الفصل الخامس: نماذج الانحدار غير الخطية

Non-linear Regression

أ. لولوه بن سعيد

أولاً:متعدد الحدود Polynomials

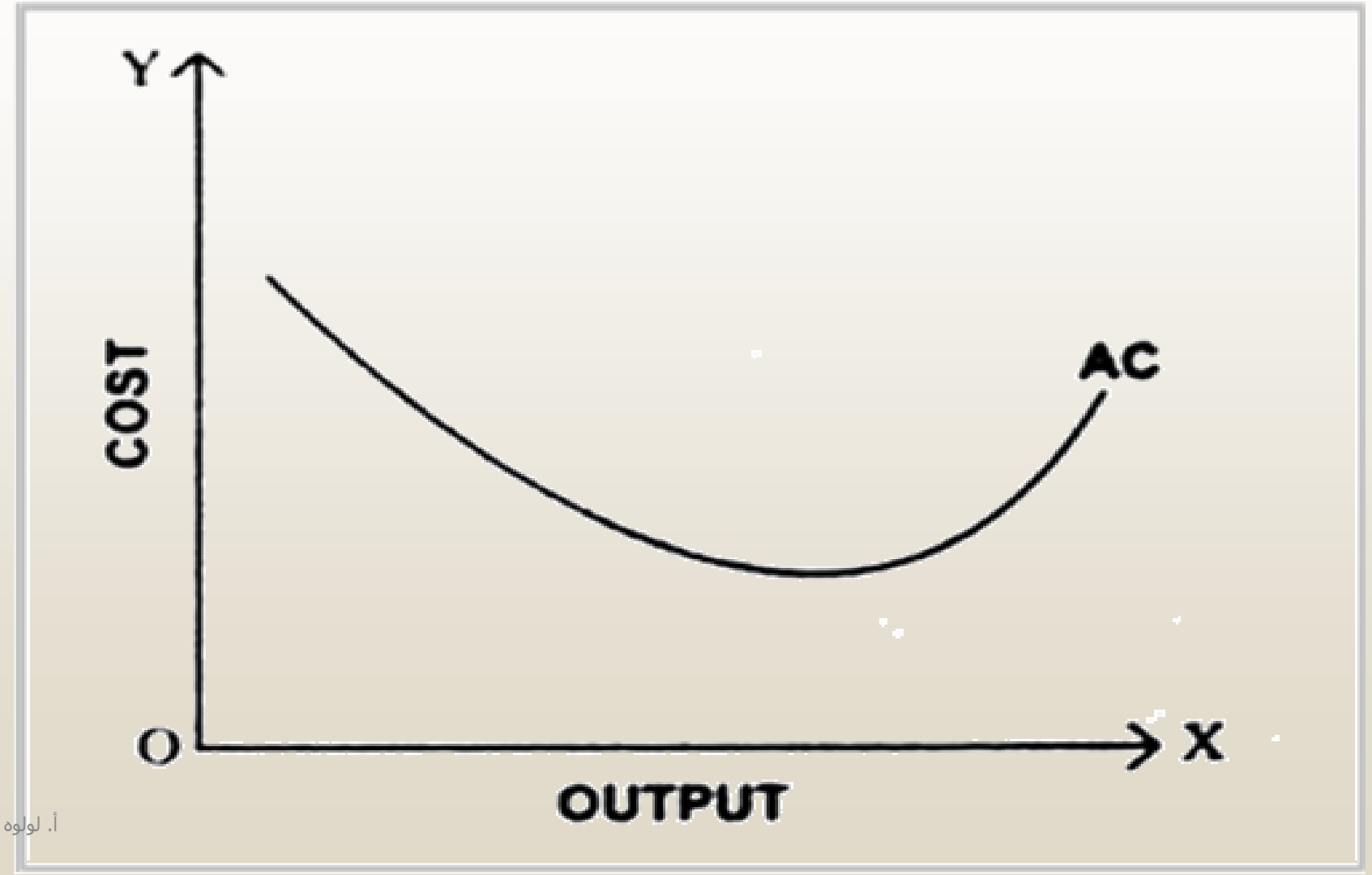
- نعلم بأنه يتم أخذ الشكل التكعيبي المعياري لمنحنيات متوسط التكلفة و التكلفة الحدية و منحنيات الإنتاج للمنشآت، و التي يعد ميلها غير ثابت فلا يمكن توضيحه باستخدام نماذج الانحدار الخطية.

