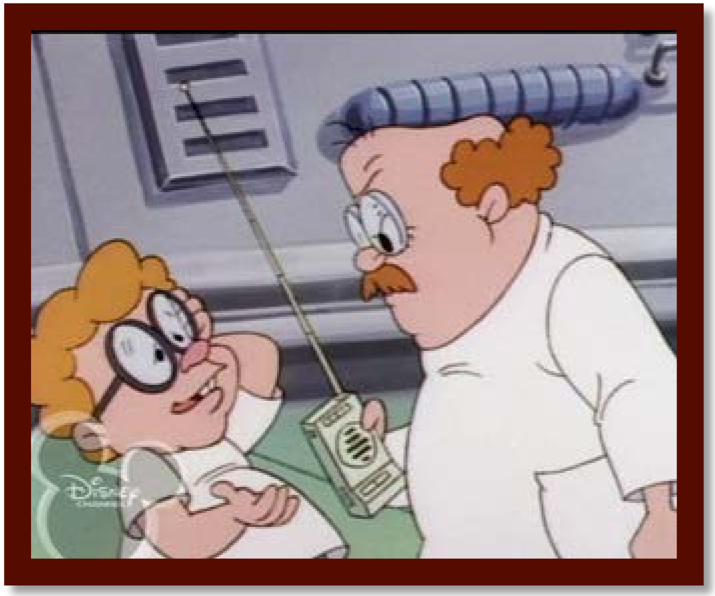
Effect of 1 unit sustained increase in X

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



Example

- The Index of Consumer Sentiment
- Standard & Poor's 500 Index
- Does consumer sentiment affect the stock market?



Regression of SP500 on Consen

Dependent Variable: SP500
Method: Least Squares

Sample(adjusted): 1978:04 2003:03
Included observations: 300 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4280.864	412.7262	-10.37216	0.0000
CONSEN	19.22990	15.94315	1.206155	0.2287
CONSEN(-1)	4.430760	21.88401	0.202466	0.8397
CONSEN(-2)	4.567627	21.83672	0.209172	0.8345
CONSEN(-3)	36.02170	15.91065	2.264000	0.0243

R-squared	0.393625	Mean dependent var	1361.881
Adjusted R-squared	0.385403	S.D. dependent var	1287.904
S.E. of regression	1009.669	Akaike info criterion	16.68916
Sum squared resid	3.01E+08	Schwarz criterion	16.75089
Log likelihood	-2498.374	F-statistic	47.87436
Durbin-Watson stat	0.031236	Prob(F-statistic)	0.000000

ESTIMATION WITH EXOGENOUS REGRESSORS

Exogeneity in time series regression

- **Exogeneity** (*past and present*)

X is ***exogenous*** if $E(u_t|X_t, X_{t-1}, X_{t-2}, \ldots) = 0$.

- **Strict Exogeneity** (*past, present, and future*)

X is ***strictly exogenous*** if $E(u_t|\ldots, X_{t+1}, X_t, X_{t-1}, \ldots) = 0$

- Strict exogeneity implies exogeneity
- For now we suppose that X is exogenous – we'll return (briefly) to the case of strict exogeneity later.
- If X is exogenous, then we can use OLS to estimate the dynamic causal effect on Y of a change in X….

The Distributed Lag Model: Assumptions

1. $E(u_t|X_t, X_{t-1}, X_{t-2}, \ldots) = 0$ (X is exogenous)
2. (a) Y and X have stationary distributions;
   (b) (Y_t, X_t) and (Y_{t-j}, X_{t-j}) become independent as j gets large
3. Y and X have eight nonzero finite moments
4. There is no perfect multicollinearity.

Under the Distributed Lag Model Assumptions

- OLS yields **consistent*** estimators of $\beta_1, \beta_2, \ldots, \beta_r$ (of the dynamic multipliers) (**consistent but possibly biased!*)
- The sampling distribution of β's is normal
- ***BUT*** the formula for the variance of this sampling distribution is not the usual one from cross-sectional (i.i.d.) data, because u_t is not i.i.d. – u_t can be serially correlated!
- This means that the usual OLS standard errors are wrong!
- We need to use, instead, *SE*s that are robust to autocorrelation as well as to heteroskedasticity…

Heteroskedasticity and Autocorrelation-Consistent (HAC) Standard Errors

- When u_t is serially correlated, the variance of the sampling distribution of the OLS estimator is different.
- Consequently, we need to use a different formula for the standard errors.

HAC SEs

$$\text{var}(\hat{\beta}_1) = \left[\frac{1}{T}\frac{\sigma_v^2}{(\sigma_X^2)^2}\right] \times f_T, \text{ where } f_T = 1 + 2\sum_{j=1}^{T-1}\left(\frac{T-j}{T}\right)\rho_j$$

The most commonly used estimator of f_T is:

$$\hat{f}_T = 1 + 2\sum_{j=1}^{m-1}\left(\frac{m-j}{m}\right)\tilde{\rho}_j \quad \text{(Newey-West)}$$

- $\tilde{\rho}_j$ is an estimator of ρ_j
- This is the "Newey-West" HAC *SE* estimator
- *m* is called the ***truncation parameter***
- Why not just set $m = T$?
- Then how should you choose *m*?
    - Use the Goldilocks method
    - Or, use the rule of thumb, $m = 0.75T^{1/3}$

Estimation with Strictly Exogenous Regressors

- *Let* X is strictly exogenous if $E(u_t|\ldots, X_{t+1,} X_t, X_{t-1}, \ldots) = 0$
- If X is strictly exogenous, there are more efficient ways to estimate dynamic causal effects than by a distributed lag regression:
  - Generalized Least Squares (GLS) estimation
  - Autoregressive Distributed Lag (ADL) estimation
- But the condition of strict exogeneity is very strong, so this condition is rarely plausible in practice.
- So we won't cover GLS or ADL estimation of dynamic causal effects.

Computation of cumulative multipliers – 1

- The **cumulative multipliers** can be computed by estimating the distributed lag model, then adding up the coefficients.
- However, you should also compute **standard errors** for the cumulative multipliers and while this can be done directly from the distributed lag model it requires some modifications.
- One easy way to compute cumulative multipliers and standard errors of cumulative multipliers is to realize that cumulative multipliers are **linear combinations of regression coefficients**.

Computation of cumulative multipliers – 2

Rewrite the distributed lag model with 1 lag:

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

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

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

or

$$Y_t = \beta_0 + \beta_1 \Delta X_t + (\beta_1 + \beta_2)\, X_{t-1} + u_t$$

Computation of cumulative multipliers – 3

- So, let $W_{1t} = \Delta X_t$ and $W_{2t} = X_{t-1}$ and estimate the regression,

$$Y_t = \beta_0 + \delta_1 W_{1t} + \delta_2 W_{2t} + u_i$$

- Then
- $\delta_1 = \beta_1$ = impact effect
- $\delta_2 = \beta_1 + \beta_2$ = the first cumulative multiplier
- the (HAC) standard errors on δ_1 and δ_2 are the standard errors for the two cumulative multipliers.

