คู่มือการใช้โปรแกรม EViews เบื้องต้น:
สำหรับการวิเคราะห์ทางเศรษฐมิติ

โดย

นายอัครพงศ์ อั้นทอง
สถาบันวิจัยสังคม มหาวิทยาลัยเชียงใหม่

กันยายน 2550

สารบัญ

	หน้า
สารบัญ	**ii**
1. บทนำ	**1**
2. พื้นฐานโปรแกรม EViews	**1**
3. การใช้งานโปรแกรม EViews เบื้องต้น	**4**
3.1 การสร้าง Workfile ใหม่	**4**
3.2 การสร้าง Objects Series ของข้อมูล	**8**
➢ *การป้อนข้อมูลเข้าโดยตรง*	
➢ *การนำข้อมูลเข้าจากโปรแกรม Spreadsheet*	
➢ *การ Copy และ Paste ข้อมูลจาก Spreadsheet สู่โปรแกรม EViews*	
➢ *เครื่องหมายทางคณิตศาสตร์ในโปรแกรม EViews*	
➢ *คำสั่งพื้นฐานในการ Transform ข้อมูล*	
3.3 การสร้างกราฟ	**18**
3.4 การตรวจสอบข้อมูล	**21**
4. วิธีการประมาณค่าแบบกำลังสองน้อยที่สุด (Ordinary Least Square Method)	**24**
4.1 ข้อสมมติพื้นฐานของวิธีการกำลังสองน้อยที่สุด	**24**
4.2 รูปแบบฟังก์ชันที่ใช้ในสมการถดถอย	**24**
4.3 การใช้โปรแกรม EViews ในการประมาณค่าสัมประสิทธิ์ของสมการถดถอย	**25**
4.4 การอ่านผลการประมาณค่าสัมประสิทธิ์ที่ได้จากโปรแกรม EViews	**28**
4.5 การแสดงผลค่าจริง ค่าพยากรณ์ และค่าความคลาดเคลื่อน	**28**
4.6 การสร้างค่าพยากรณ์ และค่าความคลาดเคลื่อน	**29**
4.7 การทดสอบสมมติฐานข้อจำกัดของค่าสัมประสิทธิ์ด้วย Wald Test	**32**
5. Multicollinearity	**34**
5.1 ปัญหา Multicollinearity	**34**
5.2 การตรวจสอบปัญหา Multicollinearity	**34**
➢ *การตรวจสอบด้วย Simple Correlation Coefficients*	
➢ *การตรวจสอบด้วย Variance Inflation Factors (VIF)*	
5.3 การแก้ไขกรณีเกิดปัญหา Multicollinearity	**40**

สารบัญ

หน้าที่

6. Heteroskedasticity 42

6.1 ปัญหา Heteroskedasticity 42

6.2 การตรวจสอบปัญหา Heteroskedasticity 43

6.3 การแก้ไขกรณีเกิดปัญหา Heteroskedasticity 45

- ***Weighted Least Square (WLS)***
- ***Heteroskedasticity-Corrected Standard Errors***

7. Autocorrelation 50

7.1 ปัญหา Autocorrelation 50

7.2 การตรวจสอบปัญหา Autocorrelation 50

- ***การตรวจสอบจากค่า Durbin-Watson statistic (D.W.)***
- ***การตรวจสอบจากกราฟของค่าความคลาดเคลื่อน***
- ***การตรวจสอบจาก Correlogram – Q-statistics***
- ***Breusch-Godfrey Test***

7.3 การแก้ไขกรณีเกิดปัญหา Autocorrelation 54

8. สมการถดถอยเมื่อตัวแปรตามเป็นตัวแปรหุ่น (Regression on Dummy Dependent Variable) 56

8.1 แบบจำลองความน่าจะเป็นเชิงเส้น (Linear Probability Model) 57

8.2 แบบจำลองโพรบิตและโลจิต (Probit and Logit Model) 60

9. การวิเคราะห์ข้อมูลอนุกรมเวลา (Time Series Analysis) 64

9.1 การทดสอบข้อมูลอนุกรมเวลา 64

- ***Correlogram and Q-statistics***
- ***Unit Root Test***
- ***Granger Causality Tests***

9.2 วิธีการพยากรณ์ 75

- ***วิธีการวิเคราะห์แนวโน้ม (Time trend)***
- ***วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential smoothing)***
- ***วิธีการ Box and Jenkins (ARIMA)***

