



เอกสารประกอบการสอน

วิชา ศ. 325

การใช้โปรแกรม EViews ร่วมกับ

Gujarati, D. Basic Econometrics. 4th edition. 2003.

โดย

อ. เฉลิมพงษ์ คงเจริญ

chaleampong@econ.tu.ac.th

คณะเศรษฐศาสตร์ มหาวิทยาลัยธรรมศาสตร์

ก.พ. 2547

สารบัญ

	หน้า
0. บทนำ	1
1. พื้นฐานเบื้องต้นของการวิเคราะห์การถดถอยเชิงเส้นโดยใช้ EViews	3
1.1 การสร้างแฟ้มงาน(workfile) EViews	3
1.2 การใส่ข้อมูลลงในแฟ้มงาน(workfile) EViews	4
1.3 การสร้างกลุ่มข้อมูล (group) ใน EViews	5
1.4 การวาดกราฟ (graph) ด้วย EViews	6
2. วิธีการกำลังสองน้อยที่สุด (Ordinary Least Square Method)	8
2.1 การคำนวณการถดถอยอย่างง่าย (Simple regression)	8
2.2 ส่วนประกอบต่างๆของหน้าต่างสมการใน EViews	9
2.3 การสร้างแฟ้มงานสำหรับตัวอย่างการใช้จ่ายเพื่อบริโภคในอินเดีย	10
2.4 การนำเข้าข้อมูลจากโปรแกรม spreadsheet ต่างๆ	10
2.5 การแสดงผล Actual, Fitted, Residual	12
2.6 การสร้างอนุกรมของ Residual	13
2.7 การเลือกรูปแบบฟังก์ชัน (Functional form)	15
3. ปัญหา Multicollinearity	16
3.1 Perfect multicollinearity	16
3.2 การตรวจสอบปัญหา multicollinearity ด้วย simple correlation coefficients	16
4. ปัญหา Heteroscedasticity	18
4.1 การตรวจสอบปัญหา Heteroscedasticity ด้วยรูปภาพ	18
4.2 การตรวจสอบปัญหา Heteroscedasticity ด้วย Park test	19
4.3 การตรวจสอบปัญหา Heteroscedasticity ด้วย Goldfeld-Quandt test	20
4.4 การตรวจสอบปัญหา Heteroscedasticity ด้วย White's test	21
4.5 การบรรเทาปัญหา Heteroscedasticity ด้วย Weighted least square	22
4.6 การบรรเทาปัญหา Heteroscedasticity ด้วย White's heteroscedasticity corrected standard error	25
5. ปัญหา Autocorrelation	27
5.1 การตรวจสอบปัญหา Autocorrelation ด้วยรูปภาพ	27
5.2 การตรวจสอบปัญหา Autocorrelation ด้วย Durbin-Watson d stat	28
5.3 การตรวจสอบปัญหา Autocorrelation ด้วย Breusch-Godfrey (BG) Test	28
5.4 การบรรเทาปัญหา Autocorrelation ด้วย First-Difference Method	30
5.5 การบรรเทาปัญหา Autocorrelation ด้วย Generalized Least Square (GLS) โดยใช้ coefficient of autocovariance (ρ) จาก Residual	31
5.6 การบรรเทาปัญหา Autocorrelation ด้วย Generalized Least Square (GLS) โดยใช้ Iterative method of estimating ρ ด้วยวิธี Cochrane-Orcutt procedure	32
5.7 วิธีการ Newey-West เพื่อแก้ไข Standard Error ของ OLS	33

0. บทนำ

พื้นฐานเกี่ยวกับโปรแกรม EViews และ objects

โปรแกรม EViews ถูกออกแบบมาให้ทำงานในลักษณะของ objects ซึ่งในการทำงานของเราก็จะสัมพันธ์กับ objects เหล่านี้ โดยที่แต่ละ objects ก็จะมีข้อมูลและลักษณะงานที่แตกต่างกัน ซึ่ง objects ที่สำคัญที่สุดในโปรแกรมนี้ก็คือ workfile (แฟ้มงาน) ซึ่งในการทำการวิเคราะห์ใดๆก็ตาม เราจำเป็นต้องสร้าง หรือเปิด workfile ขึ้นมาก่อน โดยที่ชนิด objects ต่างๆจะบ่งบอกถึงลักษณะของการทำงาน เช่น series (อนุกรม) ก็จะเกี่ยวข้องกับข้อมูล equation (สมการ) ก็จะเกี่ยวข้องกับความสัมพันธ์ระหว่างตัวแปร เป็นต้น โดยในแฟ้มงานหนึ่งๆอาจมีสมการหลายสมการ ซึ่งเราสามารถเรียกกลับมาใช้ได้เมื่อต้องการ (โดยทำการ double click ที่ objects สมการ) ดังนั้นเราจึงควรจัดการ objects ให้เข้าใจง่าย เพื่อประโยชน์ในการวิเคราะห์ผล ทดสอบสมมุติฐาน หรือ ทำนายค่า

objects พื้นฐานในโปรแกรม EViews ได้แก่ workfiles series และ equation นอกจากนี้ยังมี objects อื่นๆที่ทำหน้าที่เฉพาะด้าน โดยสรุปแล้ว objects ใน EViews จะประกอบด้วย Coefficient Vector, Databases, Equation, Graph, Group, Model, Pool (Time series/ Cross-section), Sample, Series, State Space, System, SYM(Symmetric Matrix), Table, Text, VAR(Vector Autoregression), Vector/Row, และ Vector Scalar ซึ่ง objects เหล่านี้ ยกเว้น Workfiles และ Databases จะมี icons ของตัวเอง เมื่อเราเปิด workfile ใหม่ขึ้นมาทำงาน ในหน้าต่างนั้นจะมี objects สองชนิดเปิดขึ้นมาเสมอ คือ Coefficient Vector (ซึ่งสมาชิกทุกตัวเท่ากับ 0) กับ residual series (ซึ่งสมาชิกทุกตัวเท่ากับ NA's)

	Coefficient Vector		Scalar
	Equation		Series
	Graph		Sspace (State Space)
	Group		System
	Logl (Log Likelihood)		Sym (Symmetric Matrix)
	Matrix		Table
	Model		Text
	Pool (Time Series / Cross-Section)		Var (Vector Autoregression)
	Sample		Vector/Row Vector

เพื่อสร้าง objects ใหม่เราจะเลือก **Objects/New Object** จาก main menu หรือ workfile menu แล้วเลือกชนิดของ objects ที่เราต้องการ ตั้งชื่อ แล้ว คลิก **OK**

เครื่องหมายทางคณิตศาสตร์ในโปรแกรม EViews:

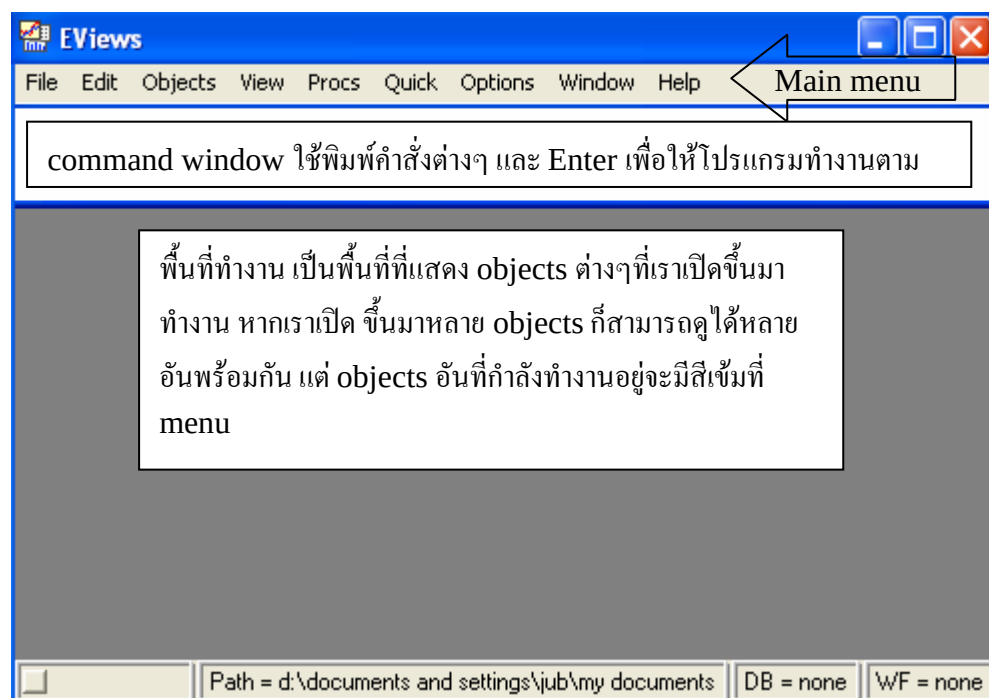
ในโปรแกรม EViews จะบรรจุเครื่องหมายและฟังก์ชันทางคณิตศาสตร์ไว้มากมาย ในขณะที่เป็นพิมพ์มีจำนวนจำกัด โดยที่ผู้ใช้สามารถดูได้จาก Function Reference ในเมนู Help นอกจากนี้เครื่องหมายทางคณิตศาสตร์และสถิติมาตรฐานแล้ว EViews ยังได้สร้างฟังก์ชันพิเศษสำหรับข้อมูลอนุกรมเวลา (time series) ที่หลายเช่น leads, lags และ differences ไว้อีกด้วย

โปรแกรม Eview จะปฏิบัติตามคำสั่งจากซ้ายไปขวา โดยที่เครื่องหมายต่างๆจะมีลำดับความสำคัญดังนี้

1. ^ (ยกกำลัง)
2. * (คูณ), / (หาร)
3. + (บวก), - (ลบ)
4. < (น้อยกว่า), > (มากกว่า), <= (น้อยกว่าหรือเท่ากับ), >= (มากกว่าหรือเท่ากับ), = (เท่ากับ)
5. and, or

หากต้องการดูเครื่องหมายหรือฟังก์ชันทั้งหมด ให้คลิก **Help/Function Reference**

พื้นที่ต่างในหน้าต่างหลักของโปรแกรม EViews

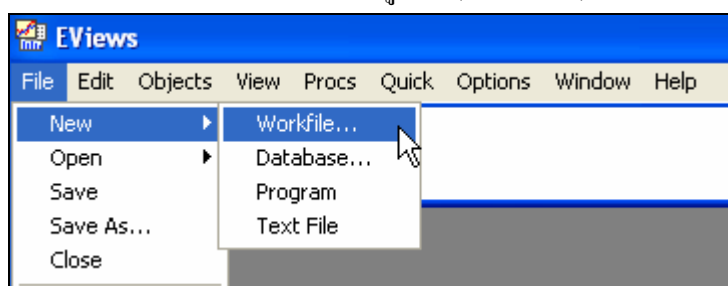


1. พื้นฐานเบื้องต้นของการวิเคราะห์การถดถอยเชิงเส้นโดยใช้ EViews

เพื่อให้นักศึกษาได้เห็นภาพของการใช้โปรแกรม EViews ควบคู่ไปกับเนื้อหาที่เรียน ผู้เขียนจะใช้ตัวอย่างจากหนังสือ Gujarati, 2003. *Basic Econometrics*. 4th edition. (ต่อไปจะย่อด้วย G) เริ่มต้นจะใช้ตัวอย่างเกี่ยวกับ รายได้ กับการใช้จ่ายเพื่อการบริโภค (Gujarati ตารางที่ 2.4, หน้า 48)

1.1 การสร้างแฟ้มงาน (workfile) EViews

ขั้นตอนที่ 1. เลือก **File/New/Workfile** ในเมนูหลัก (main menu) ของ EViews

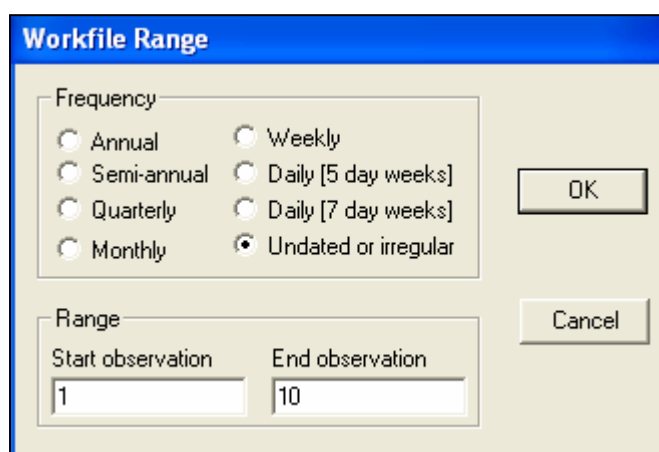


ขั้นตอนที่ 2. ตั้งความถี่ของข้อมูล ซึ่ง

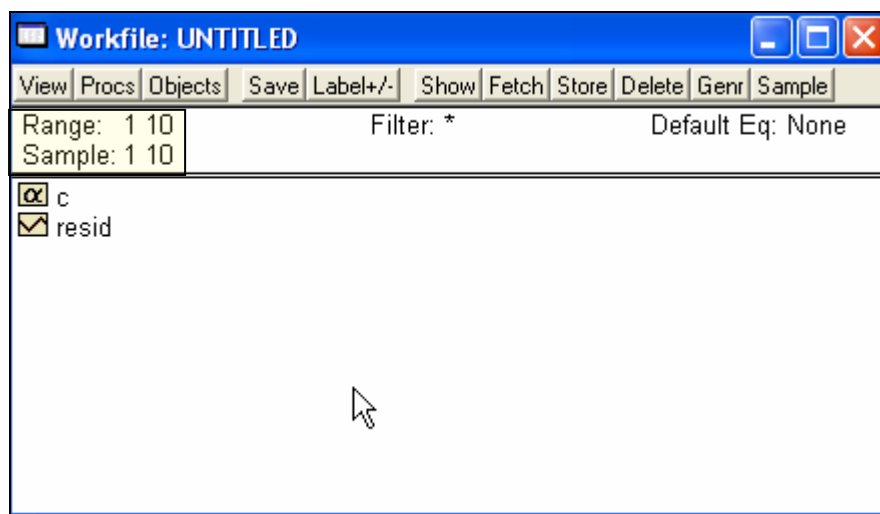
อาจเป็นข้อมูลอนุกรมเวลา รายปี, รายครึ่งปี, ..., รายวัน หรือ เป็นข้อมูลภาคตัดขวาง เช่นกรณีนี้เราจะเลือก Undated irregular