Computation of cumulative multipliers – 4

- In general, the ADL model can be rewritten as

$$Y_t = \delta_0 + \delta_1 \Delta X_t + \delta_2 \Delta X_{t-1} + \ldots + \delta_{q-1} \Delta X_{t-q+1} + \delta_q X_{t-q} + u_t$$

where

$$\delta_1 = \beta_1$$
$$\delta_2 = \beta_1 + \beta_2$$
$$\delta_3 = \beta_1 + \beta_2 + \beta_3$$
$$\ldots$$
$$\delta_q = \beta_1 + \beta_2 + \ldots + \beta_q$$

- Cumulative multipliers and their HAC SEs can be computed directly using this transformed regression

FORMS OF LAGS

Unrestricted lags (no structure)

- It is always finite!

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \ldots + \beta_n x_{t-n} + e_t$$

  - *n lags and no structure in parameters*
  - *OLS works*
- but
  - *n observations lost*
  - *high multicollinearity*
  - *imprecise, large s.e., low t, lots of d.f. lost*

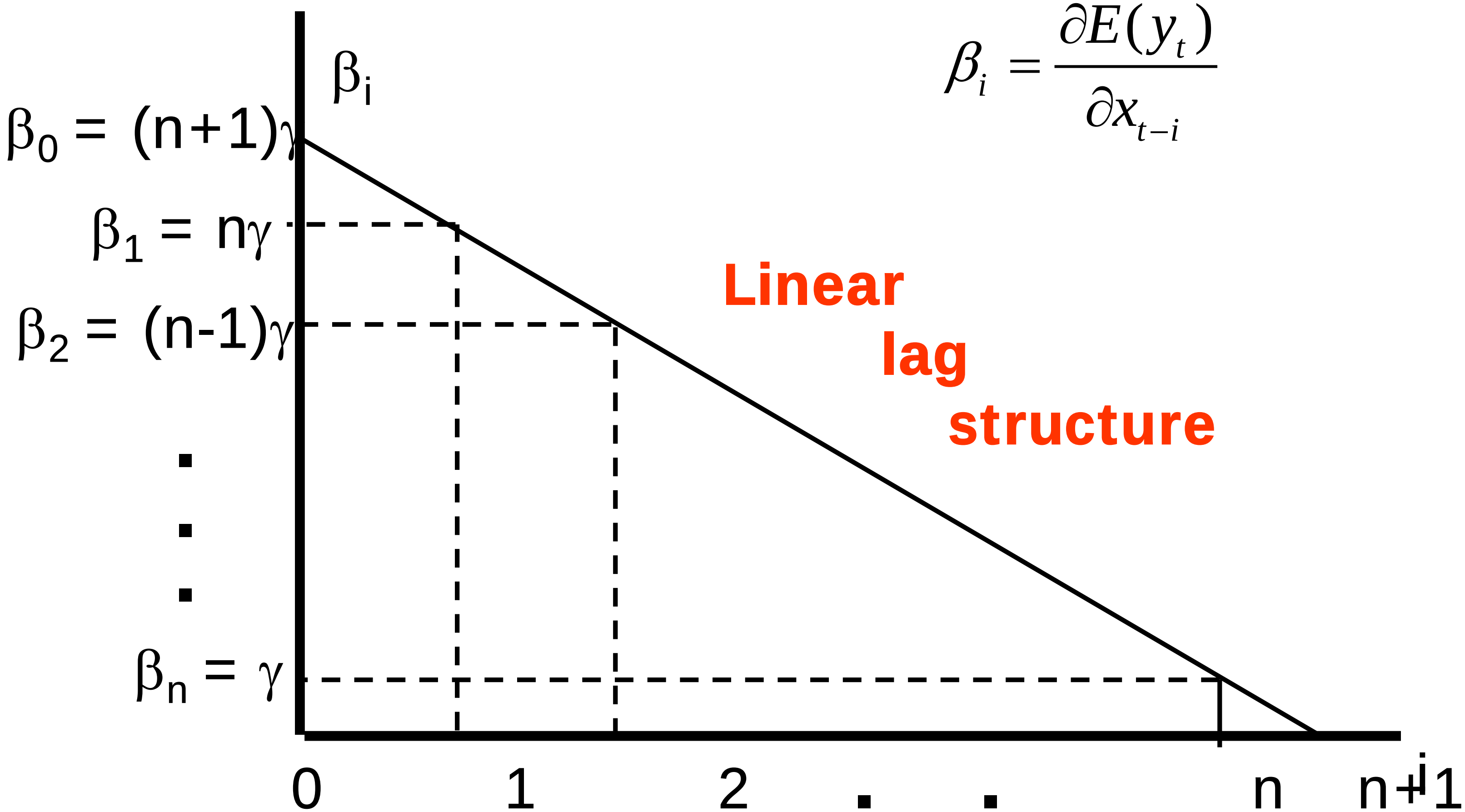
- Structure could help

Arithmetic lag

- The effect of X eventually zero
- Linearly!
- The coefficients not independent of each other
  - *effect of each lag less than previous*
  - *exactly like arithmetic series:* $u_n=u_1+d*(n-1)$

Arithmetic lag - structure

$$\beta_i = \frac{\partial E(y_t)}{\partial x_{t-i}}$$



Arithmetic lag – pros & cons

<u>Advantages</u>:

- Only one parameter to be estimated!
  - t-statistics are ok., better s.e., results more reliable
- Straightforward interpretation

<u>Disadvantages</u>:

- If restriction untrue, estimators biased and inconsistent

<u>Solution? F-test!</u>

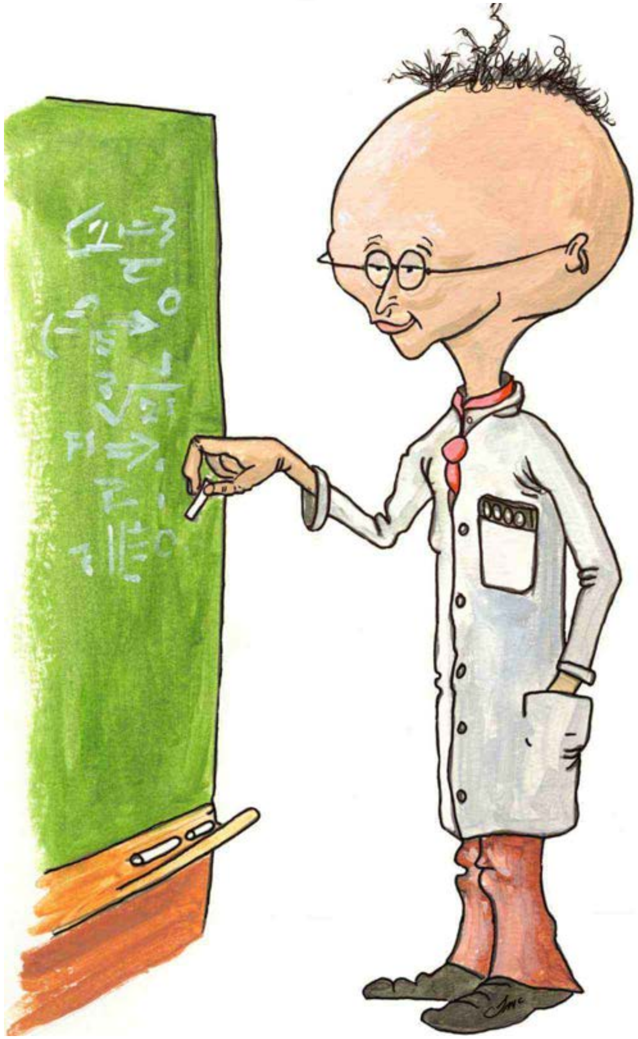
F-tests of restrictions

- Estimate the unrestricted model
- Estimate the restricted (any lag) model
- Calculate the test statistic

$$F = \frac{(RSS_R - RSS_U)/df_1}{RSS_U/df_2}$$

- Compare with critical value F(df_1,df_2)
    - df1 = number of restrictions
    - df2 = number of observations-number of variables in the unrestricted model (incl. constant)
- H_0: residuals are 'the same', restricted model OK