- على سبيل المثال، يكون شكل نموذج متوسط التكلفة كما يلي:

$$AC = \beta_1 + \beta_2 Q + \beta_3 Q^2 + e$$

و بالتالي يأخذ منحنى الدالة التربيعية شكل حرف U (لأنه غير خطي)

مثال: شكل دالة التكاليف المتوسطة:



مثال: معادلة الأجور

- عندما تكون الأجور دالة في التعليم و عدد سنوات الخبرة للعاملين، فمن المتوقع أن تكون أجور العمال الجدد منخفضة مقارنة بمن لديهم خبرات أكبر. و لأخذ نمط دورة الحياة في التحليل سندخل الخبرة و مربع الخبرة لشرح مستوى الأجور، فتكون الدالة كما يلي:

$$wage = \beta_1 + \beta_2 educ + \beta_3 exper + \beta_4 exper^2 + e$$

- و بافتراض أنه تم تقدير الدالة باستخدام بيانات معينة لمتغيرات تحتوي على عينة ١٠٠٠ مشاهدة، فإنه الأمر للشكل السابق في برنامج E-views يتم إدراجه بالصيغة التالية:

```
wage c educ exper exper^2
```

- مخرجات التقدير الافتراضية من E-views:

Dependent Variable: WAGE
Method: Least Squares
Da 0 Time: 19:31
Sample: 1 1000
Included observations: 1000

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-9.8177	1.054964	-9.3062	0.0000
EDUC	1.210072	0.070238	17.22821	0.0000
EXPER	0.340949	0.051431	6.629208	0.0000
EXPER^2	-0.00509	0.001198	-4.25151	0.0000

R-squared	0.270934	Mean dependent var	10.21302
Adjusted R-squared	0.268738	S.D. dependent var	6.246641
S.E. of regression	5.341743	Akaike info criterion	6.192973
Sum squared resid	28420.08	Schwarz criterion	6.212604
Log likelihood	-3092.49	Hannan-Quinn criter.	6.200434
F-statistic	123.3772	Durbin-Watson stat	0.491111
Prob(F-statistic)	0.000000		

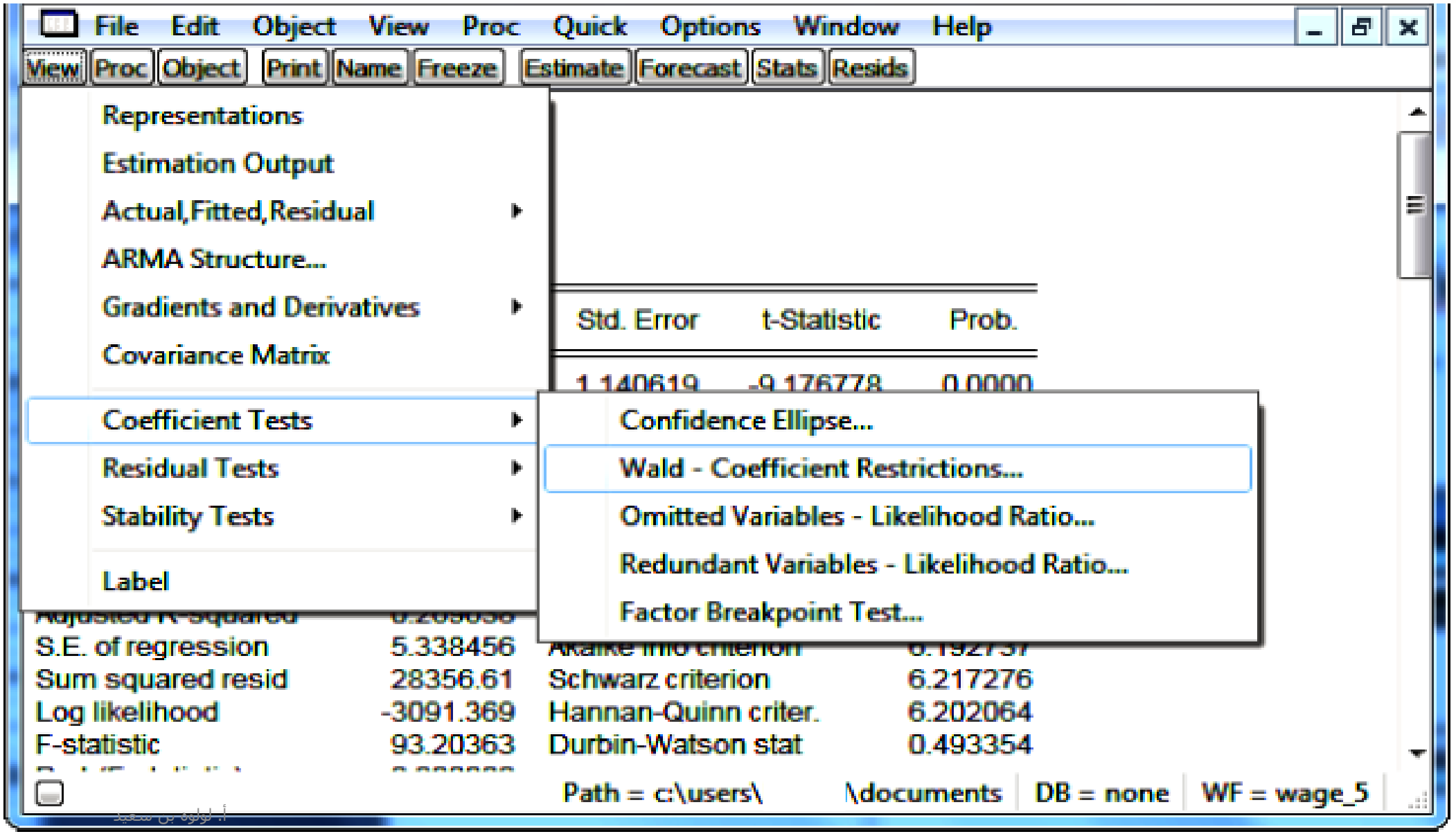
- إن تشكيل نموذج غير خطي متعدد الحدود يتطلب بعض الجهد، فيتوقع مما سبق أن يكون تأثير التعليم على الأجور يوضحه المعامل ١.٢١. مما يعني أن كل سنة تعليم يقدر لها أن تؤدي إلى زيادة أجر ساعة العمل بمقدار ١.٢١ مع بقاء العوامل الأخرى على حالها.