ขั้นตอนที่ 3. นอกจากนี้ เรายังต้องให้ข้อมูลเกี่ยวกับ จุดเริ่มต้นของข้อมูล และ จุดสิ้นสุดของข้อมูลด้วย ในกรณีของข้อมูลภาคตัดขวาง ก็ให้ใส่

จำนวนตัวอย่าง เช่น กรณีนี้ คือจาก 1 ถึง 10



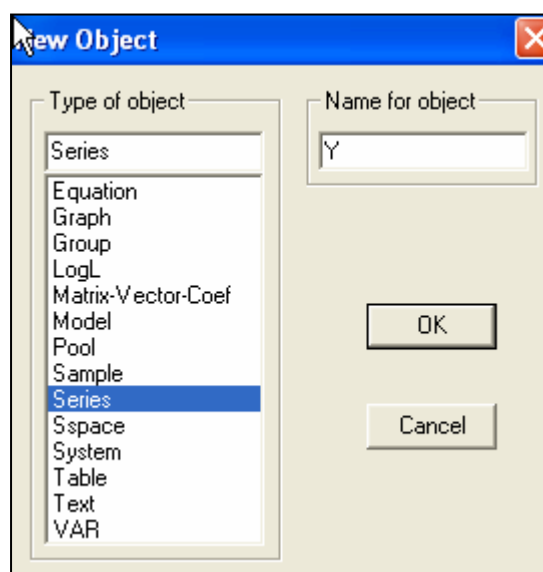
ขั้นตอนที่ 4. หลังจากเลือกความถี่และช่วงของแฟ้มงาน ที่สอดคล้องกับข้อมูลของเราแล้ว ก็คลิก **OK** EViews ก็จะสร้างแฟ้มงาน ชื่อ 'UNTITLED' และแสดงหน้าต่างดังกล่าวในพื้นที่ทำงาน ซึ่งในหน้าต่างแฟ้มงาน จะปรากฏตัวเลขสองตัว คือ ช่วง (range) และ ตัวอย่าง (sample) ซึ่งเราสามารถเปลี่ยนแปลงได้ นอกจากนี้ เรายังสังเกตเห็นว่าแฟ้มงานใหม่จะมี object 2 อันปรากฏอยู่ คือ coefficient vector 'c' หรือ เวกเตอร์ค่าคงที่ กับ residual series 'resid'



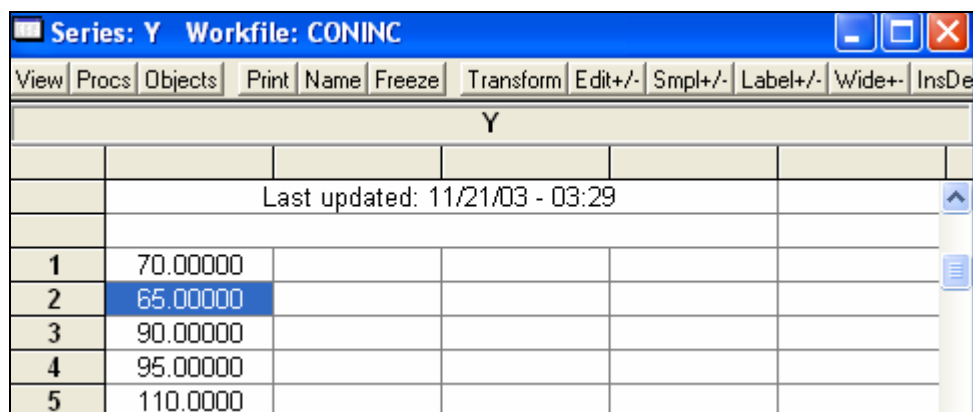
ขั้นตอนที่ 5. เพื่อเก็บ (save)แฟ้มงาน ที่เราได้สร้างขึ้น ให้เลือก **Save** ในเมนูของแฟ้มงาน หรือ **File/Save** หรือ **File/Save As** ในเมนูหลัก และ ใส่ชื่อ coninc.wfl ในช่อง File name และเลือก drive ที่เราต้องการเก็บข้อมูลไว้ แล้วคลิก **OK**

1.2 การใส่ข้อมูลลงในแฟ้มงาน (workfile) EViews

ขั้นตอนที่ 1. เพื่อสร้างชุดของข้อมูล (Series) ใหม่ สำหรับตัวแปร การใช้จ่ายเพื่อการบริโภคต่อสัปดาห์ (Y) เลือก **Object/New Object/Series** จากเมนูหลัก หรือ เมนูแฟ้มงาน แล้วใส่ชื่อ 'Y' ในช่อง Name of Object แล้วคลิก **OK** จะปรากฏ object 'Y' ในพื้นที่ทำงาน ซึ่งสมาชิกทุกตัวในชุดข้อมูลใหม่นี้เป็น NA (Not Available)



ขั้นตอนที่ 2. เพื่อจะใส่ข้อมูลให้ชุดข้อมูลนี้ ให้ double-click ที่ชุดข้อมูล Y จะปรากฏหน้าต่างของชุดข้อมูล ให้คลิก **Edit+/-** ในเมนูชุดข้อมูล แล้วใส่ข้อมูลจากตาราง 2.4 ลงในหน้าต่างดังกล่าว แทนที่ NA เมื่อเสร็จแล้วให้ แล้วคลิก **Edit+/-** เป็นการเก็บข้อมูลดังกล่าว และออกจากการแก้ไขข้อมูล ปิดหน้าต่างดังกล่าวโดยคลิก  บนมุมบนซ้ายของหน้าต่างชุดข้อมูล



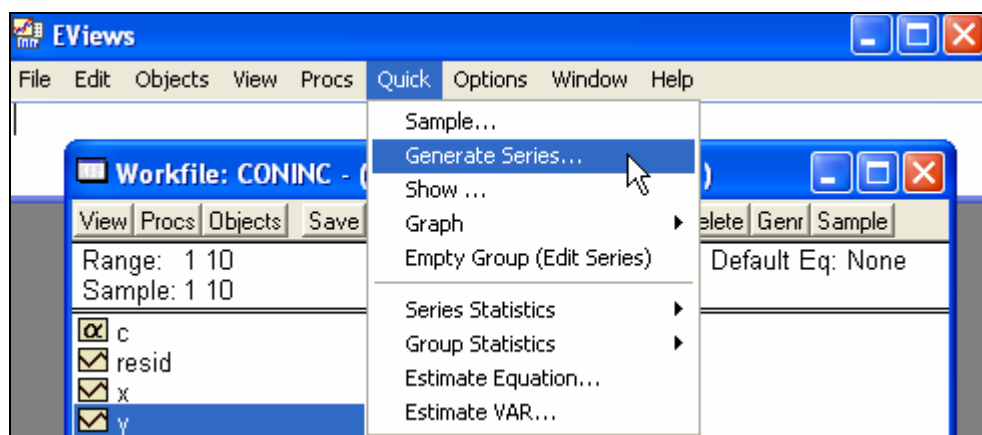
Series: Y Workfile: CONINC					
View Procs Objects Print Name Freeze Transform Edit+/- Smpl+/- Label+/- Wide+/- InsDe					
Y					
Last updated: 11/21/03 - 03:29					
1	70.00000				
2	65.00000				
3	90.00000				
4	95.00000				
5	110.0000				

ขั้นตอนที่ 3. ทำกระบวนการดังกล่าวซ้ำกับตัวแปร รายได้ต่อสัปดาห์ (X)

ขั้นตอนที่ 4. เพื่อเก็บการเปลี่ยนแปลงดังกล่าวที่เกิดขึ้นในแฟ้มงานให้ คลิก **Save** บนเมนูแฟ้มงาน

ข้อมูลส่วนใหญ่ที่เรานำมาใช้มักจะเป็นข้อมูลทศนิยมซึ่งอยู่ในรูป spreadsheet file หรือ ASCII text file ที่เราสามารถนำเข้า (import) ข้อมูลดังกล่าวมาสู่แฟ้มงานได้เลย ซึ่งกระบวนการดังกล่าวจะกล่าวถึงในบทต่อไป อย่างไรก็ตามผู้ใช้อาจใส่ข้อมูลโดยการคลิก **Quick/Empty Group (Edit Series)** เพื่อเปิดตารางว่างสำหรับกรอกข้อมูล หรือ copy & paste จากแฟ้ม spreadsheet อื่นๆก็ได้ (ในกรณีที่ข้อมูลไม่มากจนเกินไป)

ในบางกรณีเราอาจสร้างชุดข้อมูลใหม่จากข้อมูลชุดเดิม โดยคลิก **Quick/Generate Series** หรือคลิก **Genr** ในเมนูแฟ้มงาน และใช้เครื่องหมายต่างๆในการสร้างข้อมูลชุดใหม่ เช่น อนุกรมการบริโภคต่อรายได้ ($Z=Y/X$)



1.3 การสร้างกลุ่มข้อมูล (group) ใน EViews

EViews มีเครื่องมือโดยเฉพาะในการทำงานกับกลุ่มข้อมูล (Group of series) ในที่นี้เราจะสร้างกลุ่มข้อมูลที่ประกอบด้วยชุดข้อมูล X และ Y

ขั้นตอนที่ 1. เปิดแฟ้มงาน 'coninc.wfl' โดยเลือก **File/Open/Workfile** ในเมนูหลัก

ขั้นตอนที่ 2. สร้าง object group สำหรับ X และ Y โดยกดปุ่ม **Ctrl** ค้างไว้ขณะเดียวกันใช้เมาส์คลิก X และ Y แล้วเลือก **Show** จากเมนูแฟ้มงาน

ขั้นตอนที่ 3. เพื่อตั้งชื่อกลุ่มข้อมูล ให้คลิก **Name** หรือ **Object/Name** บนเมนูกลุ่มข้อมูล และใส่ชื่อ **GROUP01** ในช่อง Name to identify object แล้วคลิก **OK**

ขั้นตอนที่ 4. เพื่อเก็บการเปลี่ยนแปลงดังกล่าวที่เกิดขึ้นในแฟ้มงานให้ คลิก **Save** บนเมนูแฟ้มงาน

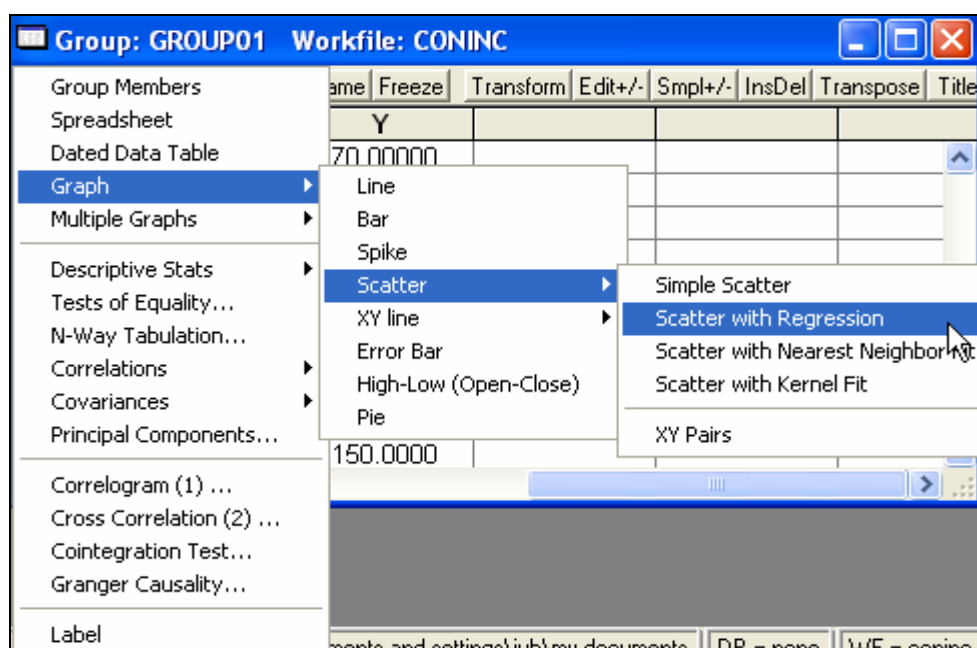
1.4 การสร้างกราฟด้วย EViews

ตามทฤษฎีการบริโภคของเคนส์บอกว่าเมื่อรายได้เพิ่มขึ้น คนจะบริโภคเพิ่มมากขึ้น ดังนั้นตัวแปร Y และ X น่าจะมีความสัมพันธ์เชิงบวก ในขั้นแรกเราจะตรวจสอบความสัมพันธ์ดังกล่าวโดย plot ตัวแปรทั้งสองเพื่อพิจารณาความสัมพันธ์ด้วยสายตา