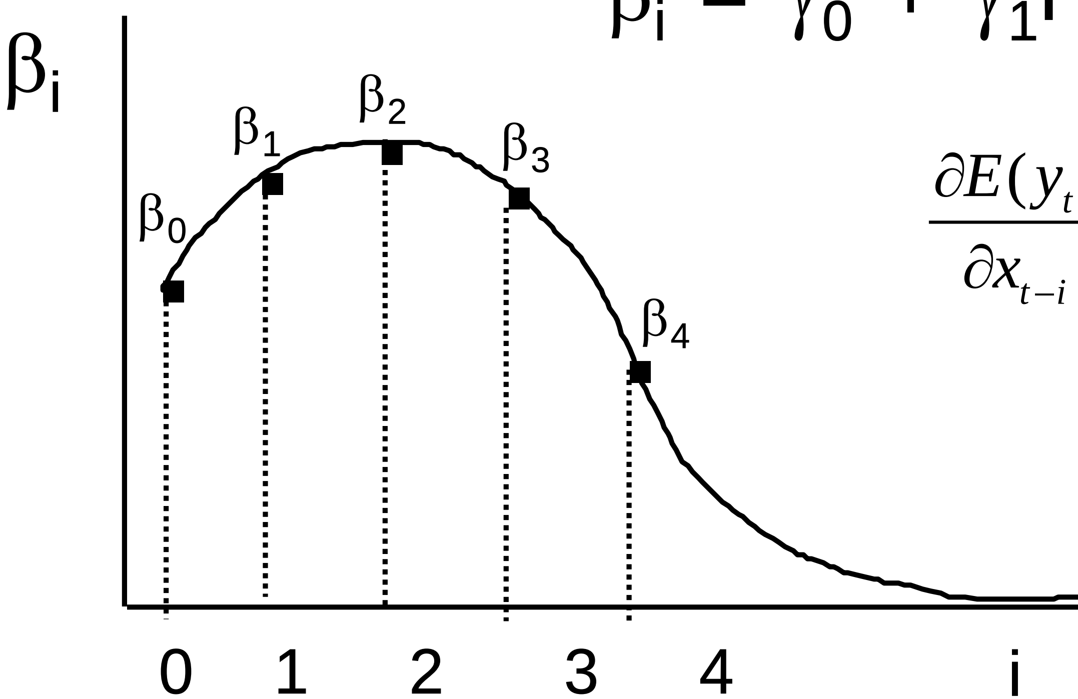
Polynomial lag



- a non-linear shape
- finite: the effect eventually goes to zero (by DEFINITION and not by nature!)
- The coefficients still related to each other, BUT not a uniform pattern (decline)

Polynomial lag: structure

$$\beta_i = \gamma_0 + \gamma_1 i + \gamma_2 i^2$$



$$\frac{\partial E(y_t)}{\partial x_{t-i}} = \beta_i$$

Polynomial lag: maths

- n – the lenght of the lag
- p – degree of the polynomial

$$\beta_i = \gamma_0 + \gamma_1 i + \gamma_2 i^2 + \ldots + \gamma_p i^p,$$

where i=1, . . . , n

- For example a quadratic polynomial

$$\beta_i = \gamma_0 + \gamma_1 i + \gamma_2 i^2 \text{, where } p=2 \text{ and } n=4$$

$\beta_0 = \gamma_0$ $\beta_1 = \gamma_0 + \gamma_1 + \gamma_2$

$\beta_2 = \gamma_0 + 2\gamma_1 + 4\gamma_2$ $\beta_3 = \gamma_0 + 3\gamma_1 + 9\gamma_2$

$\beta_4 = \gamma_0 + 4\gamma_1 + 16\gamma_2$

Polynomial lag – pros & cons

Advantages:

- Fewer parameters to be estimated than in the unrestricted lag structure that leads to more precise than unrestricted estimates
- If the polynomial restriction likely to be true, it is more flexible than arithmetic DL