- و نستنتج أيضاً أن الأثر الحدي للخبرة على الأجور مع بقاء التعليم و العوامل الأخرى على حالها هو:

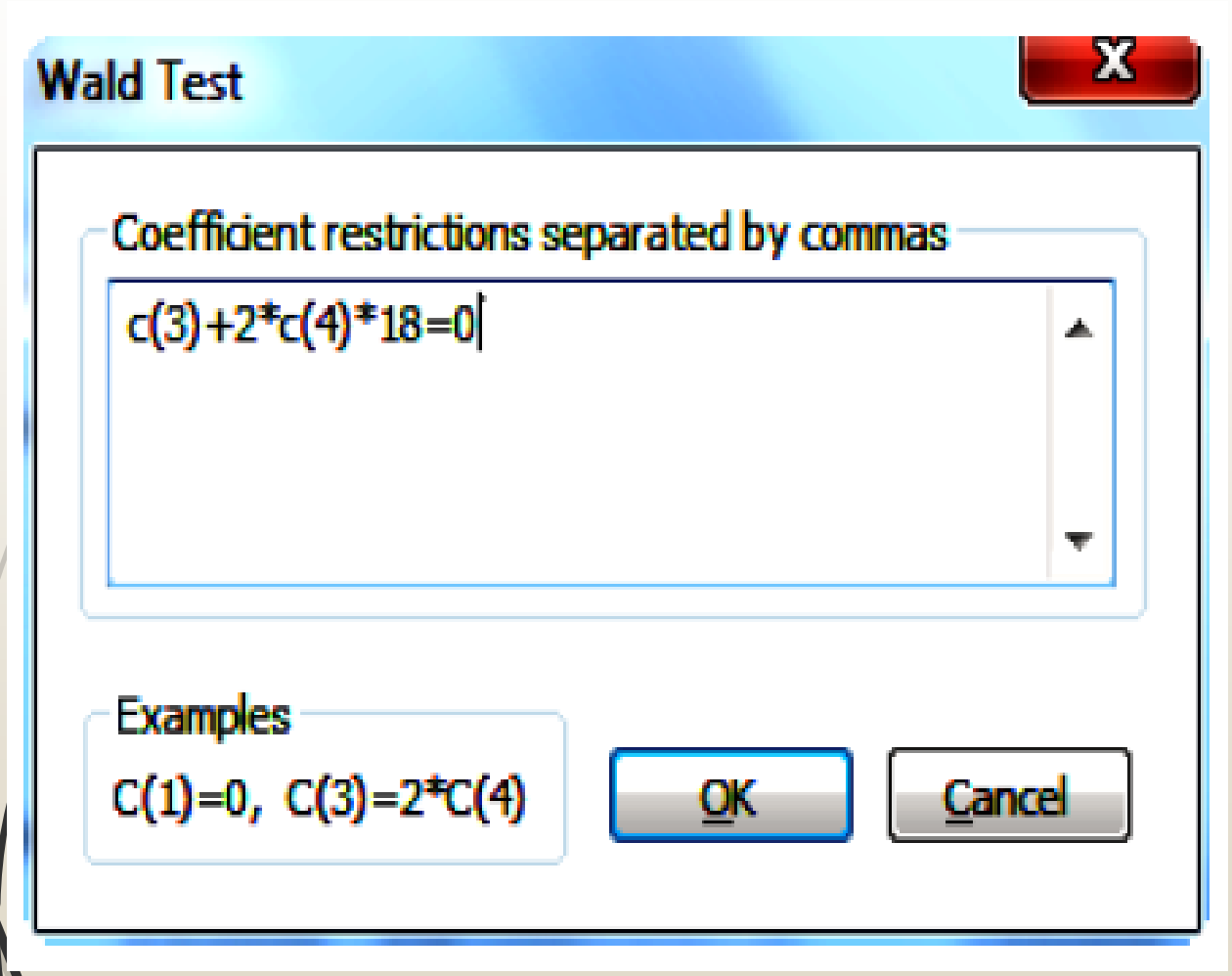
$$\frac{\partial E(wage)}{\partial exper} = \beta_3 + 2\beta_4\ Exper$$

و لتقييم الأثر الحدي على المستوى الفردي مثل سنوات الخبرة التي تساوي ١٨ سنة في E-views، فإنه يتم اختيار الأمر التالي:

View – Coefficient Test – Wald Coefficient Restriction


File Edit Object View Proc Quick Options Window Help
View Proc Object Print Name Freeze Estimate Forecast Stats Resids
Representations
Estimation Output
Actual,Fitted,Residual
ARMA Structure...
Gradients and Derivatives
Covariance Matrix
Coefficient Tests
Residual Tests
Stability Tests
Label
Confidence Ellipse...
Wald - Coefficient Restrictions...
Omitted Variables - Likelihood Ratio...
Redundant Variables - Likelihood Ratio...
Factor Breakpoint Test...
Std. Error t-Statistic Prob.
S.E. of regression 5.338456
Sum squared resid 28356.61
Log likelihood -3091.369
F-statistic 93.20363
Schwarz criterion 6.217276
Hannan-Quinn criter. 6.202064
Durbin-Watson stat 0.493354
Path = c:\users\ \documents DB = none WF = wage_5

ثم يتم كتابة المعادلة التالية:



حيث أن:

$$C(3) = b3$$

$$C(4) = b4$$

أما القيد المدخل Coefficient Restriction هو الأثر لحدي الذي يساوي صفر في الشكل. و هذا الأمر لا يختبر فرضية الأثر الحدي الذي يساوي الصفر فقط، بينما يحسب الأثر الحدي و الانحراف المعياري للأثر كما في الشكل التالي:

File Edit Object View Proc Quick Options Window Help

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Wald Test:
Equation: WAGE_QUADRATIC

F-test of hypothesis

Test Statistic	Value	df	Probability
F-statistic	0.016256	(1, 995)	0.8986
Chi-square	0.016256	1	0.8985

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(3) + 36*C(4)	-0.014883	0.116727

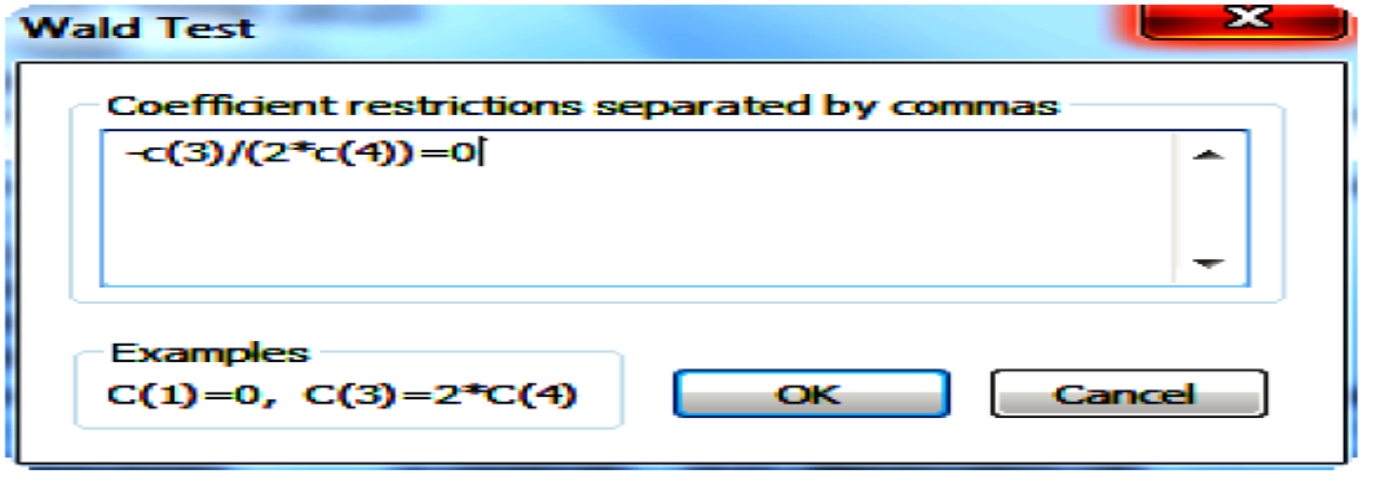
Restrictions are linear in coefficients.

calculated marginal effect

- و عند إجراء الإحصاء الوصفي للمتغيرات السابقة Wage, exper, educ من خلال أيقونة Descriptive stats نستنتج ما يلي:

	EDUC	EXPER	WAGE
Mean	13.28500	18.78000	10.21302
Median	13.0000	18.00000	8.790000
Maximum	18.0000	52.00000	60.19000
Minimum	1.00000	0.000000	2.030000
Std. Dev.	2.468171	11.31882	6.246641

- يبلغ الأجر أقصاه عندما تساوي الخبرة ß3/2ß4 ، و من الايفيوز نفتح wege_quadrartic ونختار view – coefficient Test/ Wald coefficient Restriction و نتحصل على ما يلي:



و نستنتج التالي:

Normalized Restriction (= 0)	Value	Std. Err.
-1 / 2 * C(3) / C(4)	33.47192	3.393876

و بذلك يكون أقصى أجر نحصل عليه عند مستوى خبرة 33 سنة و نستخدم الانحراف المعياري لحساب فترة التقدير.

استخدامات النماذج غير الخطية:

- المقياس في اختيار النماذج الغير خطية هي النظريات الاقتصادية، فإذا كانت علاقة النظرية خطية فتقدر بمعادلة خطية و العكس إذا كانت العلاقات في النظريات غير الخطية فتطبق النماذج غير الخطية.