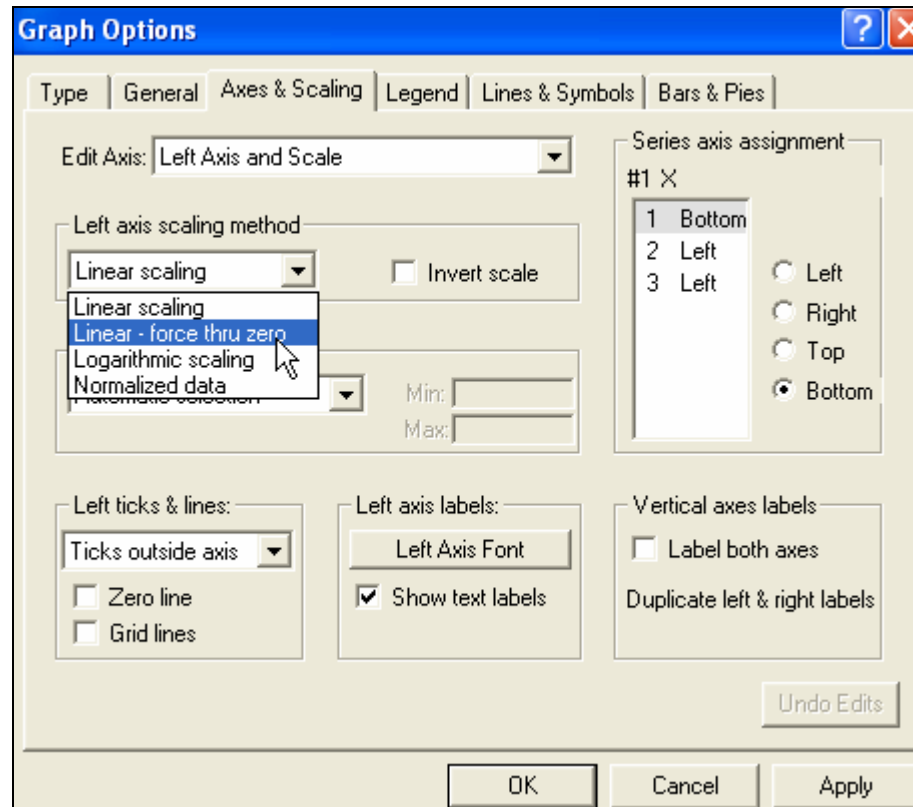
ขั้นตอนที่ 1. เปิดแฟ้มงาน 'coninc.wfl'

ขั้นตอนที่ 2. เพื่อ plot ตัวแปร Y ต่อ X ให้เปิดชุดข้อมูลทั้งสองในหน้าต่างกลุ่มข้อมูล (เราใส่ตัวแปร X ก่อน Y เนื่องจาก EViews จะกำหนดให้ตัวแปรแรกในกลุ่มเป็นแกน 'X' และตัวแปรที่สองเป็นแกน 'Y')

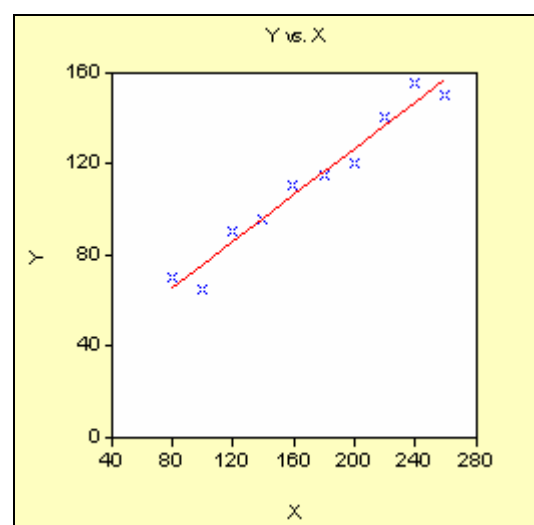
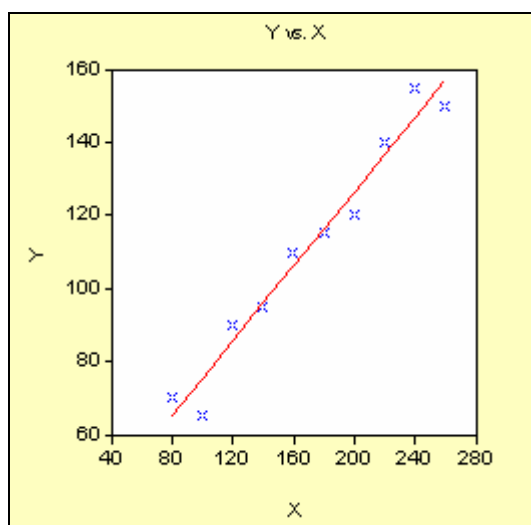
ขั้นตอนที่ 3. เลือก **View/Graph/Scatter/Scatter with Regression** แล้วคลิก **OK** จะได้รูปดังตัวอย่างข้างล่าง ซึ่งเหมือนกับรูปที่ 2.4 หน้า 48 จะเห็นได้ว่าตัวแปร Y และ X มีความสัมพันธ์เชิงบวก EViews จะใช้ **Optimal-Linear Scaling** เป็นค่าตั้งต้นของรูป



ขั้นตอนที่ 4. ถ้าต้องการเปลี่ยนแปลงมาตราส่วนของรูปทำได้โดยการคลิกขวาที่ใดก็ได้ในรูปภาพ และเลือก Options



ขั้นตอนที่ 5. การเปลี่ยน Grape Scaling เป็น Linear-Force Through Zero แล้วคลิก OK จะได้รูปทางขวามือ



2. วิธีการกำลังสองน้อยที่สุด (Ordinary Least Square Method)

การวิเคราะห์การถดถอยด้วยวิธีการกำลังสองน้อยที่สุดถือเป็นแกนหลักของการวิเคราะห์เชิงเศรษฐมิติ โดยที่เราจำเป็นต้องลองประมาณค่าสัมประสิทธิ์ของการถดถอยโดยไม่ใช้โปรแกรมทางสถิติสักครั้ง เพื่อที่จะเข้าใจกระบวนการทำงานของวิธีการกำลังสองน้อยที่สุด ซึ่งกระบวนการดังกล่าวเราจะไม่ได้ทำในการประมาณค่าโดยปกติ เนื่องจากเราสามารถประมาณค่าได้ง่ายด้วยโปรแกรมทางสถิติ ในหัวข้อนี้เราจะประมาณค่าแบบจำลองสมการถดถอยอย่างง่ายเพื่อให้เราเข้าใจกระบวนการทำงานของโปรแกรม EViews และสิ่งที่โปรแกรมแสดงผลออกมา

2.1 การคำนวณการถดถอยอย่างง่าย (Simple regression) [ตัวอย่างในหัวข้อ 3.6]

การประมาณค่าสมการถดถอยใน EViews จะใช้ Object สมการโดยที่มีกระบวนการดังนี้

ขั้นตอนที่ 1. เปิดแฟ้มงาน 'coninc.wfl' โดยเลือก **File/Open/Workfile** ในเมนูหลัก

ขั้นตอนที่ 2. เลือก **Objects/New Object/Equation** จากเมนูแฟ้มงาน หรือเราอาจเลือก **Quick/Estimate Equation** จากเมนูหลักก็ได้

ขั้นตอนที่ 3. ใส่ชื่อของสมการ เช่น EQ01 ในช่อง **Name for Object** แล้วคลิก **OK**

ขั้นตอนที่ 4. ใส่ตัวแปรตามการบริโภค (Y), แล้วตามด้วยตัวแปรอิสระ คือ ค่าคงที่ (C) และ รายได้ (X) ลงในช่อง **Equation Specification** โดยที่ตัวแปรแรกจะต้องเป็นตัวแปรตามเสมอ

ขั้นตอนที่ 5. เลือกวิธีการที่ใช้ในการประมาณค่า ในช่อง **Method** ในกรณีนี้คือ **LS - Least Square (NLS and ARMA)** ซึ่งจะเป็นค่าตั้งต้นสำหรับทุกครั้งที่เปิดหน้าต่างนี้ขึ้นมา

ขั้นตอนที่ 6. ช่วงของตัวอย่างที่ใช้ในการประมาณค่า จะถูกตั้งในเท่ากับช่วงตัวอย่างของแฟ้มงาน ซึ่งเราอาจเปลี่ยนแปลงช่วงตัวอย่างที่ใช้ในการประมาณค่าได้ตามที่เราต้องการ ในช่อง **Sample** แล้วคลิก **OK** เราจะได้ตารางผลการวิเคราะห์ถดถอยด้วยวิธีการกำลังสองน้อยที่สุดโดยโปรแกรม EViews

ขั้นตอนที่ 7. เพื่อเก็บการเปลี่ยนแปลงดังกล่าวที่เกิดขึ้นในแฟ้มงานให้ คลิก **Save** บนเมนูแฟ้มงาน

2.2 ส่วนประกอบต่างๆของหน้าต่างสมการใน EViews

EViews จะแสดงข้อมูลเชิงสถิติที่สำคัญที่เกี่ยวข้องกับการวิเคราะห์เชิงถดถอยส่วนใหญ่ไว้ในหน้าต่างสมการ (ดังรูปข้างล่าง) ซึ่งประกอบด้วยข้อมูลทั่วไปในส่วนบน ค่าสัมประสิทธิ์และตัวสถิติที่เกี่ยวข้องกับตัวแปรแต่ละตัวจะอยู่ส่วนกลาง และส่วนล่างจะประกอบด้วยสถิติโดยสรุป

ส่วนแรก: ข้อมูลทั่วไป

- บรรทัด 1: ชื่อ

ของตัวแปรตาม

- บรรทัด 2:

วิธีการที่ใช้

วิเคราะห์

- บรรทัด 3: วัน

เวลาที่ทำการ

วิเคราะห์

- บรรทัด 4: ช่วง

ของตัวอย่างที่ใช้

ในการวิเคราะห์

- บรรทัดที่ 5: จำนวนตัวอย่างที่รวมในการวิเคราะห์

- บรรทัดที่ 6: จำนวนตัวอย่างที่กันออกไป (excluded observation)

ส่วนที่สอง: ค่าสัมประสิทธิ์ ข้อมูลที่เกี่ยวข้องกับค่าสัมประสิทธิ์ที่ถูกประมาณขึ้น (Estimated coefficient) จะถูกรายงานในส่วนที่สอง โดยที่คอลัมน์แรกเป็นชื่อตัวแปร คอลัมน์ที่สองเป็นค่าสัมประสิทธิ์ ในขณะที่คอลัมน์ 3 – 5 เป็นค่า standard error of coefficient, t-statistics, Probability value(P-value) ตามลำดับ ซึ่งค่าเหล่านี้มีความสำคัญต่อการทดสอบสมมุติฐาน

ส่วนที่สาม: สถิติโดยสรุป สถิติที่สำคัญจะถูกแสดงในสี่คอลัมน์ข้างล่าง ซึ่งแต่ละตัวจะมีความหมายดังที่สรุปไว้ข้างล่างและสามารถอ้างอิงได้จากหน้าใน Gujarati ซึ่งจะวงเล็บไว้ท้ายหัวข้อ

1. R^2 : Coefficient of Determination คือสัดส่วนของความแปรปรวนของตัวแปรตามที่สามารถอธิบายได้ด้วยตัวแปรอิสระ (p.217-219)

2. Adjusted R^2 : (p.217-219)

3. Standard Error of Regression หรือ Standard Error of Estimate: $\hat{\sigma} = \sqrt{\frac{\sum \hat{u}^2}{n-k}}$ (p.78)

4. Sum of square resid: $\sum \hat{u}^2$ ซึ่ง OLS พยายามที่จะเลือกค่าสัมประสิทธิ์ที่ทำให้ค่านี้น้อยที่สุด

5. Log likelihood: ใช้ประโยชน์ในการทดสอบสมมุติฐาน

Dependent Variable: Y					
Method: Least Squares					
Date: 11/28/03 Time: 02:43					
Sample: 1 10					
Included observations: 10					
Variable	Coefficient	Std. Error	t-Statistic	Prob.	
C	24.45455	6.413817	3.812791	0.0051	
X	0.509091	0.035743	14.24317	0.0000	
R-squared	0.962062	Mean dependent var		111.0000	
Adjusted R-squared	0.957319	S.D. dependent var		31.42893	
S.E. of regression	6.493003	Akaike info criterion		6.756184	
Sum squared resid	337.2727	Schwarz criterion		6.816701	
Log likelihood	-31.78092	F-statistic		202.8679	
Durbin-Watson stat	2.680127	Prob(F-statistic)		0.000001	

6. Durbin-Watson stat: ตัวสถิติที่ใช้ทดสอบ Serial Correlation ของ residual (p.467-472)
7. Mean dependent var: วัดแนวโน้มส่วนกลาง (ค่าเฉลี่ย) ของตัวแปรตาม
8. S.D. dependent var: วัดการกระจาย (ค่าเบี่ยงเบนมาตรฐาน) ของตัวแปรตาม
9. Akaike info criterion: ใช้ในการเลือกแบบจำลอง
10. Schearz criterion: ใช้ในการเลือกแบบจำลอง
11. F-statistic: ใช้ทดสอบสมมติฐานว่าสัมประสิทธิ์ทุกตัวมีค่าเท่ากับศูนย์หรือไม่
12. Prob(F-statistic): ค่า p-value ของ F-statistic

2.3 การสร้างแฟ้มงานสำหรับตัวอย่างการใช้จ่ายเพื่อบริโภคในอินเดีย (ตัวอย่างที่ 3.2)

ตัวอย่าง 3.2 นี้ใช้ข้อมูลที่ปรากฏในตารางที่ 2.8 หน้า 56 ซึ่งมีตัวอย่างที่สุ่มมาจำนวน 55 ครัวเรือน เพื่อสร้างแฟ้มงาน

ขั้นตอนที่ 1. เลือก **File/New/Workfile** ในเมนูหลัก

ขั้นตอนที่ 2. ตั้งค่าความถี่ของแฟ้มงานเป็น **Undated irregular**

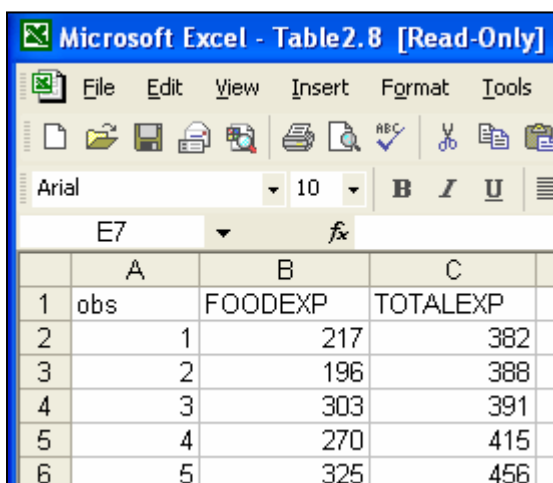
ขั้นตอนที่ 3. ใส่ **Start date (1)** และ **End Date (55)**