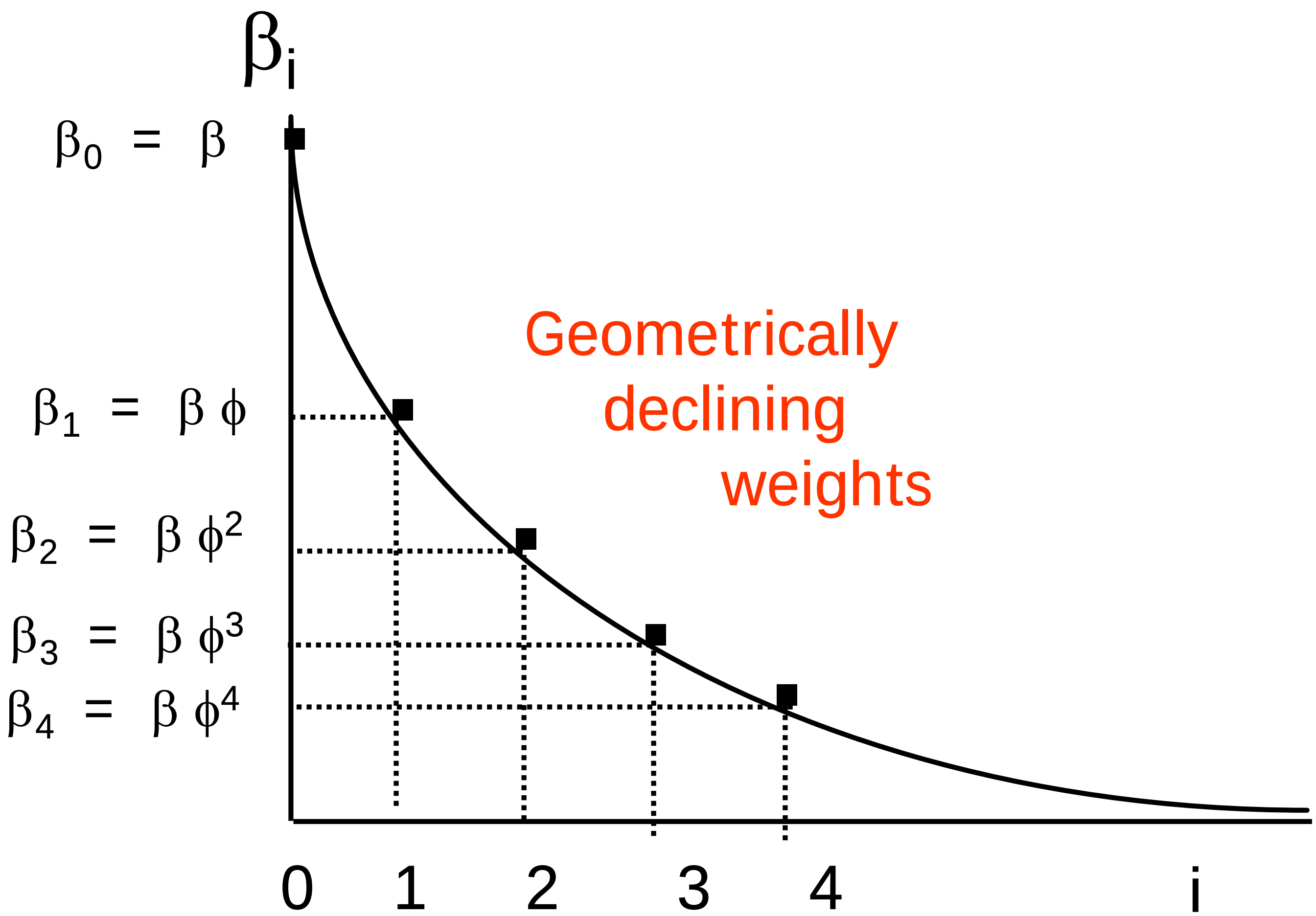
Disadvantages

- If the restriction untrue, estimates are biased and inconsistent

Geometric lag (Koyck)

- Distributed lag is infinite $\Rightarrow$ infinite lag length (no time limits)
- BUT cannot estimate an infinite number of parameters!
- We need to restrict the lag coefficients to follow a pattern
- For the geometric lag the pattern is one of continuous decline at decreasing rate

Geometric lag: structure



Geometric lag: maths

- Infinite distributed lag model

  $y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \ldots + e_t$

  $y_t = \alpha + \Sigma \beta_i x_{t-i} + e_t$

- Geometric lag structure

  $\beta_i = \beta \phi^i$ where $|\phi| < 1$ and $\beta\phi^i > 0$

- Infinite unstructured geometric lag model

  $y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \ldots + e_t$

  AND: $\beta_0 = \beta$, $\beta_1 = \beta\phi$, $\beta_2 = \beta\phi^2$, $\beta_3 = \beta\phi^3$ ...

- Substitute $\beta_i = \beta \phi^i$ => infinite geometric lag

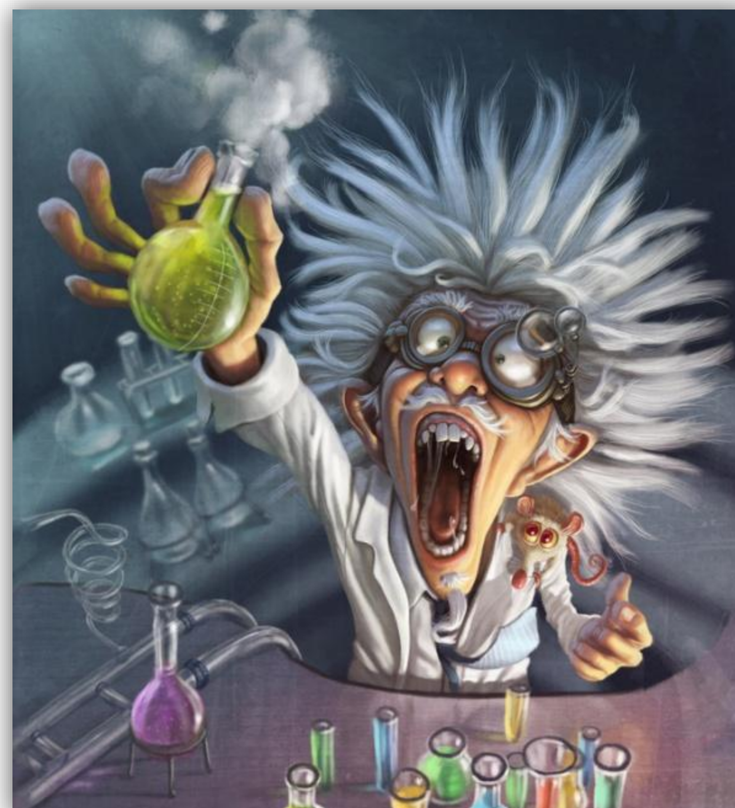
  $y_t = \alpha + \beta(x_t + \phi x_{t-1} + \phi^2 x_{t-2} + \phi^3 x_{t-3} + \ldots) + e_t$

Geometric lag: problems

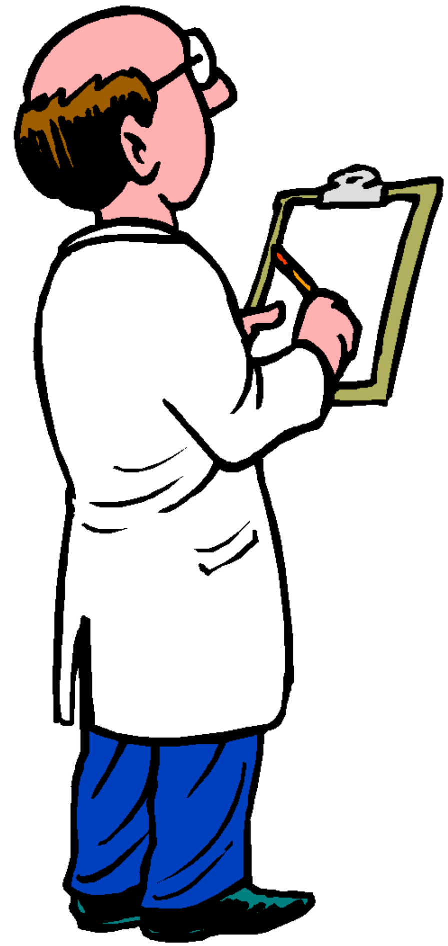
- Cannot estimate using OLS (y_{t-1} is dependent on e_{t-1} $\Rightarrow$ cannot alow that)
- Need Koyck transformation
- Have to do some algebra to rewrite the model in form that can be estimated.