- كذلك هو الحال في المعادلات التربيعية أو التكعيبية أو اللوغاريتمية.

$$Y = \beta_1 + \beta_2 X_2^2 + \beta_3 \sqrt{X_3} + \beta_4 \log X_4 + u$$

- يمكن إعادة تحويل شكل المعادلة السابقة بحيث تكون معادلة خطية، فنستبدل:

$Z2 = X^2$ ، $Z3 = \sqrt{X3}$ ، $Z4 = \log X\ 4$

و يصبح شكل المتغيرات الآن خطية كما هو الحال في المعلمات، و هذا النوع من التحويل هو شكلي فقط:

$$Y = \beta_1 + \beta_2 Z_2 + \beta_3 Z_3 + \beta_4 Z_4 + u$$

مثال: منحنى فيليبس Philips Curve:

- يمثل منحنى فيليبس العلاقة غير الخطية بين معدل البطالة و نسبة التغير في الأجور:

$$W = \beta_1 + \beta_2 \frac{1}{U} + u$$

- تفترض الفرضية السابقة أن نسبة الأجور W ترتبط بعلاقة عكسية مع معدل البطالة U، و حيث يتم التقدير بطريقة المربعات الصغرى العادية OLS و نحصل على التقدير الافتراضي التالي:

$$\hat{W} = 0.00679 + 0.1842 \frac{1}{U} + u$$

الدوال اللوغاريتميةLog-models:

- تنقسم النماذج اللوغاريتمية إلى نوعين:

لوغاريتمية كاملة	شبه لوغاريتمية
بحيث يكون المتغير التابع و المتغيرات المفسرة جميعها تحتوي على اللوغاريتم	بحيث يكون إما المتغير التابع يحتوي على لوغاريتم أو المتغيرات المفسرة

- إذاً: قد يوجد اللوغاريتم في الطرف الأيسر من المعادلة فيكون لوغاريتم المتغير التابع log-y، و قد يكون في الطرف الأيمن للمعادلة فيكون لوغاريتم المتغير المستقل log-x أو كليهما.

- لا نستطيع أن نختار شكل النموذج لوغاريتمي أو شبه لوغاريتمي كما نريد، بل تحتم علينا النظريات شكل النموذج.
- يرمز للوغاريتم بالرمز log أو ln للوغاريتم الطبيعي.
- يتم تحويل الدوال غير الخطية إلى معادلات خطية باستخدام أسلوب التحويل اللوغاريتمي Logarithmic Transformation ، فمثلاً: يتم أخذ اللوغاريتم للطرفين فتصبح معادلة خطية كما يلي:

$$\log Y = \log \beta_1 X^{\beta_2}$$

$$= \log \beta_1 + \log X^{\beta_2}$$

$$= \log \beta_1 + \beta_2 \log X$$

- و بالتالي يمكن تقديرها بطريقة المربعات الصغرى العادية

الحالة الأولى: لوغاريتم طرفي المعادلة:
log-log model/ double-log/ constant elasticity model:

- نأخذ على سبيل المثال دالة إنتاج كوب-دوغلاس Cobb-Douglas production function والتي تفسر السلوك الإنتاجي و علاقته بعناصر الإنتاج، و الذي يمكن أن يستخدم في دراسة عملية الإنتاج على مستوى المنشأة أو على مستوى الاقتصاد الوطني ككل.

$$Q_i = \beta_1 L_i^{\beta 2} K_i^{\beta 3}$$

حيث أن:

Q تمثل مستوى الإنتاج **K تمثل عنصر رأس المال**

L تمثل عنصر العمل **β_1 معامل التناسب**

β_2 ، β_3 هي ثوابت تحددها التكنولوجيا

- و يتم تحويل المعادلة غير الخطية السابقة إلى الشكل اللوغاريتمي كالتالي:

$$\ln Q_i = \ln ß1 + \ln ß2\ Li + ln\ ß3\ K + ui$$

و تم ذلك بأخذ اللوغاريتم للطرفين في النموذج السابق، و حيث أن المعادلة غير خطية في المتغيرات Variables و خطية في المعلمات Parameters فيمكن تقديرها بطريقة المربعات الصغرى العادية.

- **من خصائص معلمات دالة كوب- دوغلاس:**

- اذا كان مجموع $\beta_3 + \beta_2$ يساوي ١ فإن عوائد السعة الثابتة **constant returns to scale** أي زيادة عوامل الإنتاج بحجم معين تؤدي إلى زيادة الإنتاج بنفس الحجم.

- أما اذا كانت أكبر من واحد فإن عوائد السعة متزايدة **increasing returns to scale** أي زيادة عوامل الإنتاج بحجم معين تؤدي إلى زيادة الإنتاج بحجم أكبر.

- و اذا كانت أقل من واحد فإن عوائد السعة متناقصة **decreasing returns to scale** أي زيادة عوامل الإنتاج بحجم معين تؤدي إلى زيادة الإنتاج بحجم أقل.