ขั้นตอนที่ 4. คลิก **OK**

2.4 การนำเข้าข้อมูลจากโปรแกรม spreadsheet ต่างๆ

เมื่อเราสร้างแฟ้มงานแล้ว เราสามารถนำเข้าข้อมูลจากแฟ้มข้อมูลของโปรแกรม Spreadsheet อื่นๆได้ง่าย แต่อย่างไรก็ตาม เราจำเป็นต้องทราบตำแหน่งของข้อมูลในแฟ้มนั้นๆก่อน

ขั้นตอนที่ 1. เปิดโปรแกรม Excel และเปิดแฟ้มชื่อ **Table2.8.xls**



	A	B	C
1	obs	FOODEXP	TOTALEXP
2	1	217	382
3	2	196	388
4	3	303	391
5	4	270	415
6	5	325	456

ขั้นตอนที่ 2. ข้อมูลตัวเลขชุดแรกอยู่ในเซลล์ **B2** และมีข้อมูลอีกสองคอลัมน์ที่ติดกัน (ดังตัวอย่างข้างล่าง)

ขั้นตอนต่อไปเราก็จะนำข้อมูลดังกล่าวจากแฟ้ม **Table2.8.xls** เข้าสู่แฟ้ม EViews ใหม่ที่สร้างขึ้น

ขั้นตอนที่ 1. ปิดแฟ้ม Excel

ขั้นตอนที่ 2. คลิก **Procs/Import/Read Text-Lotus-Excel** บนเมนูแฟ้มงาน

ขั้นตอนที่ 3. เลือก

drive และ folder
ซึ่งเป็นที่อยู่ของ
แฟ้ม Table2.8.xls

ขั้นตอนที่ 4. เลือก

Excle.xls ในช่อง

Files of type

ขั้นตอนที่ 5. คลิก

สองครั้ง (double
click) ที่

Table2.8.xls

ขั้นตอนที่ 6. ใส่

B2 ลงในช่อง

Upper-left data

cell และ จำนวน

ของ series ซึ่ง

เท่ากับสอง

(EViews จะใส่ชื่อ

Series ดังกล่าว

โดยนำมาจาก แถว

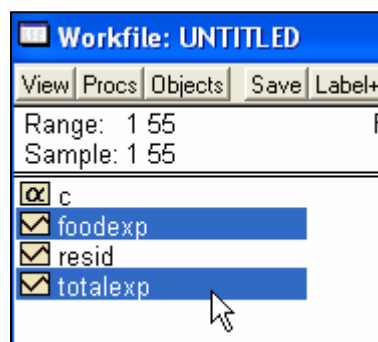
บนของข้อมูล)

ขั้นตอนที่ 7. ช่วงของข้อมูลจะถูกตั้งตามตัวอย่างของแฟ้มงาน

ขั้นตอนที่ 8. คลิก OK เพื่อสิ้นสุดกระบวนการนำเข้าข้อมูล

Run Regression ตามตัวอย่างที่ 3.2

หลังจากนำเข้าข้อมูลจะปรากฏ object อนุกรมใหม่ 2 อนุกรมคือ FoodExp และ TotalExp ซึ่งเราสามารถวิเคราะห์เชิงถดถอยได้โดยการ



ขั้นตอนที่ 1. โดยกด Ctrl ค้างไว้ แล้วคลิก ตัวแปรที่ละตัว เรียงลำดับจากตัวแปรตาม และตัวแปรอิสระ (โดยไม่ต้องใส่ค่าคงที่ เพราะ EViews จะตั้งให้เสมอ)

ขั้นตอนที่ 2. คลิกขวาขณะที่ ลูกศรอยู่บนแท็บที่สีน้ำเงิน เลือก

Open/As Equation เราจะได้หน้าต่าง Equation Specification พร้อม

ตัวแปรที่เราใส่ไว้ และค่าคงที่ (c) แล้วดำเนินขั้นตอนต่อไปตามขั้นตอนที่ 5-7 ในหัวข้อ 2.1 (โดย save ในแฟ้มงานชื่อ indianfoodexp.wfl) จะได้ผลการประมาณค่าเหมือน ตัวอย่างที่ 3.2 หน้า 91

Dependent Variable: FOODEXP				
Method: Least Squares				
Date: 12/03/03 Time: 01:01				
Sample: 1 55				
Included observations: 55				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
TOTALEXP	0.436809	0.078323	5.577047	0.0000
C	94.20878	50.85635	1.852449	0.0695
R-squared	0.369824	Mean dependent var		373.3455
Adjusted R-squared	0.357934	S.D. dependent var		83.43510
S.E. of regression	66.85575	Akaike info criterion		11.27864
Sum squared resid	236893.6	Schwarz criterion		11.35163
Log likelihood	-308.1625	F-statistic		31.10345
Durbin-Watson stat	2.083299	Prob(F-statistic)		0.000001

หรือเราอาจจะเขียนคำสั่ง “LS Foodexp c TotalExp” ในช่องคำสั่งก็ได้ โดยที่ LS คือคำสั่งให้ EViews ประมาณค่าด้วยวิธีการกำลังสองน้อยที่สุด โดยมี Foodexp เป็นตัวแปรตาม และที่เหลือเป็นตัวแปรอธิบาย

2.5 การแสดงผล Actual, Fitted, Residual

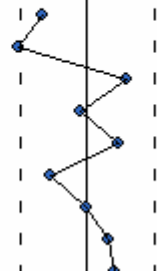
หลังจากที่เราประมาณค่าได้ผลดังตารางข้างต้นแล้ว เราสามารถคำนวณหาค่าที่คาดไว้ของตัวแปรตามที่ขึ้นอยู่กับตัวแปรอิสระ (Conditional Expected value of $Y: \hat{Y}$) ซึ่งใน EViews เรียกว่า ‘Fitted’ แล้วนำมาเปรียบเทียบกับ ค่าตัวแปรตามจริง (actual Y) และหาความแตกต่างระหว่างสองค่านี้ เราจะได้ค่า ‘Residual’

ขั้นตอนที่ 1. คลิก View/Actual, Fitted,

Residual/Actual,Fitted,Residual

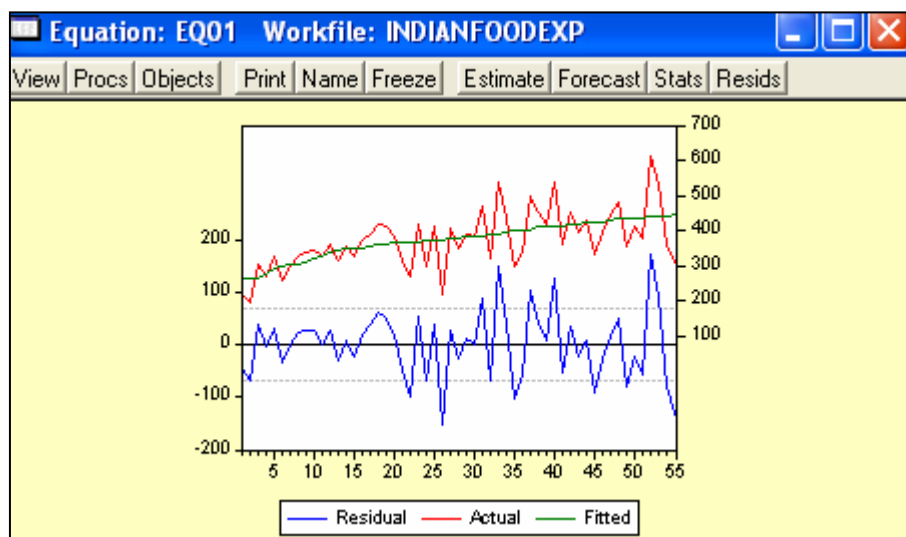
Table บนเมนูของสมการ จะได้ตารางแสดงค่า Actual Y , Fitted $Y(\hat{Y})$ และ Residual ดังรูปข้างล่าง

Equation: EQ01 Workfile: INDIANFOODEXP			
Representations	Freeze	Estimate	Forecast
Estimation Output			
Actual,Fitted,Residual	Actual,Fitted,Residual Table		
Gradients and Derivatives	Actual,Fitted,Residual Graph		
Covariance Matrix	Residual Graph		
	Standardized Residual Graph		
Coefficient Tests			
Residual Tests			
Stability Tests			
Label	Coefficient	Std. Error	t-Statistic
	0.436809	0.078323	5.577047

Equation: EQ01 Workfile: INDIANFOODEXP									
View	Procs	Objects	Print	Name	Freeze	Estimate	Forecast	Stats	Resids
obs	Actual	Fitted	Residual	Residual Plot					
1	217.000	261.070	-44.0697						
2	196.000	263.691	-67.6906						
3	303.000	265.001	37.9990						
4	270.000	275.484	-5.48441						
5	325.000	293.394	31.6064						
6	260.000	295.141	-35.1408						
7	300.000	300.383	-0.38251						
8	325.000	303.003	21.9966						
9	336.000	309.992	26.0077						

ขั้นตอนที่ 2. หากต้องการดูกราฟค่าระหว่าง Actual Y, Fitted $Y(\hat{Y})$ และ Residual เราอาจเลือกคลิก **View/Actual,Fitted,Residual** /**Actual,Fitted,Residual Graph** บนเมนูของสมการ จะได้กราฟแสดงค่า Actual Y, Fitted $Y(\hat{Y})$ และ Residual ดังรูปข้างล่าง

Equation: EQ01 Workfile: INDIANFOODEXP			
View	Procs	Objects	Print Name Freeze Estimate Forecast Stats Resids
Representations			
Estimation Output			
Actual,Fitted,Residual			Actual,Fitted,Residual Table
Gradients and Derivatives			Actual,Fitted,Residual Graph
Covariance Matrix			Residual Graph
			Standardized Residual Graph
Coefficient Tests			
Residual Tests			
Stability Tests			
Label			
			efficient Std. Error t-Statistic
			436809 0.078323 5.5770

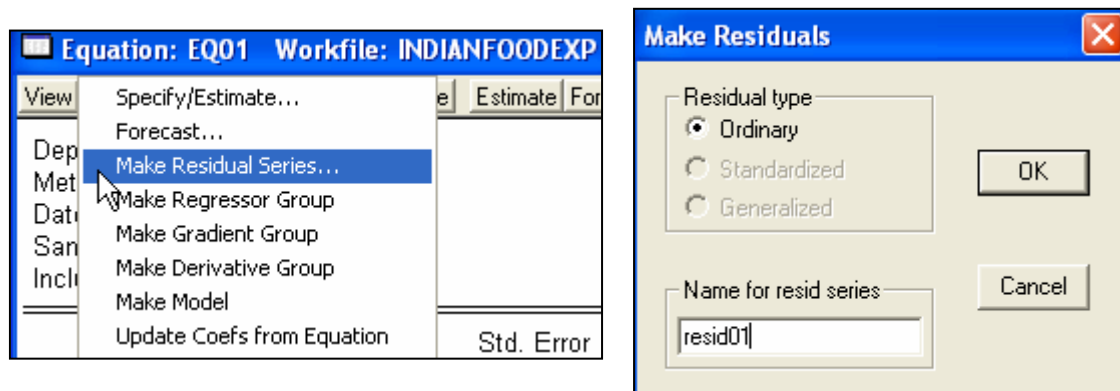


2.6 การสร้างอนุกรมของ Residual

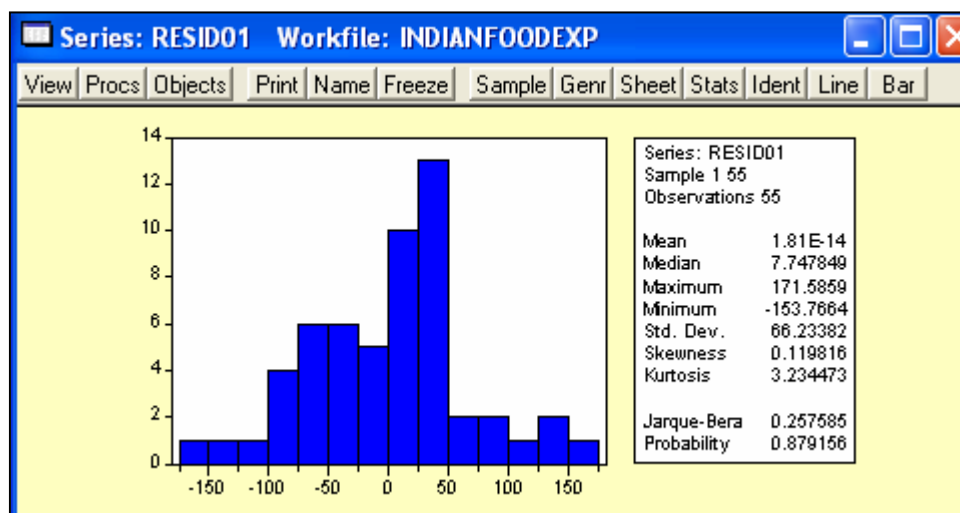
ในการวิเคราะห์ผลและตรวจสอบผลการประมาณค่าที่เราจะศึกษาต่อไปอาจต้องใช้ข้อมูลเกี่ยวกับ Residual ของสมการถดถอยเชิงเส้น ซึ่งอนุกรมดังกล่าวจะเปลี่ยนแปลงไปทุกครั้งเมื่อมีการ

ประมาณค่าสมการใหม่ ดังนั้นเราจึงควรสร้างอนุกรม Residual สำหรับสมการนั้นๆไว้เพื่อประโยชน์ในการวิเคราะห์