Geometric lag: pros & cons

- **<u>Advantages</u>**
  - You only estimate two parameters ϕ and β!
- **<u>Disadvantages</u>**
  - We allow neither for heterogenous nor for unsmooth declining



KOYCK TRANSFORMATION



Koyck transformation – 1

- Model:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \ldots + \beta_k X_{t-k} + \varepsilon_t$$

- The Koyck transformation suggests that the further back in time we go, the less important is that factor (for instance, information from 10 years ago vs. information from last year)
- The transformation suggests:

$\beta_j = \beta_0 \lambda^j$, where $0 < \lambda < 1$, $j = 1,\ldots k$

Koyck transformation – 2

- So,

$$\beta_1 = \beta_0 \lambda \ , \ \beta_2 = \beta_0 \lambda^2 \ , ...$$

- Can use the expression for β_j to rewrite the model

$$Y_t = \alpha + \beta_0(X_t + \lambda X_{t-1} + \lambda^2 X_{t-2} + \ldots + \lambda^k X_{t-k}) + \varepsilon_t$$

- This imposes the assumption that earlier information is relatively less important

- Lagging the equation and multiplying it by λ, we get:

$$\lambda Y_{t-1} = \lambda\alpha + \beta_0(\lambda X_{t-1} + \lambda^2 X_{t-2} + \lambda^3 X_{t-3} + \ldots + \lambda^k X_{t-k}) + \lambda\varepsilon_{t-1}$$

- Subtracting, we get

$$Y_t = \alpha(1-\lambda) + \beta_0 X_t + \lambda Y_{t-1} + v_t, \quad \text{where } v_t = \varepsilon_t - \lambda\, \varepsilon_{t-1}$$

Koyck transformation – 3

- **Why is this transformation useful?**
  - Allows us to take the ad-hoc lag series and condense it into a lagged endogenous variable
  - now we only lose one observation due to the lagged endogenous variable
  - the λ given by the estimation gives the coefficient of autocorrelation
- **Problem: by construction, we have first-order autocorrelation**
  - use Durbin h-statistic
  - but estimating equation might be mis-specified!

Testing autoregressive model for autocorrelation – 1

If we have the model,

$$y_t = \beta_1 + \beta_2 X_t + \beta_3 y_{t-1} + \mu_t$$

We test for autocorrelation with the Durbin h-statistic

$$h = \hat{\rho}\sqrt{\frac{n}{1\text{-}n\left(\mathrm{var}\left(\hat{\beta}_3\right)\right)}}$$

Testing autoregressive model for autocorrelation – 2

If we estimate ρ, the autocorrelation coefficient as

$\hat{\rho} = 1 - \frac{d}{2}$, where d is the traditional Durbin-Watson statistic. The Durbin h is now

$$h \cong (1 - \frac{d}{2})\sqrt{\frac{n}{1 - n\left(\hat{var}\left(\hat{\beta}_3\right)\right)}} \sim N(0,1)$$

Note h does not exist when $n[\hat{var}(b_3)] > 1$.

Economic rational for Koyck model

- Adaptive Expectations
- Partial Adjustment

Adaptive Expectation Model – 1

Basic model : $y_t = \alpha + \beta X_t^e + \mu_t$

Adjustment process

$$\left(X_t^e - X_{t-1}^e \right) = \theta (X_t - X_{t-1}^e) \; or$$

$X_t^e = \theta X_t + (1-\theta) X_{t-1}^e$, if we lag and substitute

$$X_t^e = \theta X_t + (1-\theta)\theta X_{t-1} + (1-\theta)^2 X_{t-2}^e$$

With succesive lagging and substituting, we get

$$X_t^e = \theta X_t + (1-\theta)\theta X_{t-1} + (1-\theta)^2 \theta X_{t-2} + ...$$

Adaptive Expectation Model – 2

Substituting back into the basic model, we get

$$\mathrm{y_t} = \alpha + \beta\theta X_t + \beta\theta(1-\theta)X_{t-1} + \beta\theta(1-\theta)^2 X_{t-2} + \ldots + \mu_t$$

This is the Koyck model with $\beta_0 = \beta\theta \; and \; \lambda = (1-\theta)$

Our estimating equation therefore is

$$\mathrm{y_t} = \alpha\theta + \beta\theta X_t + (1-\theta)y_{t-1} + (\mu_t - (1-\theta)\mu_{t-1})$$

Adaptive Expectation Model – 3

- Expected value of the independent variable is **weighted average** of the present and all past values of X.
- The estimating equation has a **MA(1)** process error term.

Example – 1

- In 1968, M. Friedman estimated the equation:

$$Y_t = \alpha + \beta X^*_t + u_t$$

where X^*_t - natural rate of unemployment.

- He tried to measure the natural rate of unemployment

Example – 2

- ***Using adaptive expectations we have that***

$$X^*_t - X^*_{t-1} = \delta\,(X_t - X^*_{t-1})$$

where

$X_t{}^*$ - expectation

X_t - observed

$0 < \delta < 1$

- ***Can rewrite the equation***:

$$X^*_t - (1 - \delta\,)X^*_{t-1} = \delta\,X_t$$

- ***Using a lag operator with***:

$LX_t = X_{t-1}$

$L^2X_t = X_{t-2}$

Example – 3

- We can then rewrite

$$\delta X_t = (1 - \gamma L)X^*_t, \text{ where } \gamma = (1- \delta)$$

- This can be rewritten as:

$$X_t^* = \frac{\delta}{(1-\gamma L)} X_t$$

- It is the natural rate of unemployment in terms of the observed rate of unemployment

Example – 4

- Substituting into the model we get:

$$Y_t = a + b\left[\frac{\delta}{(1-\gamma L)}\right]X_t + u_t$$

- Upon further multiplication and substitution we arrive at:

$$Y_t = a\delta + b\delta X_t + (1-\delta)Y_{t-1} + v_t$$

where

$$v_t = u_t - (1-\delta)u_{t-1}$$

- This looks very similar to that for the Koyck transformation

Partial Adjustment model – 1

Let y_t^d – the desired level of Y in period t and

is a function of X_t, *i.e.*

$$y_t^d = \alpha + \beta X_t + \mu_t ,$$

one adjusts the actual level of y

according to the adjustment process

$$(y_t - y_{t-1}) = \theta(y_t^d - y_{t-1}),\ 0 < \theta < 1$$

substituting the first equation into the second, we get

$$y_t = \alpha\theta + \beta\theta X_t + (1-\theta) y_{t-1} + \theta\mu_t$$

Partial Adjustment model – 2

- Estimating equation looks like Koyck but is different as far as estimation is concerned
- Error term is well behaved
- In the limit the lagged dependent variable is uncorrelated with the error term
- model can be estimated consistently by OLS

Problems with the approaches

- For the lagged endogenous variables in the ad-hoc lag structure, we are uncertain as to which economic model of agent behavior underlies the estimating equation
- We have 1st-order autocorrelation by the construction of the model (need to use the Durbin h-statistic)
- Y_{t-1} and ε_{t-1} are sure to be correlated [$E(X,\varepsilon) \neq 0$]
  - *this leads to biased estimates*
  - *we'll deal with this using instrumental variables and simultaneous equations*

ESTIMATION OF DISTRIBUTED LAG MODELS

Estimation of Distributed Lag Models

Infinite Lag

$$y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + ... + \mu_t$$

Not enough data to estimate. Need restrictions

Finite Lag

$$y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + ... + \beta_K X_{t-K} + \mu_t$$

Problems of Ad-hoc Estimation

- No a priori guide to length of lag.
- Longer lags => less degrees of freedom
- Multicollinearity
- Data mining

Estimating Koyck model

- Model can be estimated by **maximum likelihood**. It is difficult.
- Simple method of estimation is **instrumental variables**.

Instrumental Variable Estimation – 1

For each right-hand sided (RHS) variable in our estimating equation, we need a variable Z with the properties that Z is correlated with the RHS variable, but uncorrelated with the error term.

$$\text{Plim}[(Z_t - E(Z_t))(X_t - E(X_t))] \neq 0 \ \textit{and}$$

$$P\lim[(Z_t - E(Z_t))(\mu_t - E(\mu_t))] = 0$$

Instrumental Variable Estimation – 1

For the Koyck model, we may use 1 as instrument for itself and X as instrument for X.