تقدير دالة إنتاج كوب-دوغلاس:

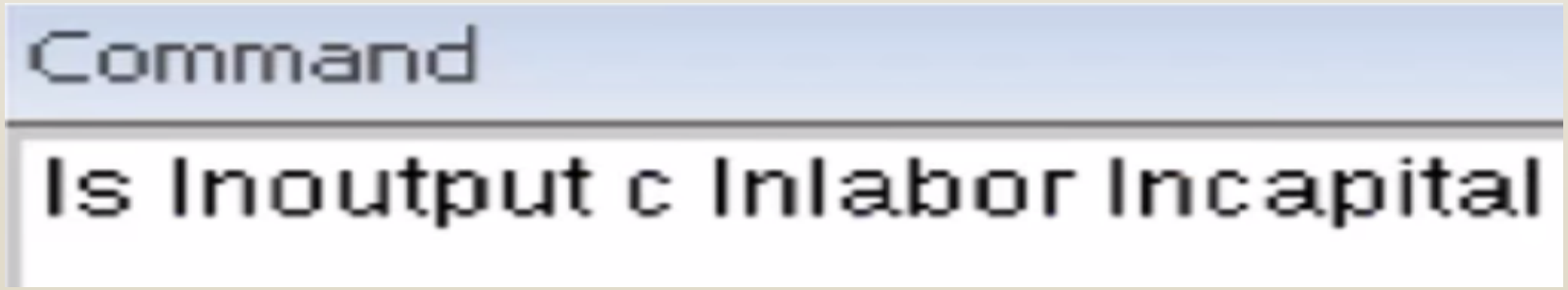
- بافتراض لدينا بيانات دالة إنتاج افتراضية لـ ٥١ ولاية أمريكية في عام ٢٠٠٥، حيث أن:

  الناتج output يقاس بالقيمة المضافة/ألف دولار

  و العمل labor تقاس بساعات العمل/بالألف

  و الإنفاق على رأس المال capital يقاس بالألف دولار

- نستخدم الأمر التالي للتقدير:



- و تصبح النتائج كما يلي:

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: LNOUTPUT
Method: Least Squares
Date: 05/06/16 Time: 16:21
Sample: 1 51
Included observations: 51

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.887600	0.396228	9.811519	0.0000
LNLABOR	0.468332	0.098926	4.734170	0.0000
LNCAPITAL	0.521279	0.096887	5.380281	0.0000

R-squared	0.964175	Mean dependent var	16.94139
Adjusted R-squared	0.962683	S.D. dependent var	1.380870
S.E. of regression	0.266752	Akaike info criterion	0.252027
Sum squared resid	3.415517	Schwarz criterion	0.365664
Log likelihood	-3.426697	Hannan-Quinn criter.	0.295451
F-statistic	645.9317	Durbin-Watson stat	1.946387
Prob(F-statistic)	0.000000		

ملاحظة:

- عندما يكون النموذج لوغاريتمي بالكامل (دالة كوب دوغلاس)، فإنه لو تغير المتغير المستقل بنسبة ١٪ يتغير المتغير التابع بمقدار المعلمة بنسبة مئوية.

- أما عندما يكون النموذج شبه لوغاريتمي (نماذج النمو مثل نمو الناتج المحلي الإجمالي)ـ (حالة لوغاريتم المتغير التابع logy)، فإن تغير المتغير المستقل بوحدة واحدة يتغير المتغير التابع بمقدار المعلمة بنسبة مئوية.

- و عندما يكون النموذج شبه لوغاريتمي (دالة إنفاق إنجلEngel)ـ (حالة لوغاريتم المتغير المفسرlogX)، فإن تغير المتغير المستقل بنسبة ١٪ يتغير المتغير التابع بمقدار المعلمة بالوحدة (ألف، مليون، سم، كيلو، جرام الخ).

الحالة الثانية: دالة شبه لوغاريتميةSemi-log :

- من أمثلتها الدالة المعكوسة Reciprocal Models/ inverse Models:

$$Y_i = \beta_1 + \beta_2(\frac{1}{X_i}) + u_i$$

- ما يهمنا هنا هو معرفة تفسير β_2

حيث أن تزايد X سيؤدي إلى انخفاض قيمة الكسر $\frac{1}{Xi}$ β_2 بالاقتراب من الصفر، و تقترب قيمة Y من قيمة β_1

- اذا كانت β_2 موجبة، فهذا يعني أن الميل سالب

أما اذا كانت β_2 سالبة، فهذا يعني أن الميل موجب

مثال: الدالة المعكوسة:

- مثال: لدينا بيانات افتراضية لإجمالي انفاق الفرد من دخله على السلع و الخدمات المختلفة EXPEND-REC و ذلك في الولايات المتحدة الأمريكية و نرغب في استخدام الدالة المعكوسة لمعرفة كم هو نصيب انفاق الفرد على الطعام فقط SFDHO بافتراض ثبات العوامل الأخرى على حالها:
- يكون الأمر في e-views كالتالي:

ls sfdho c lnexpend_rec

- حيث أن expend_rec في البيانات الأصلية هي $\frac{1}{X}$

فيظهر التالي:

Dependent Variable: SFDHO
Method: Least Squares
Date: 05/07/16 Time: 16:29
Sample: 1 869
Included observations: 869