เอกสารอ้างอิง 90

คู่มือการใช้โปรแกรม EViews เบื้องต้น: สำหรับการวิเคราะห์เศรษฐมิติ

1. บทนำ

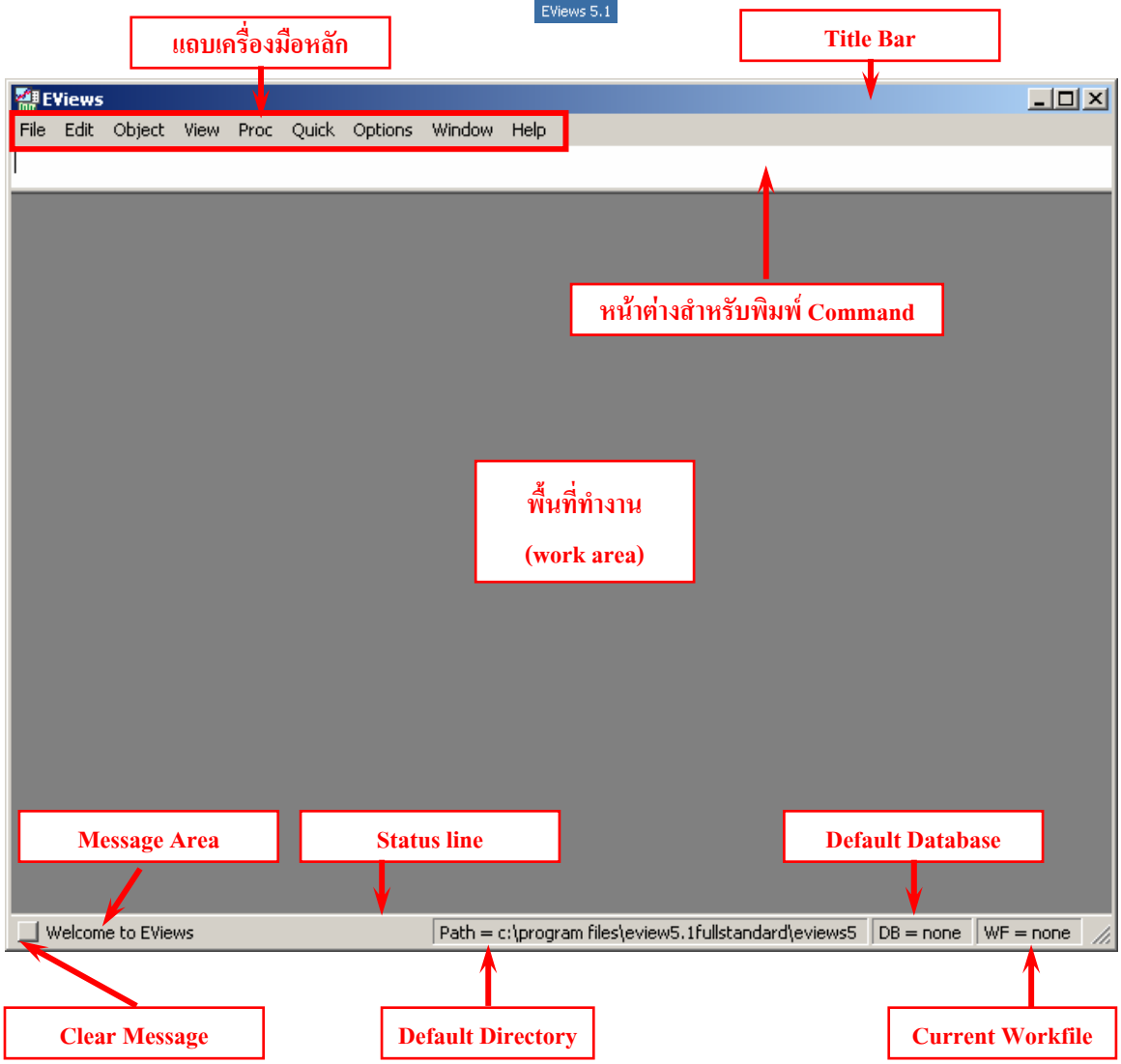
วิชาเศรษฐมิติเป็นวิชาหนึ่งที่มีความสำคัญในสาขาเศรษฐศาสตร์ วิชานี้เป็นวิชาที่คอยสร้างเครื่องมือให้กับนักเศรษฐศาสตร์ในการนำมาใช้วิเคราะห์ความเป็นจริงและปรากฏการณ์ต่างๆ ที่เกี่ยวข้องกับเศรษฐศาสตร์และธุรกิจ ในอดีตการนำเครื่องมือจากวิชานี้มาใช้ค่อนข้างทำได้ลำบาก เหตุผลหนึ่งมาจากการที่ในขณะนั้นไม่มีโปรแกรมคอมพิวเตอร์ออกมาสนับสนุนอย่างกว้างขวางเหมือนกับในปัจจุบัน ทำให้ในปัจจุบันการใช้เครื่องมือทางเศรษฐมิติมีความสะดวกมากขึ้น หนึ่งในโปรแกรมที่นักเศรษฐศาสตร์นิยมใช้ในการวิเคราะห์ทางเศรษฐมิติ ก็คือโปรแกรม EViews โปรแกรมนี้ได้รับการพัฒนามาจากโปรแกรม Micro TSP (Time Series Processor) ที่มีการนำเสนอครั้งแรกเมื่อประมาณปี พ.ศ. 2524 โปรแกรมนี้ถูกเขียนขึ้นโดยนักเศรษฐศาสตร์ เพื่อนำมาใช้ในการคำนวณข้อมูลที่เป็นข้อมูลอนุกรมเวลา (Time series data) ในระยะแรกเขียนเพื่อให้สามารถใช้ได้กับเครื่องคอมพิวเตอร์ Mainframe และมีการพัฒนาเพื่อใช้กับเครื่อง PC (Personal Computer) ในเวลาต่อมาเมื่อบริษัท Quantitative Micro Software Co.,Ltd. ได้รับลิขสิทธิ์จาก TSP ในการพัฒนาโปรแกรม Micro TSP เพื่อนำมาวิเคราะห์ข้อมูลประเภทภาคตัดขวาง (Cross section data) ซึ่งได้พัฒนาให้สามารถใช้กับระบบปฏิบัติการ DOS ต่อมาได้มีการพัฒนาโปรแกรมอย่างต่อเนื่อง และให้สามารถใช้งานได้กับระบบปฏิบัติการ Windows ซึ่งเป็นระบบปฏิบัติการที่ได้รับความนิยมในปัจจุบัน ทางบริษัท Quantitative Micro Software Co.,Ltd. จึงได้เปลี่ยนชื่อจากโปรแกรม TSP มาเป็นโปรแกรม EViews นับตั้งแต่ปี พ.ศ. 2536 และได้มีการพัฒนาโปรแกรมนี้เรื่อยมา โดยส่วนใหญ่จะเน้นและให้ความสำคัญกับการวิเคราะห์ข้อมูลอนุกรมเวลา ในปัจจุบัน (ปี พ.ศ. 2550) โปรแกรม EViews ได้พัฒนามาจนถึงเวอร์ชั่น 6.0 ซึ่งภายในเวอร์ชั่นนี้มีความสามารถในการวิเคราะห์ทางเศรษฐมิติหลายๆ อย่าง และมีความสะดวกในการใช้งานมากกว่าโปรแกรมอื่นๆ จึงทำให้โปรแกรม EViews เป็นหนึ่งในโปรแกรมที่ได้รับความนิยมจากนักเศรษฐศาสตร์ในการนำมาใช้ในการวิเคราะห์ทางเศรษฐมิติ โดยเฉพาะในการวิเคราะห์ข้อมูลอนุกรมเวลา