For y_{t-1} we need some other variable as the instrument. Choices included X_{t-1} *and* $\hat{y}_{t-1}$ *where*

$$\hat{y}_{t-1} = d_0 + d_1 X_t + d_2 X_{t-1}$$

Instrumental Variable Estimation – 2

For our Koyck style model, multiple the equation by the instrumental variable and sum across all observations.

We get the following normal equations that must be solved for our parameter estimates.

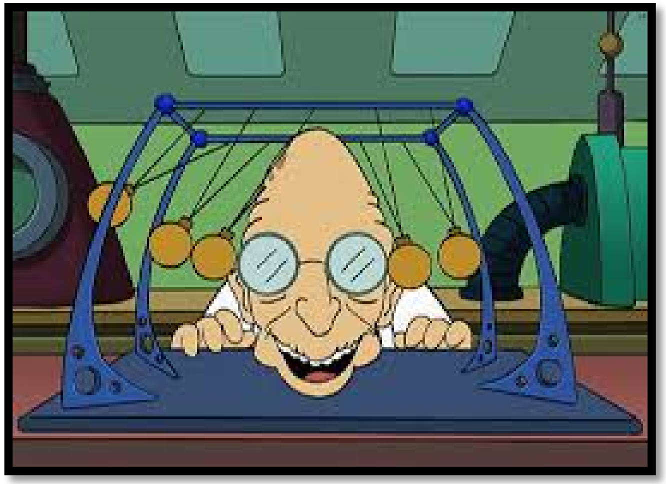
$$\sum \mathrm{y}_{\mathrm{t}} = b_1 n + b_2 \sum X_t + b_3 \sum \mathrm{y}_{\mathrm{t-1}}$$

$$\sum X_t y_t = b_1 \sum X_t + b_2 \sum X_t^2 + b_3 \sum \mathrm{X}_{\mathrm{t}} \mathrm{y}_{\mathrm{t-1}}$$

$$\sum Z_t y_t = b_1 \sum Z_t + b_2 \sum Z_t X_t + b_3 \sum \mathrm{Z}_{\mathrm{t}} \mathrm{y}_{\mathrm{t-1}}$$

Properties of IV estimators

- Estimators are consistent
- Estimators are asymptotically unbiased.
- Parameter estimates will not be as efficient as the maximum likelihood estimates, but are easier to do.



REVIEW

Inclusion of Lagged variables

- Inertia of the dependent variable, whereby a change in an explanatory variable does not immediately effect the dependent variable.
- The overreaction to ‘news’, particularly common in asset markets and often referred to as ‘overshooting’, where the asset ‘overshoots’ its long-run equilibrium position, before moving back towards equilibrium
- To allow the model to produce dynamic forecasts.

Types of models

- If the regression model includes not only the current but also the lagged (past) values of the explanatory variables (the X’s) it is called a **distributed-lag model**.
- If the model includes one or more lagged values of the dependent variable among its explanatory variables, it is called an **autoregressive model**. This model is known as a **dynamic model**.

ARDL Models

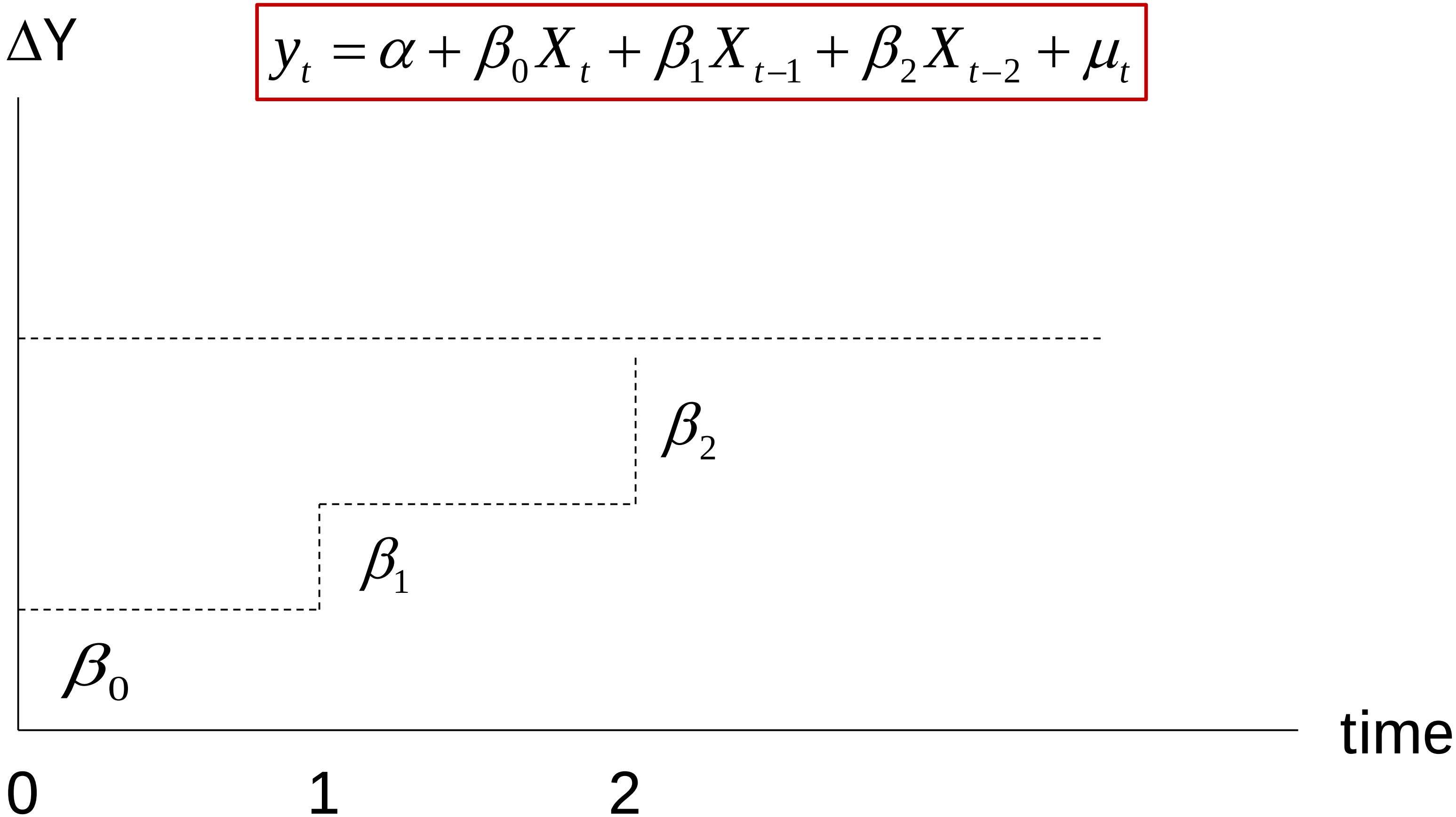
- An Autoregressive Distributed lag model or ARDL model refers to a model with lags of both the dependent and explanatory variables.
- An ARDL(1,1) model would have 1 lag on both variables:

$$y_t = \alpha_0 + \alpha_1 x_t + \alpha_2 x_{t-1} + \alpha_3 y_{t-1} + u_t$$

Demonstration of distributed Lag

Effect of 1 unit sustained increase in X

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



The Distributed Lag Model: Assumptions

1. $E(u_t|X_t,X_{t-1},X_{t-2},\ldots) = 0$ (X is exogenous)
2. (a) Y and X have stationary distributions;

   (b) (Y_t,X_t) and (Y_{t-j},X_{t-j}) become independent as j gets large
3. Y and X have eight nonzero finite moments
4. There is no perfect multicollinearity.