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.077263	0.004012	19.25950	0.0000
EXPEND_REC	1331.338	63.95713	20.81610	0.0000

R-squared	0.333236	Mean dependent var	0.144736
Adjusted R-squared	0.332467	S.D. dependent var	0.085283
S.E. of regression	0.069678	Akaike info criterion	-2.487556
Sum squared resid	4.209346	Schwarz criterion	-2.476584
Log likelihood	1082.843	Hannan-Quinn criter.	-2.483357
F-statistic	433.3100	Durbin-Watson stat	1.997990
Prob(F-statistic)	0.000000		

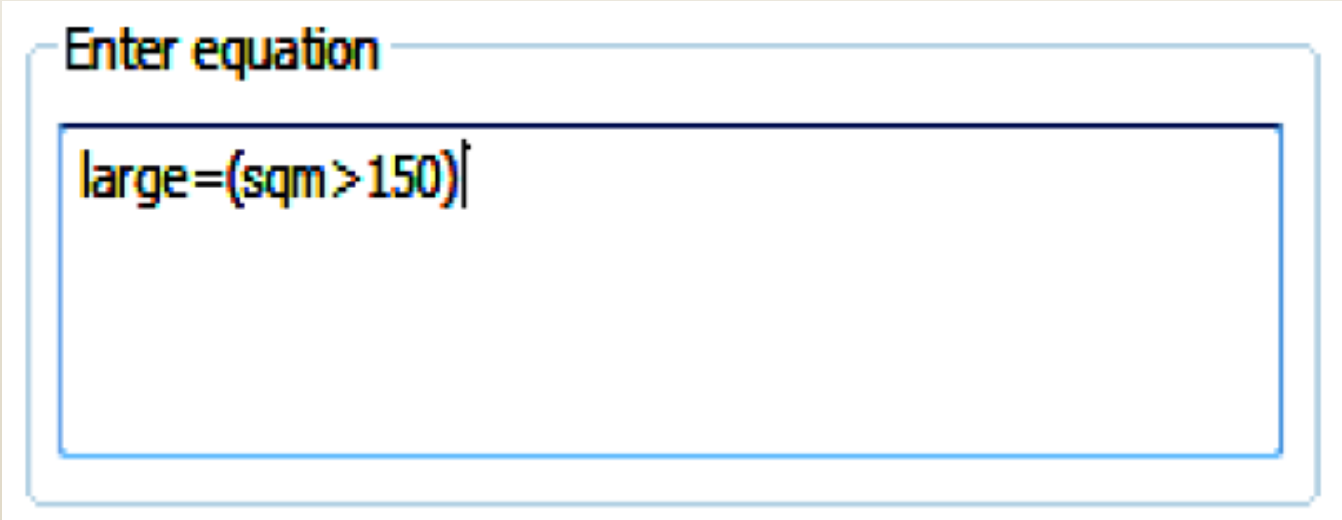
- القاطع يعني أن زيادة دخل الفرد إلى مالا نهاية، نصيب الإنفاق على الطعام يتزايد بمقدار ٠,٠٧
- و يكون تفسير المعلمة للمتغير المستقلEXPEND-REC و التي تساوي ١٣٣١ كما يلي:

بما أن قيمة β_2 ظهرت موجبة فإن ذلك يعني أن الميل سالب: أي عندما يزيد إجمالي انفاق الفرد فالمفترض أن نصيب انفاقه على الطعام من دخله الثابت يقل (علاقة عكسية).

- ملاحظة: يتم استخدام طريقة المربعات الصغرى OLS في جميع الأمثلة السابقة عندما يكون النموذج (خطي في المعلمات parameters) أي β_1 ، β_2 و β_3جميعها خطية، و خطي في المتغيرات X1 , Yخطية، فلا يمكن استخدام طريقة التقدير بالمربعات الصغرى في حالة المعلمات غير الخطية.

المتغيرات الوهمية Dummy Variables:

- هي متغيرات ثنائية تأخذ القيمتين (٠ ، ١) فقط، و تحتاج إلى وجود أو غياب شروط معينة.
- مثال: نأخذ مثال المنازل، إذا كانت كبيرة في المساحة تأخذ القيمة ١ و تأخذ القيمة ٠ في حال عكس من ذلك. فعلينا أن نعرف البيت الكبير بحيث يكون وسيط Median مساحة البيت sqm في العينة هو ١٥٠ م٢ مثلاً، فيتم افتراض أن البيت الذي يزيد مساحته عن ١٥٠م٢ هو بيت كبير. و بالتالي فإنه يتم عمل ما يلي بالنقر على Genr في نافذة E-views و إعطاء المتغير الوهمي الجديد اسم Large للبيت الكبير:



- فسينتج لدينا متغير جديد Large الذي سيعطي القيمة ١ إذا كانت sqm > ١٥٠م٢ و يعطي القيمة ٠ لغيرها:

obs	SQM	LARGE
1	120.0000	0.000000
2	140.0000	0.000000
3	180.0000	1.000000
4	100.0000	0.000000
5	200.0000	1.000000

إيجاد التفاعل بين متغيرين:

- على سبيل المثال نرغب في معرفة التفاعل بين متغير العمر age و الدخل income لبيانات معينة، فنقوم أولاً بتقدير انحدار المربعات الصغرى للمتغير التابع للبيانات نفترض أنه الطلب على البيتزا Pizza.
  و نضيف متغير تفاعل age و income
- نقوم بتقدير دالة الانحدار بطريقة المربعات الصغرى و نستنتج:

Ls pizza c age income age*income

Dependent Variable: PIZZA
Method: Least Squares
Date: 01/27/10 Time: 22:06
Sample: 1 40
Included observations: 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	161.4654	120.6634	1.338147	0.1892
AGE	-2.97742	3.352101	-0.88823	0.3803
INCOME	0.009074	0.00367	2.472717	0.0183
AGE*INCOME	-0.00016	8.67E-05	-1.84715	0.0730

R-squared	0.387319	Mean dependent var	191.55
Adjusted R-squared	0.336262	S.D. dependent var	155.8806
S.E. of regression	126.9961	Akaike info criterion	12.62083
Sum squared resid	580608.7	Schwarz criterion	12.78972
Log likelihood	-248.417	Hannan-Quinn criter.	12.68189
F-statistic	7.586038	Durbin-Watson stat	0.932029
Prob(F-statistic)	0.000468		

أ. لولوه بن سعيد

- و الان نرغب بمعرفة الأثر الحدي للعمر Age :

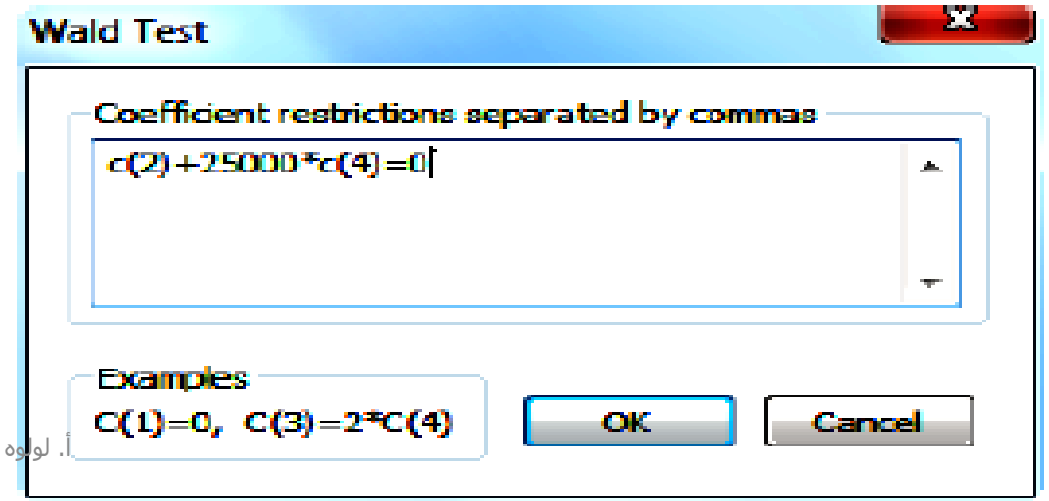
$$\frac{\partial E(Pizza)}{\partial Age} = \beta_2 + \beta_4\ Income$$

- لتقييم الأثر الحدي هذا عندما يكون الدخل income يساوي ٢٥٠٠، فنختار من نافذة الانحدار view / representation و نجد:

Estimation Equation:

PIZZA = C(1) + C(2)*AGE + C(3)*INCOME + C(4)*AGE*INCOME

View/Coefficient Test/Wald Coefficient Restriction اختر



تتضمن نتائج اختبار Wald الأثر الحدي الذي تشكلة الفرضية والإنحراف المعياري.

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(2) + 25000*C(4)	-6.982702	2.267797

الدالة التربيعية:

- مثال على الدالة من الدرجة الثانية:

$$RGDP_t = A_1 + A_2 time + A_3 time^2 + u_t$$

- فإذا أردنا معرفة أثر التغير في الزمن Time على الناتج المحلي الإجمالي الحقيقي RGDP ، فإنه يتم أخذ الأثر الحدي للتغير الزمن بسنة واحدة في الناتج المحلي الحقيقي

$$\frac{dRGDP}{time} = A_2 + 2A_3 time$$

الخلاصة:

MODEL	FORM	SLOPE $(\frac{dY}{dX})$	ELASTICITY $\frac{dY}{dX}\cdot\frac{X}{Y}$
Linear	$Y = B_1 + B_2 X$	B_2	$B_2(\frac{X}{Y})$
Log-linear	$\ln Y = B_1 + \ln X$	$B_2(\frac{Y}{X})$	B_2
Log-lin	$\ln Y = B_1 + B_2 X$	$B_2(Y)$	$B_2(X)$
Lin-log	$Y = B_1 + B_2 \ln X$	$B_2(\frac{1}{X})$	$B_2(\frac{1}{Y})$
Reciprocal	$Y = B_1 + B_2(\frac{1}{X})$	$-B_2(\frac{1}{X^2})$	$-B_2(\frac{1}{XY})$

مراجع: E-views و القياس الاقتصادي – خالد السواعي، Economic Society