ขั้นตอนที่ 1. คลิก Procs/Make Residual Series จะได้หน้าต่าง Make Residuals ขึ้นมา



ขั้นตอนที่ 2. ใส่ชื่ออนุกรม เป็น resid01 (เพื่อให้สอดคล้องกับ EQ01) ในช่อง Name for resid series
จะได้อนุกรมใหม่ ชื่อ Resid01 พร้อมทั้งกราฟแสดงการกระจายและสถิติ ดังรูปในหน้า 150



2.7 การเลือกรูปแบบฟังก์ชัน (Functional form)

ในแบบจำลองที่เราประมาณค่าในตัวอย่างข้างต้นเราสมมุติว่าตัวแปรตามและตัวแปรอธิบายมีความสัมพันธ์เชิงเส้นตรง อย่างไรก็ตามรูปแบบของฟังก์ชันที่เหมาะสมสำหรับการประมาณค่าจะต้องสอดคล้องกับทฤษฎีหรือข้อมูล ดังนั้นการระบุรูปแบบให้โปรแกรมทราบก็มีความแตกต่างกัน ดังตารางข้างล่างซึ่งนำมาจากตารางที่ 6.6

แบบจำลอง	สมการ	EViews specification
Linear	$Y = \beta_1 + \beta_2 X$	Y C X
Log-linear	$\ln Y = \beta_1 + \beta_2 \ln X$	log(Y) C log(X)
Log-lin	$\ln Y = \beta_1 + \beta_2 X$	log(Y) C X
Lin-log	$Y = \beta_1 + \beta_2 \ln X$	Y C log(X)
Reciprocal	$Y = \beta_1 + \beta_2 \left(\frac{1}{X} \right)$	Y C 1/X
Log reciprocal	$\ln Y = \beta_1 - \beta_2 \left(\frac{1}{X} \right)$	log(Y) C 1/X

3. ปัญหา Multicollinearity

3.1 Perfect multicollinearity

โปรแกรม EViews ไม่สามารถประมาณค่าสัมประสิทธิ์ของสมการเชิงถดถอยใดๆที่ตัวแปรอธิบายตั้งแต่สองตัวมีความสัมพันธ์เชิงเส้นอย่างสมบูรณ์ (perfectly collinear) ซึ่งหากเกิดเหตุการณ์ดังกล่าว โปรแกรมจะส่งข้อความออกมาว่า “Near singular matrix”

ในหัวข้อต่อไปเราจะอธิบายวิธีการที่เราใช้ตรวจสอบความรุนแรงของปัญหา multicollinearity โดยใช้ตัวอย่างในหัวข้อ 10.10 (หน้า 370, Gujarati) ซึ่งใช้ข้อมูลจากตารางที่ 10.7

3.2 การตรวจสอบปัญหา multicollinearity ด้วย simple correlation coefficients

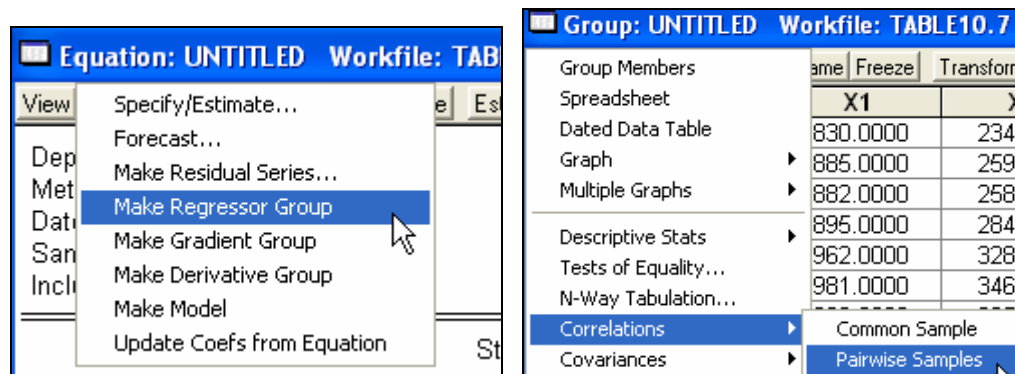
จากหัวข้อดังกล่าว เราวิเคราะห์ผลของตัวแปรอธิบายต่างๆ ($X_1, X_2, X_3, X_4, X_5, \text{Time}$) ต่อการจ้างงาน (Y) โดยใช้สมการเชิงถดถอย

ขั้นตอนที่ 1. เปิด workfile ที่มีข้อมูลดังกล่าวอยู่ชื่อ ‘table10.7.wfl’ แล้ว run regression ระหว่าง Y กับค่าคงที่, $X_1, X_2, X_3, X_4, X_5, \text{Time}$ ได้ผลดังตารางข้างล่าง

Dependent Variable: Y Method: Least Squares Date: 01/26/04 Time: 01:13 Sample: 1947 1962 Included observations: 16				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	1.506187	8.491493	0.177376	0.8631
X2	-0.035819	0.033491	-1.069516	0.3127
X3	-2.020230	0.488400	-4.136427	0.0025
X4	-1.033227	0.214274	-4.821985	0.0009
X5	-0.051104	0.226073	-0.226051	0.8262
TIME	1829.151	455.4785	4.015890	0.0030
C	77270.12	22506.71	3.433204	0.0075
R-squared	0.995479	Mean dependent var	65317.00	
Adjusted R-squared	0.992465	S.D. dependent var	3511.968	
S.E. of regression	304.8541	Akaike info criterion	14.57718	
Sum squared resid	836424.1	Schwarz criterion	14.91519	
Log likelihood	-109.6174	F-statistic	330.2853	
Durbin-Watson stat	2.559488	Prob(F-statistic)	0.000000	

จากตารางดังกล่าวเราสังเกตเห็นว่า ค่า R^2 ค่อนข้างสูง แต่ค่า t-stat ของตัวแปรบางตัวเช่น X_1 , X_2 , X_5 ไม่มีนัยสำคัญ ดังนั้นเราจึงต้องการทดสอบเบื้องต้นว่ามีปัญหา multicollinearity หรือไม่ โดยใช้ pair-wise correlation

ขั้นตอนที่ 2. สร้างกลุ่มของตัวแปรอธิบายโดยการเลือก Proc/Make Regressor Group ในเมนูของสมการ จะได้กลุ่ม (group) ของตัวแปรอธิบาย



ขั้นตอนที่ 3. เลือก View/Correlation/Pairwise Samples จากเมนูของกลุ่ม จะได้ตาราง Correlation Matrix ดังตารางข้างล่างเหมือนกับตารางที่ 10.8

Correlation Matrix							
	Y	X1	X2	X3	X4	X5	TIME
Y	1.000000	0.970899	0.983552	0.502498	0.457307	0.960391	0.971329
X1	0.970899	1.000000	0.991589	0.620633	0.464744	0.979163	0.991149
X2	0.983552	0.991589	1.000000	0.604261	0.446437	0.991090	0.995273
X3	0.502498	0.620633	0.604261	1.000000	-0.177421	0.686552	0.668257
X4	0.457307	0.464744	0.446437	-0.177421	1.000000	0.364416	0.417245
X5	0.960391	0.979163	0.991090	0.686552	0.364416	1.000000	0.993953
TIME	0.971329	0.991149	0.995273	0.668257	0.417245	0.993953	1.000000

ขั้นตอนที่ 4. เลือก Freeze ในเมนูของกลุ่ม จะได้ object ตาราง (table) แล้วเลือก Name เพื่อจัดเก็บตารางดังกล่าว

4. ปัญหา Heteroscedasticity

เราจะอธิบายการตรวจสอบและการบรรเทาปัญหา heteroscedasticity โดยใช้ตัวอย่างในบทที่ 11 ใน Gujarati(2003) โดยเริ่มด้วยตัวอย่างที่ 11.11 ซึ่งใช้ข้อมูลในตารางที่ 11.1 (table11.1.wfl)

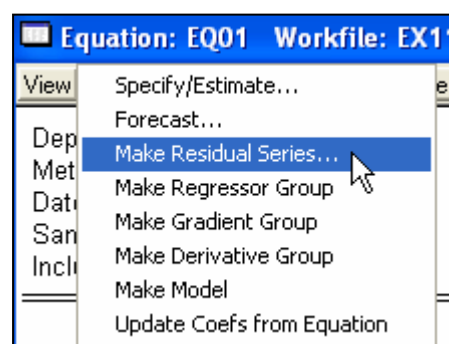
ตัวอย่างดังกล่าวศึกษาความสัมพันธ์ระหว่าง AVECOMP กับ ค่าคงที่ และ AVEPROD ได้ผลการประมาณค่าดังตารางข้างล่าง

Dependent Variable: AVECOMP Method: Least Squares Date: 01/26/04 Time: 01:49 Sample: 1 9 Included observations: 9				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1992.062	936.6123	2.126880	0.0710
AVEPROD	0.232999	0.099853	2.333428	0.0523
R-squared	0.437520	Mean dependent var	4161.778	
Adjusted R-squared	0.357166	S.D. dependent var	420.6625	
S.E. of regression	337.2744	Akaike info criterion	14.67280	
Sum squared resid	796278.0	Schwarz criterion	14.71663	
Log likelihood	-64.02760	F-statistic	5.444885	
Durbin-Watson stat	0.616592	Prob(F-statistic)	0.052349	

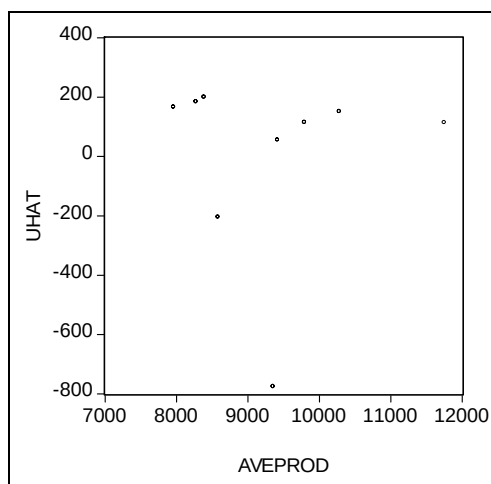
4.1 การตรวจสอบปัญหา Heteroscedasticity ด้วยรูปภาพ

ขั้นตอนที่ 1. หลังจากนั้นเราก็สร้างชุดข้อมูลของ residual โดยเลือก **Proc/Make Residual Series..** ในเมนูของสมการ แล้ว save อยู่ในชื่อ 'UHAT'

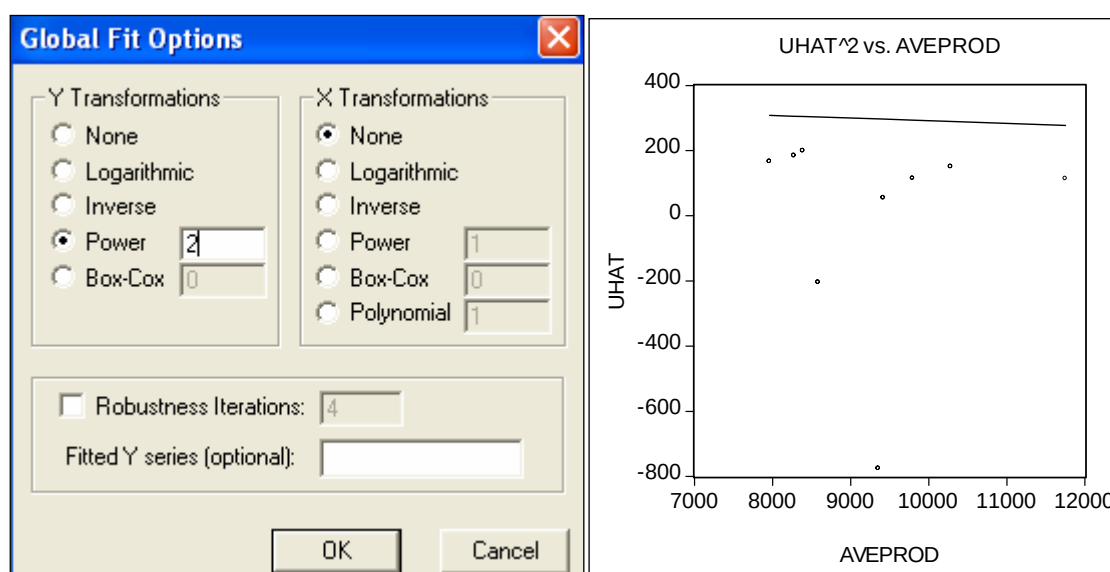
ขั้นตอนที่ 2. สร้างกลุ่มข้อมูลระหว่าง UHAT กับ AVEPROD



ขั้นตอนที่ 3. สร้าง scatter diagram โดยเลือก **View/Graph/Scatter/Simple Scatter** ได้ scatter diagram เพื่อตรวจสอบปัญหา heteroscedasticity ดังรูป



ขั้นตอนที่ 4. หากต้องการให้ UHAT อยู่ในรูปยกกำลังสองเราอาจเลือก **View/Graph/Scatter/Scatter with regression** แล้วเลือก Y Transformations เป็น Power 2 (หรือ $UHAT^2$ นั่นเอง)



4.2 การตรวจสอบปัญหา Heteroscedasticity ด้วย Park test (จากตัวอย่างที่ 11.1)

ขั้นตอนที่ 1. ประมาณค่าสมการถดถอยเชิงเส้น และสร้างชุดข้อมูล residual ตามขั้นตอนที่ 1 ในหัวข้อ 4.1

ขั้นตอนที่ 2. เลือก **Objects/New Objects/Equation** ในเมนูหลัก แล้วใส่รูปแบบของสมการที่ใช้ทดสอบ คือ ' $\log(UHAT^2) C \log(AVEPROD)$ ' ใน Equation Specification