Computation of cumulative multipliers

In general, the ARDL model can be rewritten as

$$Y_t = \delta_0 + \delta_1 \Delta X_t + \delta_2 \Delta X_{t-1} + \ldots + \delta_{q-1} \Delta X_{t-q+2} + \delta_q \Delta X_{t-q+1} + u_t$$

where

$\delta 1 = \beta 1$

$\delta 2 = \beta 1 + \beta 2$

$\delta 3 = \beta 1 + \beta 2 + \beta 3$

…

$\delta q = \beta 1 + \beta 2 + \ldots + \beta q$

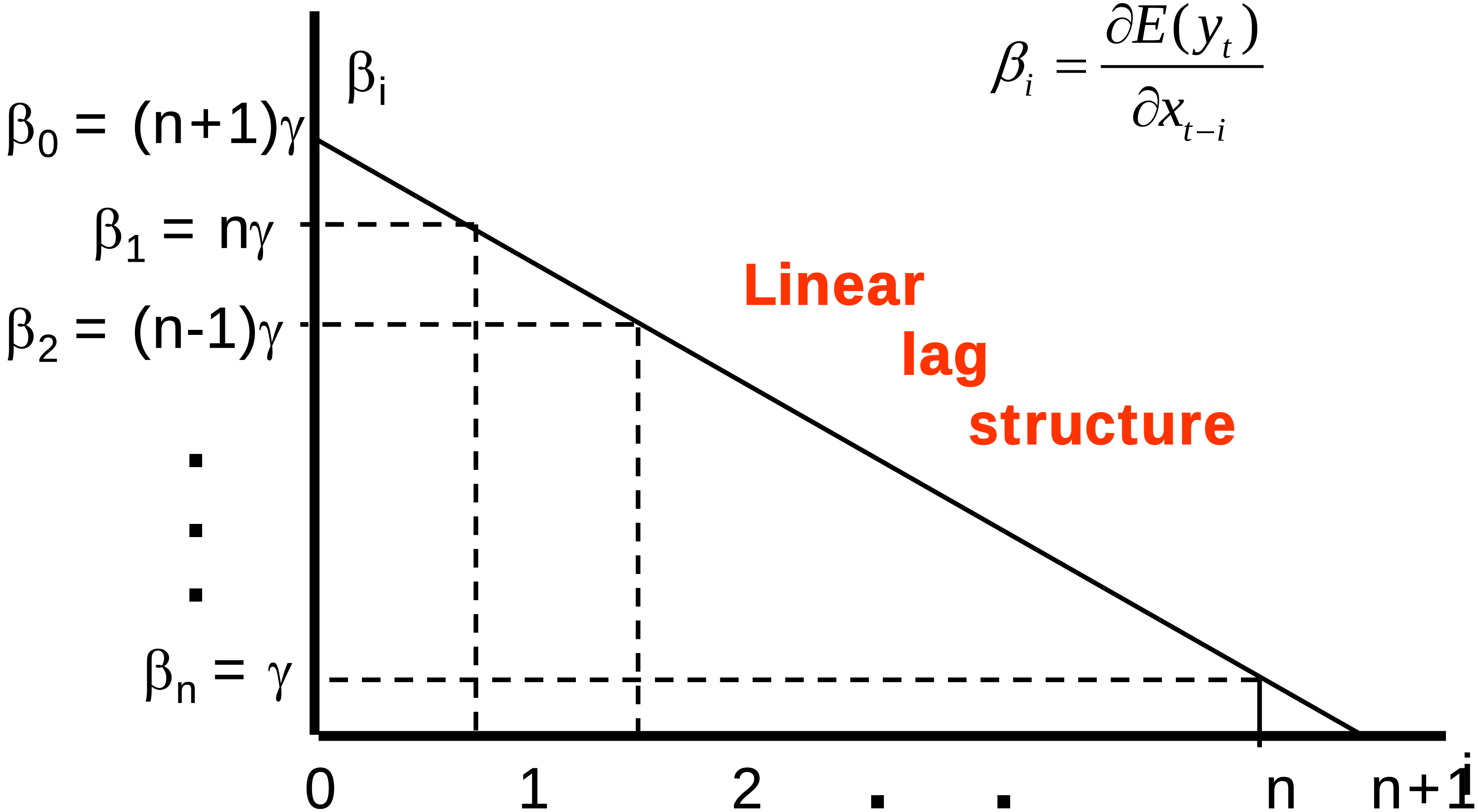
Unrestricted lags (no structure)

- It is always finite!

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \ldots + \beta_n x_{t-n} + e_t$$

  - n lags and no structure in parameters
  - OLS works
- but
  - n observations lost
  - high multicollinearity
  - imprecise, large s.e., low t, lots of d.f. Lost
- Structure could help

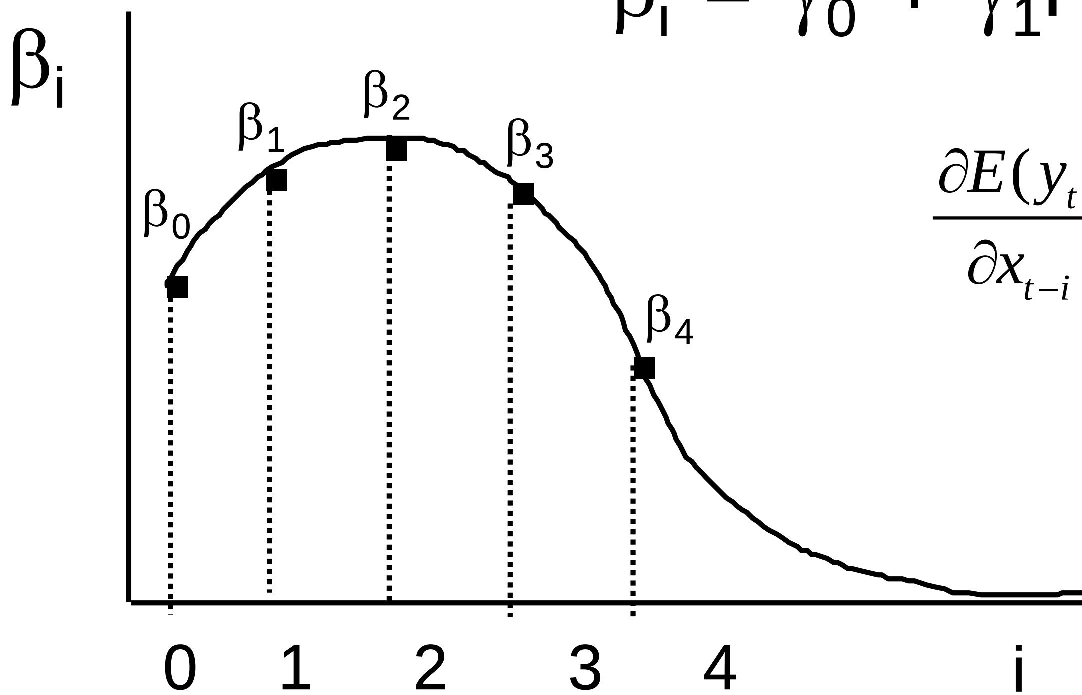
Arithmetic lag - structure



$$\beta_i = \frac{\partial E(y_t)}{\partial x_{t-i}}$$

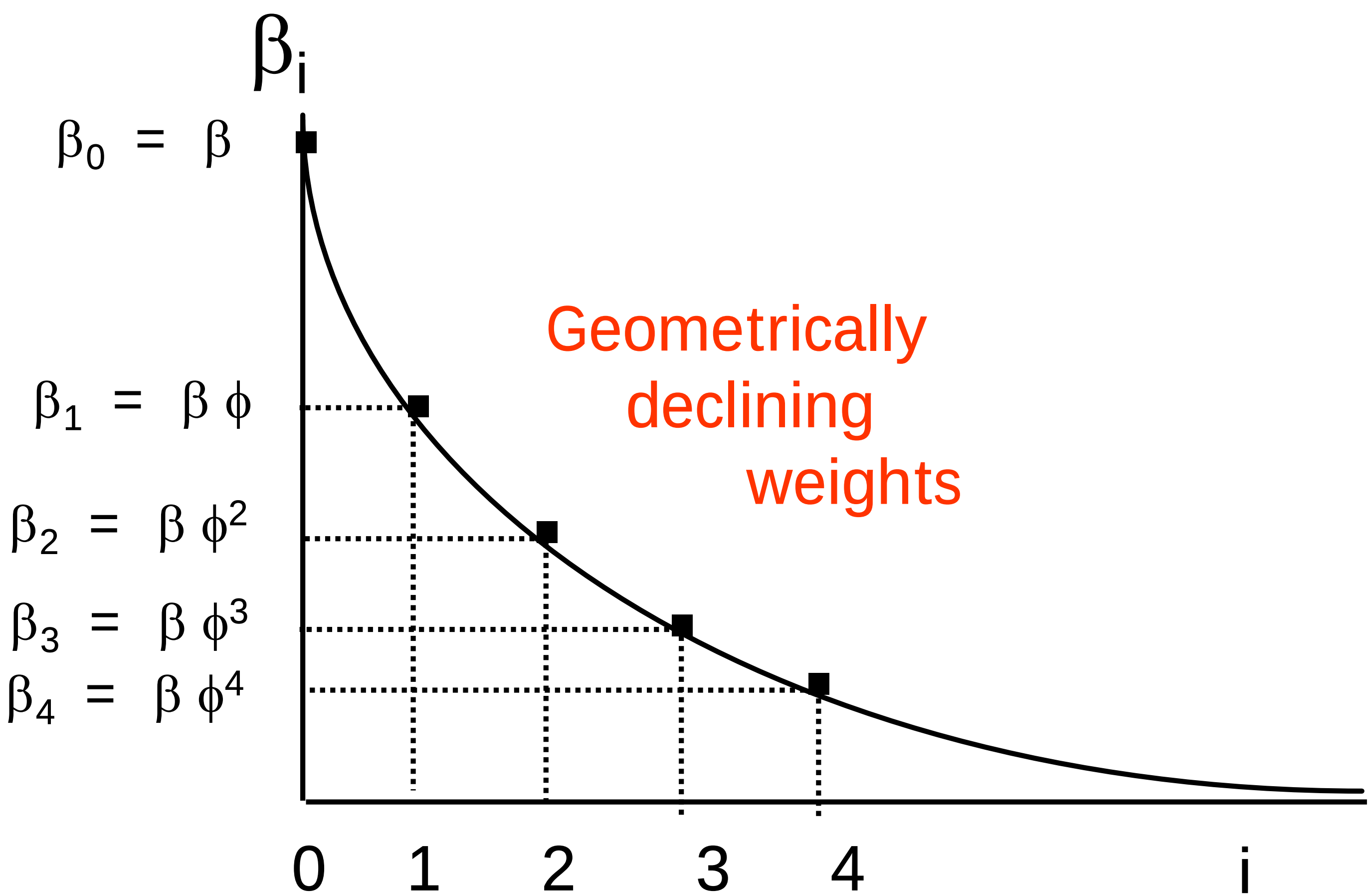
Polynomial lag: structure

$$\beta_i = \gamma_0 + \gamma_1 i + \gamma_2 i^2$$



$$\frac{\partial E(y_t)}{\partial x_{t-i}} = \beta_i$$

Geometric lag: structure



Koyck transformation

- Model:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \ldots + \beta_k X_{t-k} + \varepsilon_t$$

- The Koyck transformation suggests that the further back in time we go, the less important is that factor (for instance, information from 10 years ago vs. information from last year)
- The transformation suggests:

$$\beta_j = \beta_0 \lambda^j \quad , \quad \text{where } 0 < \lambda < 1,\ j = 1,\ldots k$$

Koyck transformation: Analysis

- ***Why is this transformation useful?***
  - Allows us to take the ad-hoc lag series and condense it into a lagged endogenous variable
  - now we only lose one observation due to the lagged endogenous variable
  - the λ given by the estimation gives the coefficient of autocorrelation
- ***Problem: by construction, we have first-order autocorrelation***
  - use Durbin h-statistic
  - but estimating equation might be mis-specified!

Economic rational for Koyck model

- Adaptive Expectations
- Partial Adjustment

Adaptive Expectation Model

Basic model : $y_t = \alpha + \beta X_t^e + \mu_t$

After calcus, we get

$$\mathrm{y_t} = \alpha + \beta\theta X_t + \beta\theta(1-\theta)X_{t-1} + \beta\theta(1-\theta)^2 X_{t-2} + \ldots + \mu_t$$

This is the Koyck model with $\beta_0 = \beta\theta \; and \; \lambda = (1-\theta)$

Our estimating equation therefore is

$$\mathrm{y_t} = \alpha\theta + \beta\theta X_t + (1-\theta)y_{t-1} + (\mu_t - (1-\theta)\mu_{t-1})$$

Partial Adjustment model

Let y_t^d – the desired level of Y in period t and

is a function of X_t, *i.e.*

$$y_t^d = \alpha + \beta X_t + \mu_t,$$

one adjusts the actual level of y

according to the adjustment process

$$(y_t - y_{t-1}) = \theta(y_t^d - y_{t-1}),\ 0 < \theta < 1$$

substituting the first equation into the second, we get

$$y_t = \alpha\theta + \beta\theta X_t + (1-\theta)y_{t-1} + \theta\mu_t$$

Estimating Koyck model

- Model can be estimated by **maximum likelihood**. It is difficult.
- Simple method of estimation is **instrumental variables**.

QUESTIONS?

THANK YOU FOR YOUR ATTENTION!