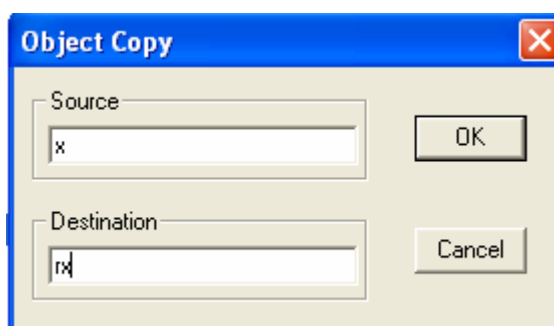
ขั้นตอนที่ 3. ทดสอบนัยสำคัญของค่าสัมประสิทธิ์ความชันสมการดังกล่าว

Dependent Variable: LOG(UHAT^2)				
Method: Least Squares				
Date: 01/26/04 Time: 01:52				
Sample: 1 9				
Included observations: 9				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	35.82684	38.32272	0.934872	0.3810
LOG(AVEPROD)	-2.802022	4.196134	-0.667763	0.5257
R-squared	0.059886	Mean dependent var	10.23844	
Adjusted R-squared	-0.074416	S.D. dependent var	1.414819	
S.E. of regression	1.466517	Akaike info criterion	3.796787	
Sum squared resid	15.05470	Schwarz criterion	3.840615	
Log likelihood	-15.08554	F-statistic	0.445907	
Durbin-Watson stat	1.137101	Prob(F-statistic)	0.525681	

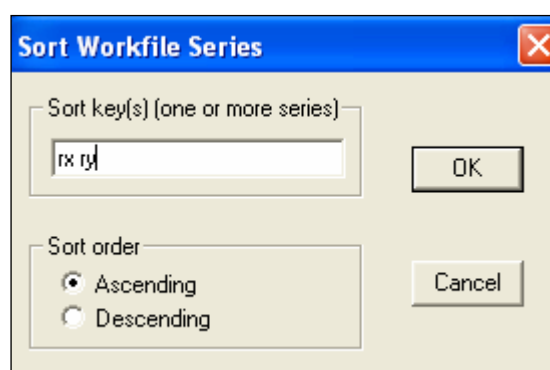
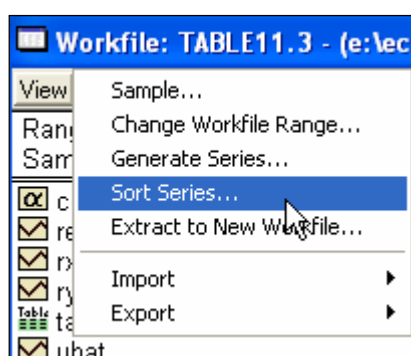
4.3 การตรวจสอบปัญหา Heteroscedasticity ด้วย Goldfeld-Quandt test (ตัวอย่างที่ 11.4)

ใช้ข้อมูลตารางที่ 11.3 โดยเปิด workfile ชื่อ 'table11.3.wfl' แล้วดำเนินการทดสอบดังนี้

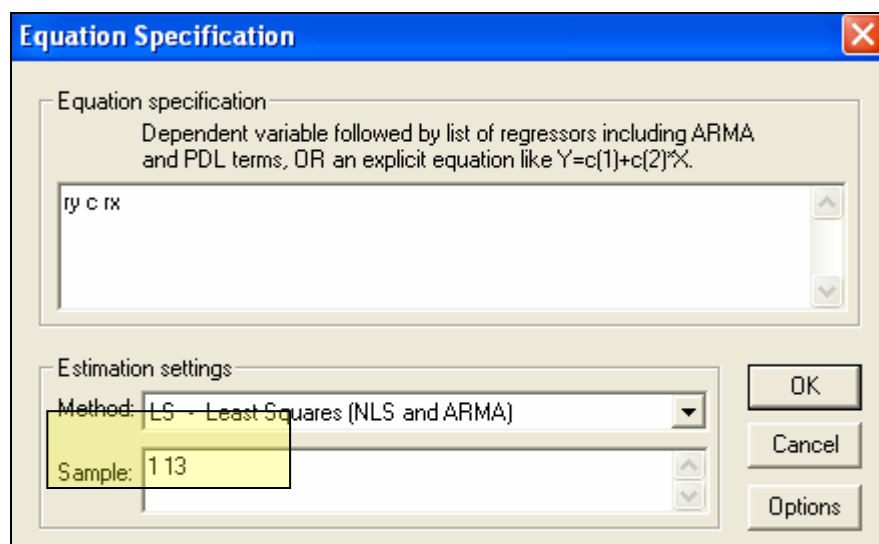
ขั้นตอนที่ 1. สร้างข้อมูลชุดใหม่เพื่อที่จะรักษาข้อมูลชุดเดิมไว้ โดยคลิกขวาที่ตัวแปร X แล้วเลือก **Object copy..** เพื่อ copy เป็นข้อมูลชุดโดยใส่ RX ในช่อง Destination (แล้วทำเช่นเดียวกันกับข้อมูล Y)



ขั้นตอนที่ 2. เรียงข้อมูลตามตัวแปรอธิบายคือ 'RX' โดยเลือก **Proc/Sort Series...** ในเมนู workfile แล้วใส่ชุดข้อมูลที่เราต้องการเรียง โดยใส่ชุดข้อมูลที่เป็นหลักก่อน เช่น rx ry (เรียงตาม ตาม rx) จากนั้นก็คลิก OK



ขั้นตอนที่ 3. ประมาณค่าสมการเชิงถดถอยระหว่าง RY กับ C และ RX ของข้อมูลชุดที่ความแปรปรวนน้อย โดยในช่อง Sample เราจำกัดเพียงตัวอย่างที่ 1-13 เราจะได้ค่า RSS_1 จาก ช่อง Sum squared resid = 377.1663



ขั้นตอนที่ 3. ประมาณค่าสมการเชิงถดถอยระหว่าง RY กับ C และ RX ของข้อมูลชุดที่ความแปรปรวนมาก โดยในช่อง Sample เราจำกัดเพียงตัวอย่างที่ 18-30 เราจะได้ค่า RSS_2 จาก ช่อง Sum squared resid = 1536.8

ขั้นตอนที่ 4. คำนวณค่าสถิติ $\lambda = \frac{RSS_2 / df}{RSS_1 / df}$ แล้วทดสอบด้วย F-test

4.4 การตรวจสอบปัญหา Heteroscedasticity ด้วย White's test

ใช้ข้อมูลจากตารางที่ 11.3 ใน Gujarati

ขั้นตอนที่ 1. ประมาณค่าสมการถดถอยเชิงเส้น ระหว่าง Y กับ C และ X

ขั้นตอนที่ 2. เพื่อตรวจสอบ

โดยวิธีของ White เลือก

View/Residual Tests/White

Heteroskedasticity(cross

term) ในเมนูของสมการ

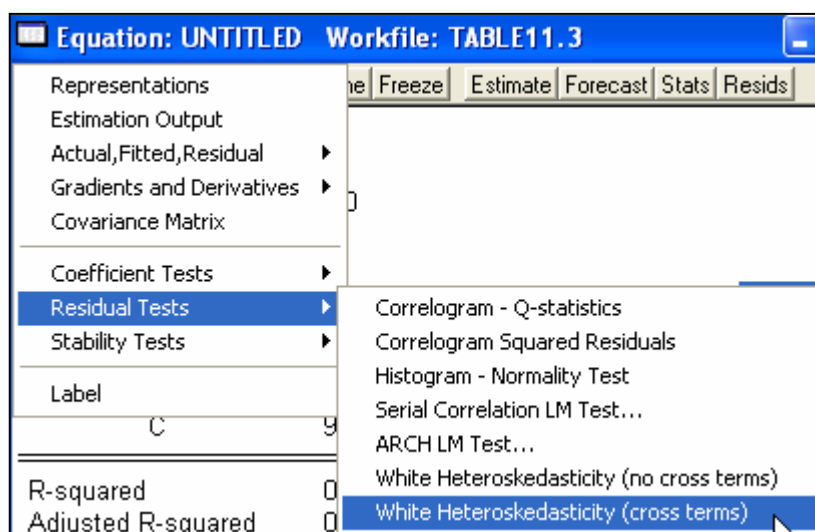
[ในการทดสอบดังกล่าว

สามารถเลือกรูปแบบได้ 2

รูปแบบ คือ มีพจน์ที่ตัวแปร

อธิบายคุณกัน (cross term)

หรือไม่มีก็ได้ (no cross term)]



โดยผลการทดสอบแสดงในตารางข้างล่าง

White Heteroskedasticity Test:				
F-statistic	2.917301	Probability	0.071274	
Obs*R-squared	5.330902	Probability	0.069568	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 01/28/04 Time: 01:27				
Sample: 1 30				
Included observations: 30				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-12.29621	191.7731	-0.064119	0.9493
X	0.197385	2.368760	0.083329	0.9342
X^2	0.001700	0.006707	0.253503	0.8018
R-squared	0.177697	Mean dependent var	78.70511	
Adjusted R-squared	0.116785	S.D. dependent var	112.5823	
S.E. of regression	105.8043	Akaike info criterion	12.25570	
Sum squared resid	302252.7	Schwarz criterion	12.39582	
Log likelihood	-180.8355	F-statistic	2.917301	
Durbin-Watson stat	0.791307	Prob(F-statistic)	0.071274	

ขั้นตอนที่ 4. นำค่าตัวสถิติ nR^2 ซึ่ง EViews ได้คำนวณไว้ในแถบสีฟ้า เพื่อในการทดสอบ White กับ critical χ^2 ซึ่งค่าดังกล่าวสามารถหาได้โดยใช้ EViews โดยพิมพ์คำสั่ง “=@qchisq(.95,2)” ในช่องคำสั่งสีขาว โดยที่ .95 คือ 1- ระดับนัยสำคัญที่ต้องการ และ 2 คือ degree of freedom (จำนวนตัวแปรอธิบายในสมการที่ใช้ทดสอบ White) จะได้ค่า critical χ^2 ในมุมซ้ายล่างของหน้าจอ  พบว่าค่า $nR^2 < \text{critical } \chi^2$ จึงไม่สามารถปฏิเสธสมมติฐานที่ว่า “ไม่มีปัญหา heteroscedasticity ได้” หรือเราสามารถตรวจสอบได้จากค่า Probability ก็ได้ พบว่า ค่า P-value = 0.069568 ($>$ นัยสำคัญที่เราเลือกคือ 0.05)

4.5 การบรรเทาปัญหา Heteroscedasticity ด้วย Weighted Least Square(WLS) (ตัวอย่างที่ 11.7)

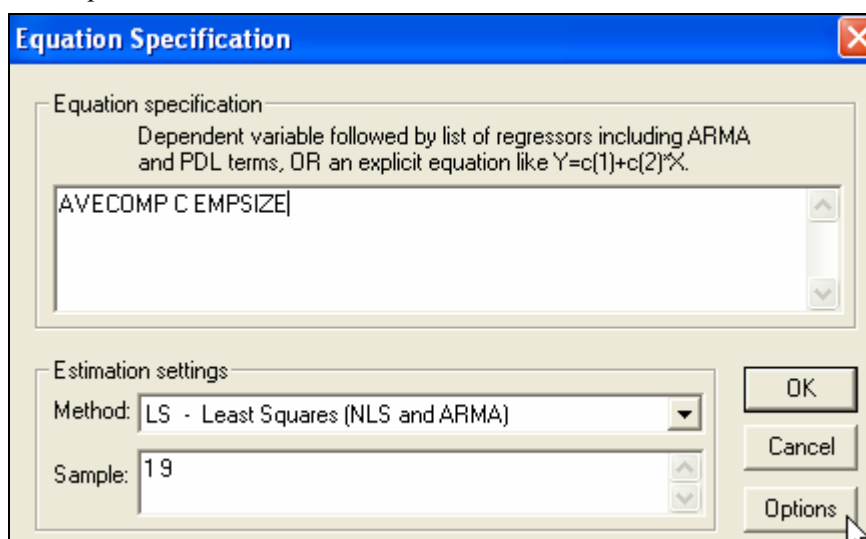
ใช้ข้อมูลตารางที่ 11.1 เพื่อประมาณค่าสมการระหว่างค่าจ้างเฉลี่ย(AVECOMP) กับขนาดแรงงาน(EXPSIZE) ในกรณีดังกล่าวเราทราบว่าปัญหา heteroscedasticity และทราบค่าความเบี่ยงเบนมาตรฐานของแต่ละตัวอย่างเราสามารถบรรเทาปัญหาโดยใช้ WLS ซึ่งสามารถสั่งโปรแกรมได้สองวิธีคือ

วิธีที่หนึ่ง ในการสร้างสมการถดถอย ในช่อง Equation Specification ใส่รูปแบบเหมือนสมการที่ 11.6.1 คือ ‘AVECOMP/STDEV 1/STDEV EMPSIZE/STDEV’ แล้วคลิก OK จะได้ผลดังตารางข้างล่าง

Dependent Variable: AVECOMP/STDEV				
Method: Least Squares				
Date: 01/28/04 Time: 01:55				
Sample: 1 9				
Included observations: 9				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
1/STDEV	3406.213	80.92452	42.09123	0.0000
EMPSIZE/STDEV	154.2425	16.94475	9.102671	0.0000
R-squared	0.964435	Mean dependent var	4.376112	
Adjusted R-squared	0.959354	S.D. dependent var	0.670698	
S.E. of regression	0.135218	Akaike info criterion	-0.970720	
Sum squared resid	0.127988	Schwarz criterion	-0.926892	
Log likelihood	6.368239	Durbin-Watson stat	1.184707	

วิธีการที่สอง

ขั้นตอนที่ 1. ในช่อง Equation Specification ก็ใส่ตัวแปรตามปกติ คือ AVECOMP C EMPSIZE แล้วเลือก Options



ขั้นตอนที่ 2. จะปรากฏหน้าต่าง Equation Options ขึ้นมา ให้คลิกที่ช่อง ☒ Weighted LS/TLS แล้วใส่ตัวถ่วงน้ำหนัก ซึ่งก็คือ '1/STDEV' ในช่อง Weight แล้วคลิก OK จะทำให้กลับมาสู่หน้าต่าง Equation Specification แล้วคลิก OK อีกครั้ง จะได้ผลการประมาณค่าดังตารางข้างล่าง ซึ่งค่าสัมประสิทธิ์ของตัวแปรอธิบายก็จะเหมือนกับวิธีการที่หนึ่ง

Estimation Options
✕

LS & TSLS options

☐ Heteroskedasticity Consistent Coefficient Covariance

☒ White
☐ Newey-West

☒ Weighted LS/TSLS
(not available with ARMA)

Weight:

Iteration control

Max Iterations:

Convergence:

☐ Display settings

ARMA options

Starting coefficient values

☒ Backcast MA terms

Derivatives

Select method to favor:

☒ Accuracy
☐ Speed

☐ Use numeric only

Dependent Variable: AVECOMP
 Method: Least Squares
 Date: 01/28/04 Time: 02:05
 Sample: 1 9
 Included observations: 9
 Weighting series: 1/STDEV

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3406.213	80.92452	42.09123	0.0000
EMPSIZE	154.2425	16.94475	9.102671	0.0000

Weighted Statistics

R-squared	0.964435	Mean dependent var	4098.466
Adjusted R-squared	0.959354	S.D. dependent var	628.1452
S.E. of regression	126.6394	Akaike info criterion	12.71370
Sum squared resid	112262.8	Schwarz criterion	12.75752
Log likelihood	-55.21163	F-statistic	82.85861
Durbin-Watson stat	1.184707	Prob(F-statistic)	0.000040

Unweighted Statistics

R-squared	0.935613	Mean dependent var	4161.778
Adjusted R-squared	0.926415	S.D. dependent var	420.6625
S.E. of regression	114.1114	Sum squared resid	91149.92
Durbin-Watson stat	1.142176		

4.6 การบรรเทาปัญหา Heteroscedasticity ด้วย White's heteroscedasticity corrected standard error

ขั้นตอนที่ 1. ประมาณค่าสมการถดถอยโดยในช่อง Equation Specification ก็ใส่ตัวแปรตามปกติ คือ AVECOMP C EMPSIZE แล้วเลือก Options

☒ Heteroskedasticity Consistent Coefficient Covariance
☒ White
☐ Newey-West

ขั้นตอนที่ 2. จะปรากฏหน้าต่าง Equation Options ขึ้นมา ให้คลิกที่ช่อง และเลือกวิธีการของ White แล้วคลิก OK จะทำให้กลับมาสู่หน้าต่าง Equation Specification แล้วคลิก OK อีกครั้ง จะได้ผลการประมาณค่าดังตารางข้างล่าง

Dependent Variable: AVECOMP Method: Least Squares Date: 01/28/04 Time: 02:17 Sample: 1 9 Included observations: 9 White Heteroskedasticity-Consistent Standard Errors & Covariance				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3417.778	107.1010	31.91173	0.0000
EMPSIZE	148.8000	16.86644	8.822253	0.0000
R-squared	0.938425	Mean dependent var	4161.778	
Adjusted R-squared	0.929628	S.D. dependent var	420.6625	
S.E. of regression	111.5918	Akaike info criterion	12.46070	
Sum squared resid	87169.16	Schwarz criterion	12.50453	
Log likelihood	-54.07316	F-statistic	106.6823	
Durbin-Watson stat	1.223659	Prob(F-statistic)	0.000017	

เมื่อเปรียบเทียบค่า standard deviation (s.d.) กับสมการที่ไม่ได้มีการแก้ไขปัญหา heteroscedasticity (ตารางข้างล่าง) พบว่าค่าสัมประสิทธิ์เท่ากัน แต่ s.d. กรณีไม่ได้แก้ไขมีค่าน้อยกว่าทำให้ค่า t สูง ในขณะที่ s.d. ของ White's heteroscedasticity corrected standard error มีค่าสูงขึ้นและลดขนาดของค่า t ทำให้การทดสอบน่าเชื่อถือขึ้น (ตัวอย่างนี้อาจเห็นไม่ค่อยชัดเจนเนื่องจากค่า t สูงด้วยกันทั้งคู่)

Dependent Variable: AVECOMP Method: Least Squares Date: 01/28/04 Time: 02:16 Sample: 1 9 Included observations: 9				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3417.778	81.06958	42.15857	0.0000
EMPSIZE	148.8000	14.40644	10.32871	0.0000
R-squared	0.938425	Mean dependent var	4161.778	
Adjusted R-squared	0.929628	S.D. dependent var	420.6625	
S.E. of regression	111.5918	Akaike info criterion	12.46070	
Sum squared resid	87169.16	Schwarz criterion	12.50453	
Log likelihood	-54.07316	F-statistic	106.6823	
Durbin-Watson stat	1.223659	Prob(F-statistic)	0.000017	

5. ปัญหา Autocorrelation

เราจะอธิบายการตรวจสอบและการบรรเทาปัญหา Autocorrelation โดยใช้ตัวอย่างในบทที่ 12 ใน Gujarati (2003) โดยเริ่มด้วยตัวอย่างในหัวข้อ 12.5 (หน้า 460) ซึ่งใช้ข้อมูลในตารางที่ 12.4 (table12.4.wf1)

ตัวอย่างดังกล่าวศึกษาความสัมพันธ์ระหว่าง Real Compensation(Y) กับ ค่าคงที่ และ Productivity(X) ได้ผลการประมาณค่าดังตารางข้างล่าง

Dependent Variable: Y Method: Least Squares Date: 02/04/04 Time: 02:34 Sample: 1959 1998 Included observations: 40				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	29.51925	1.942347	15.19773	0.0000
X	0.713659	0.024105	29.60658	0.0000
R-squared	0.958449	Mean dependent var	85.64500	
Adjusted R-squared	0.957356	S.D. dependent var	12.95632	
S.E. of regression	2.675533	Akaike info criterion	4.854881	
Sum squared resid	272.0220	Schwarz criterion	4.939325	
Log likelihood	-95.09761	F-statistic	876.5495	
Durbin-Watson stat	0.122904	Prob(F-statistic)	0.000000	

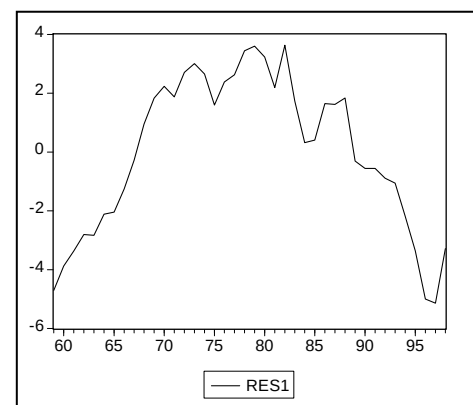
5.1 การตรวจสอบปัญหา Autocorrelation ด้วยรูปภาพ

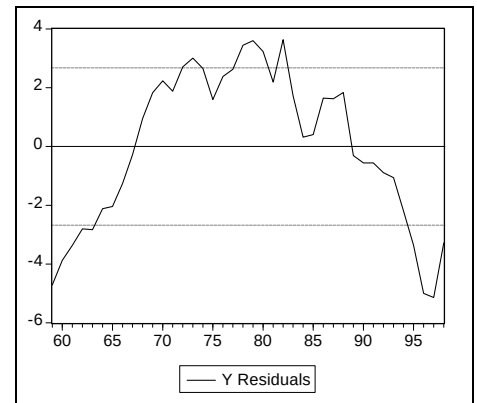
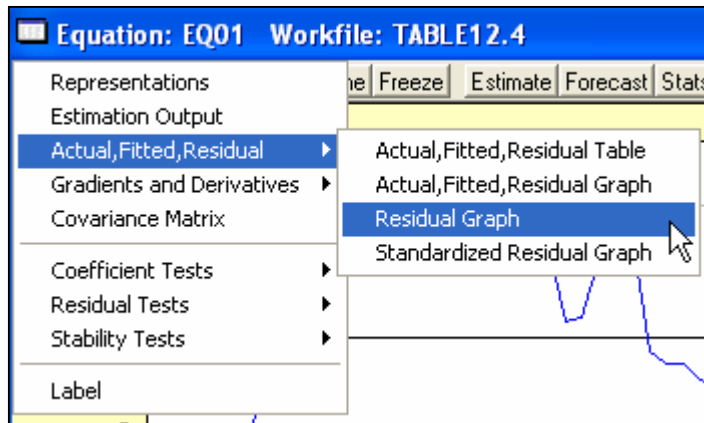
* การ plot residual กับเวลา

ขั้นตอนที่ 1. หลังจากนั้นเราก็สร้างชุดข้อมูลของ residual โดยเลือก **Proc/Make Residual Series..** ในเมนูของสมการ แล้ว save อยู่ในชื่อ 'res1' เราจะได้กราฟแสดงค่า residual ของแต่ละตัวอย่าง ดังรูปข้างขวามือ(เหมือนกับรูปที่ 12.8)

หรือเราอาจจะเลือก **View/Actual,Fitted,Residual/Residual**

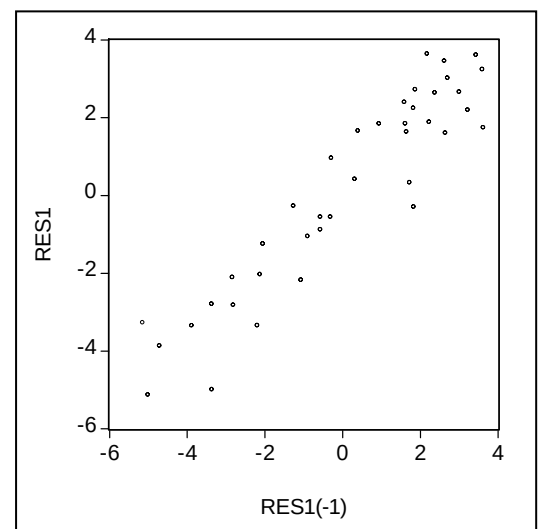
Graph ในเมนูของสมการ จะได้รูปข้างล่างซึ่งเหมือนกับรูปที่ 12.8





* การ plot residual กับ residual ในปีที่ติดกัน

เราอาจ plot กราฟระหว่าง residual ด้วยกันได้โดยเขียนคำสั่งในช่องคำสั่ง “**graph gr1.scad res1(-1) res1**” โดยที่คำสั่งดังกล่าวหมายความว่า “[สร้าง graph] [ชื่อกราฟ].[ชนิดกราฟ] [ตัวแปร] [ตัวแปร]” จะได้กราฟดังรูปขวามือ (เหมือนกับรูปที่ 12.9)



5.2 การตรวจสอบปัญหา Autocorrelation ด้วย Durbin-Watson d stat

ในผลการประมาณค่าสมการถดถอย จะคำนวณค่า Durbin-Watson d stat ให้เสมอในช่อง Durbin-Watson stat ในตัวอย่างข้างต้นมีค่า Durbin-Watson stat เท่ากับ 0.122904 แล้วนำมาเปรียบเทียบกับ critical d_L กับ d_U ในตาราง Durbin-Watson โดยที่ $n=40$, k (จำนวนตัวแปรอธิบาย)=1 พบว่า $d_L=1.44$ กับ $d_U=1.54$ จากค่า DW d stat = 0.1229 < $d_L(1.44)$ เราสามารถปฏิเสธสมมุติฐานว่างที่ว่า “ไม่มี positive autocorrelation”

5.3 การตรวจสอบปัญหา Autocorrelation ด้วย Breusch-Godfrey (BG) Test

ขั้นตอนที่ 1. ประมาณค่าสมการถดถอย แล้วสร้าง residual ชื่อว่า “**res1**”

ขั้นตอนที่ 2. ประมาณค่าสมการถดถอยระหว่าง residual กับ ตัวแปรอธิบาย และ lagged ของ residual โดยใส่ “**res1 c x res1(-1 to -6)**” ในช่อง Equation Specification [(-1 to -6) หมายถึง lagged ที่ 1 ถึง 6] จะได้ผลการประมาณค่าดังตารางข้างล่าง

ขั้นตอนที่ 3. คำนวณ

ค่าสถิติ $(n-p)R^2 = (40-$ $6) \times 0.8920 = 30.328$

แล้วนำไป

เปรียบเทียบกับตัว

สถิติ $\chi_p^2 (=18.5476$

ที่นัยสำคัญเท่ากับ

0.005) ดังนั้น เรา

ปฏิเสธสมมติฐาน

ที่ว่าค่า coefficient of

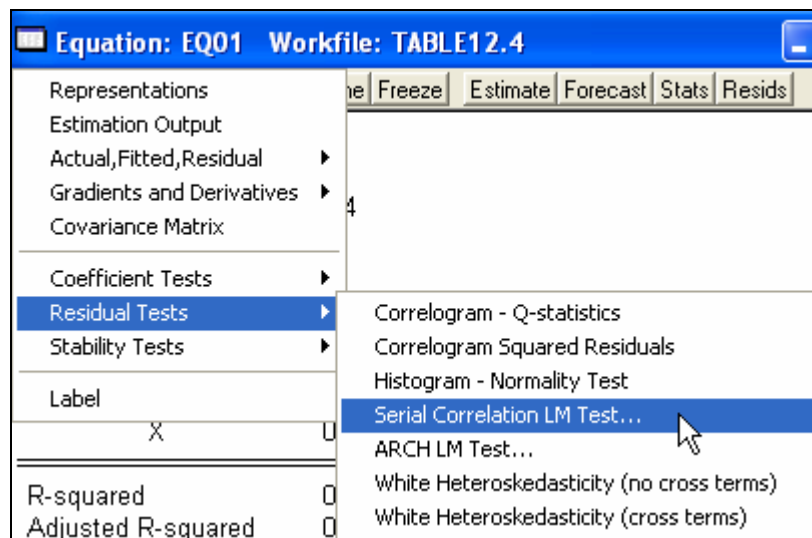
covariance (ρ) ทุกตัว

เท่ากับศูนย์ หรือมีค่า

coefficient of covariance บางตัวไม่เท่ากับศูนย์นั่นเอง

Dependent Variable: RES1 Method: Least Squares Date: 02/05/04 Time: 21:26 Sample(adjusted): 1965 1998 Included observations: 34 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.590462	1.963603	2.847043	0.0085
X	-0.066605	0.023469	-2.838058	0.0087
RES1(-1)	0.814971	0.216231	3.768978	0.0009
RES1(-2)	-0.268651	0.273887	-0.980882	0.3357
RES1(-3)	-0.106017	0.272780	-0.388652	0.7007
RES1(-4)	0.305630	0.273258	1.118467	0.2736
RES1(-5)	-0.064375	0.280577	-0.229438	0.8203
RES1(-6)	0.216156	0.222160	0.972976	0.3395
R-squared	0.892012	Mean dependent var	0.578821	
Adjusted R-squared	0.862938	S.D. dependent var	2.411439	

นอกจากนี้ในโปรแกรม EViews ได้สร้างคำสั่งที่ใช้ทดสอบปัญหา Autocorrelation โดยที่เราสามารถเลือก **View/Residual Tests/Serial Correlation LM Test...**



จะได้ผลตามตารางข้างล่าง ซึ่งได้คำนวณตัวสถิติ $(n-p)R^2$ และค่า P-value

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	23.01054	Probability	0.000000	
Obs*R-squared	32.47339	Probability	0.000013	
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Date: 02/05/04 Time: 21:28				
Presample missing value lagged residuals set to zero.				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.140995	1.007706	0.139917	0.8896
X	-0.002318	0.012893	-0.179767	0.8585
RESID(-1)	1.003107	0.182609	5.493200	0.0000
RESID(-2)	-0.088115	0.257674	-0.341963	0.7346
RESID(-3)	-0.076036	0.258113	-0.294585	0.7702
RESID(-4)	0.199826	0.258333	0.773520	0.4449
RESID(-5)	-0.107547	0.261060	-0.411963	0.6831
RESID(-6)	-0.059385	0.198127	-0.299734	0.7663
R-squared	0.811835	Mean dependent var	-5.76E-15	
Adjusted R-squared	0.770674	S.D. dependent var	2.641008	

5.4 การบรรเทาปัญหาวิธีการ First-Difference Method

ภายใต้ข้อสมมุติที่ว่า error term มีลักษณะเป็น AR(1) เราสามารถบรรเทาปัญหา Autocorrelation ได้โดยประมาณค่าสมการ $\Delta Y_t = \beta_2 \Delta X_t + \varepsilon_t$ ดังสมการ (12.9.7) สำหรับการใส่รูปแบบของ function ใน coefficient of autocovariance ในโปรแกรม EViews d() จะหมายถึง difference ของตัวแปรดังกล่าว] จะได้ผลการประมาณค่าดังตารางข้างล่าง เหมือนดังสมการ 12.9.9 และค่า Durbin-Watson stat เพิ่มขึ้น จนอยู่ในช่วงที่เราไม่สามารถปฏิเสธสมมุติฐานว่างที่ว่า “ไม่มี positive autocorrelation”

Dependent Variable: D(Y)				
Method: Least Squares				
Date: 02/05/04 Time: 21:39				
Sample(adjusted): 1960 1998				
Included observations: 39 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(X)	0.719956	0.078194	9.207333	0.0000
R-squared	0.361092	Mean dependent var	1.194872	
Adjusted R-squared	0.361092	S.D. dependent var	1.173378	
S.E. of regression	0.937901	Akaike info criterion	2.734962	
Sum squared resid	33.42701	Schwarz criterion	2.777617	
Log likelihood	-52.33175	Durbin-Watson stat	1.509651	

5.5 การบรรเทาปัญหา Autocorrelation ด้วย Generalized Least Square (GLS) โดยใช้ coefficient of autocovariance (ρ) จาก Residual

ขั้นตอนการประมาณค่าโดยใช้ GLS คือ

- 1) แปลงค่าตัวแปรโดยคำนึงถึงปัญหา Autocorrelation (หากไม่ทราบค่า coefficient of autocovariance (ρ) ก็ต้องประมาณขึ้นมา
- 2) ประมาณค่าตัวแปรที่แปลงค่าด้วย OLS

เพื่อบรรเทาปัญหา Autocorrelation เราสามารถใช้การประมาณค่าด้วย Generalized Least Square (GLS) แต่เราไม่ทราบค่า coefficient of autocovariance (ρ) เราจำเป็นต้องคำนวณค่าดังกล่าวก่อน ซึ่งในข้อนี้ใช้การประมาณค่าของ

residual

ขั้นตอนที่ 1. ประมาณค่าสมการถดถอยด้วยรูปแบบ “res1 res1(-1)” ได้ผลตามตารางข้างขวามือ และ $\rho = 0.914245$

ขั้นตอนที่ 2. แปลงตัวแปร X และ Y ให้อยู่ในรูป $Y^* = (Y_t - \rho Y_{t-1})$ และ $X^* = (X_t - \rho X_{t-1})$

โดยเลือก Quick/Generate Series... ใส่ “YSTAR=Y-0.914245*Y(-1)” ลงในช่องว่างแล้วคลิก OK จะได้

series YSTAR ขึ้นมา (สร้าง series XSTAR ด้วยวิธีการเดียวกัน)

ขั้นตอนที่ 3. ประมาณค่าสมการถดถอยระหว่าง YSTAR กับ XSTAR จะได้ผลการประมาณค่าดังตารางข้างขวามือ และจากการสังเกต Durbin-Watson stat พบว่าสมการดังกล่าวไม่มีปัญหา Autocorrelation แล้ว

Dependent Variable: RES1 Method: Least Squares Date: 02/05/04 Time: 21:41 Sample(adjusted): 1960 1998 Included observations: 39 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
RES1(-1)	0.914245	0.056337	16.22811	0.0000
R-squared	0.873615	Mean dependent var	0.120615	
Adjusted R-squared	0.873615	S.D. dependent var	2.561492	
S.E. of regression	0.910629	Akaike info criterion	2.675945	
Sum squared resid	31.51134	Schwarz criterion	2.718600	
Log likelihood	-51.18093	Durbin-Watson stat	1.472987	

Dependent Variable: YSTAR Method: Least Squares Date: 02/05/04 Time: 21:51 Sample(adjusted): 1960 1998 Included observations: 39 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.108209	0.656933	6.253621	0.0000
XSTAR	0.528896	0.077424	6.831186	0.0000
R-squared	0.557761	Mean dependent var	8.496580	
Adjusted R-squared	0.545808	S.D. dependent var	1.273241	
S.E. of regression	0.858085	Akaike info criterion	2.581693	
Sum squared resid	27.24346	Schwarz criterion	2.667004	
Log likelihood	-48.34301	F-statistic	46.66510	
Durbin-Watson stat	1.620601	Prob(F-statistic)	0.000000	

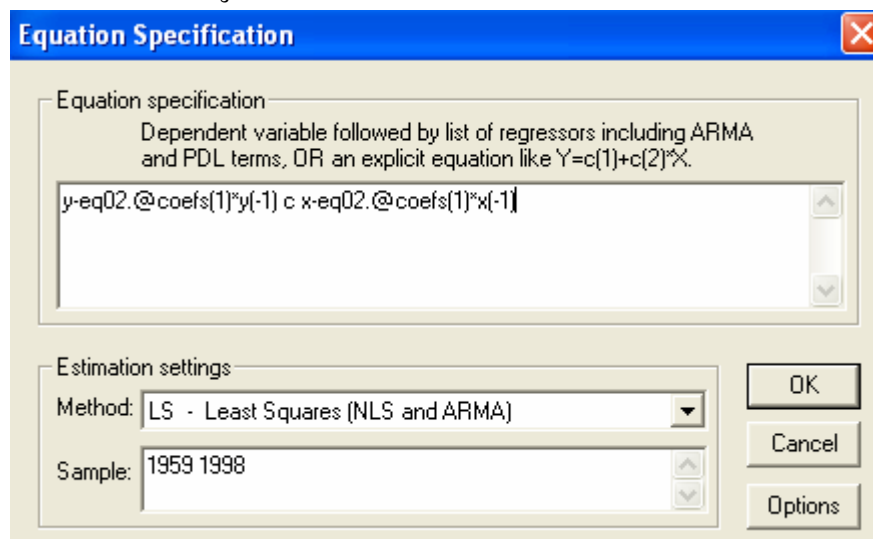
5.6 การบรรเทาปัญหา Autocorrelation ด้วย Generalized Least Square (GLS) โดยใช้ Iterative method of estimating ρ ด้วยวิธี Cochrane-Orcutt procedure

การประมาณค่า coefficient of autocovariance (ρ) ในข้อ 5.5 เป็นการประมาณเพียงครั้งเดียว แต่วิธีการ Cochrane-Orcutt จะประมาณค่า ρ ด้วยกระบวนการซ้ำๆ (Iterative)

เราให้สมการตั้งต้น 'Y C X' ในหน้า 27 เป็น สมการ 'eq01'

ขั้นตอนที่ 1: ประมาณค่า ρ เหมือนกับหัวข้อที่แล้ว และเราให้ save สมการดังกล่าวในชื่อ 'eq02' เราจะได้ค่า rho เริ่มต้น คือ 0.914245

ขั้นตอนที่ 2: ประมาณค่าสมการซึ่งผนวกปัญหา autocorrelation เข้าไปแล้ว (YSTAR กับ XSTAR) ด้วยการระบุสมการเป็น 'y-eq02.@coefs(1)*y(-1) c x-eq02.@coefs(1)*x(-1)' ดังรูปข้างล่าง แล้วคลิก OK จะได้สมการที่แก้ปัญหาลงดังหน้าทีแล้ว



ขั้นตอนที่ 3: สร้าง residual series ขึ้นมาใหม่จากสมการข้างต้น โดยพิมพ์ 'series resid1=y-(eq03.@coefs(1)/(1-eq02.@coefs(1))+eq03.@coefs(2)*x)' ในช่องคำสั่ง แล้ว Enter จะปรากฏรูป

 RESID1 successfully computed. ในช่องมุมซ้ายล่าง

ขั้นตอนที่ 4: Run สมการ eq02, eq03 และสร้าง residual series ขึ้นใหม่ โดย

(1) เปิดสมการ eq02 ขึ้นมาแล้วคลิก **Estimate** และ **OK**

(2) เปิดสมการ eq03 ขึ้นมาแล้วคลิก **Estimate** และ **OK**

(3) series resid1=y-(eq03.@coefs(1)/(1-eq02.@coefs(1))+eq03.@coefs(2)*x)

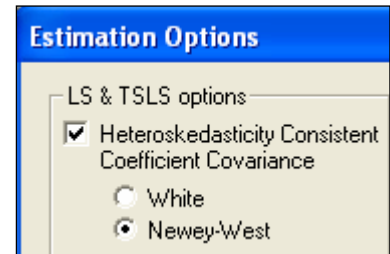
แล้วย้อนกลับไปยัง (1) ใหม่ จนกระทั่ง ค่า ρ (ค่าสัมประสิทธิ์ของ resid1 ในสมการ eq02) ไม่ค่อยเปลี่ยนแปลง เช่น น้อยกว่า 0.001 เราจะได้ค่า ρ เท่ากับ 0.8979 แล้ว

ขั้นตอนที่ 5: ใช้ค่า ρ ดังกล่าวในการประมาณค่าสมการระหว่าง YSTAR กับ XSTAR ดังเช่นขั้นตอนที่ 3 ในข้อ 5.5

5.7 วิธีการ Newey-West เพื่อแก้ไข Standard Error ของ OLS

ขั้นตอนที่ 1. ประมาณค่าสมการถดถอยโดยในช่อง Equation Specification ก็ใส่ตัวแปรตามปกติ คือ Y C X แล้วเลือก Options

ขั้นตอนที่ 2. จะปรากฏหน้าต่าง Equation Options ขึ้นมา ให้คลิกที่ช่อง Heteroskedasticity Consistent Coefficient Covariance และเลือกวิธีการของ Newey-West แล้วคลิก OK จะทำให้กลับมาสู่หน้าต่าง Equation Specification แล้วคลิก OK อีกครั้ง จะได้ผลการประมาณค่าดังตารางข้างล่าง



Dependent Variable: Y Method: Least Squares Date: 02/04/04 Time: 02:53 Sample: 1959 1998 Included observations: 40 Newey-West HAC Standard Errors & Covariance (lag truncation=3)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	29.51925	4.118072	7.168223	0.0000
X	0.713659	0.051281	13.91661	0.0000
R-squared	0.958449	Mean dependent var	85.64500	
Adjusted R-squared	0.957356	S.D. dependent var	12.95632	
S.E. of regression	2.675533	Akaike info criterion	4.854881	
Sum squared resid	272.0220	Schwarz criterion	4.939325	
Log likelihood	-95.09761	F-statistic	876.5495	
Durbin-Watson stat	0.122904	Prob(F-statistic)	0.000000	

เมื่อเปรียบเทียบค่า standard error (Std. Error) กับสมการที่ไม่ได้มีการแก้ไขปัญหา Autocorrelation (ตารางหน้า 27) พบว่าค่าสัมประสิทธิ์และ R^2 เท่ากัน แต่ standard error. กรณีไม่ได้แก้ไขมีค่าน้อยกว่าทำให้ค่า t สูง ในขณะที่ standard error ของ Newey-West method for correcteing standard error มีค่าสูงขึ้น และลดขนาดของค่า t แสดงว่า standard error ของ OLS underestimate true standard error บางคนอาจสังเกตเห็นว่า ค่า Durbin-Watson ของสมการ Newey-West method มีค่าเท่ากับ OLS แต่ไม่ต้องกังวล เพราะวิธีการ Newey-West ได้แก้ไขปัญหา Autocorrelation เรียบร้อยแล้ว