2. พื้นฐานโปรแกรม EViews

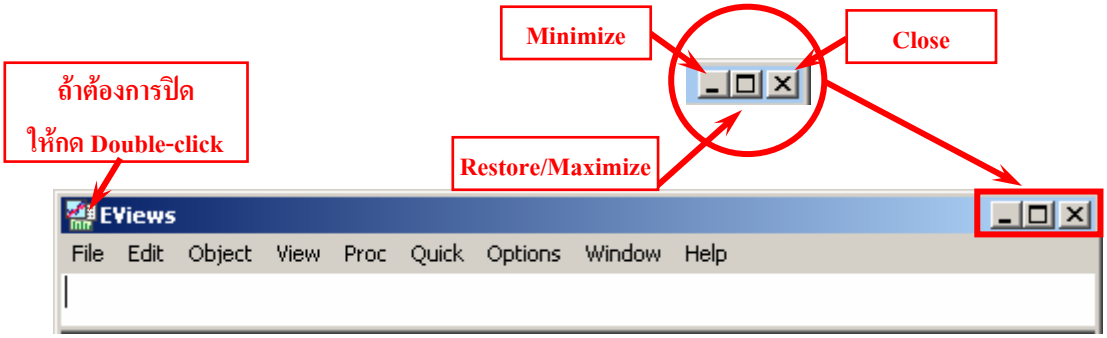
โปรแกรม EViews ถูกออกแบบมาให้ทำงานในลักษณะของ Objects ซึ่งผู้ใช้งานจะต้องทำงานและจัดการข้อมูลต่างๆ บน Objects ที่มีความแตกต่างกัน เช่น Objects Series (ใช้สำหรับเก็บข้อมูล) Objects Equation (ใช้สำหรับการประมาณค่าสัมประสิทธิ์) Objects โดยส่วนใหญ่ถูกเก็บไว้ที่ Workfile ดังนั้นในขั้นตอนแรกของการใช้งานโปรแกรม EViews ผู้ใช้จะต้องสร้าง Workfile ก่อน ซึ่งเป็นหน้าต่างหลักที่ใช้ทำงาน สำหรับในหน้าต่าง Workfile จะสามารถมี Objects ต่างๆ ได้หลายๆ Objects ดังนั้นผู้ใช้ควรมีการจัดการและตั้งชื่อของ Objects ต่างๆ ให้สามารถใช้งานได้สะดวกและรวดเร็ว ในที่นี้จะขอสรุปขั้นตอนการเข้าสู่โปรแกรม EViews การสร้างหน้าต่าง Workfile และวิธีการสร้าง Objects ต่างๆ สำหรับการทำงาน ซึ่งในเอกสารฉบับนี้ได้ใช้โปรแกรม EViews Version 5.1 เป็นตัวอย่างในการสาธิต ดังมีรายละเอียดพอสังเขปดังนี้

➢ เปิดโปรแกรม EViews โดยเลือกไอคอน

แล้วจะเข้าสู่โปรแกรมดังรูป

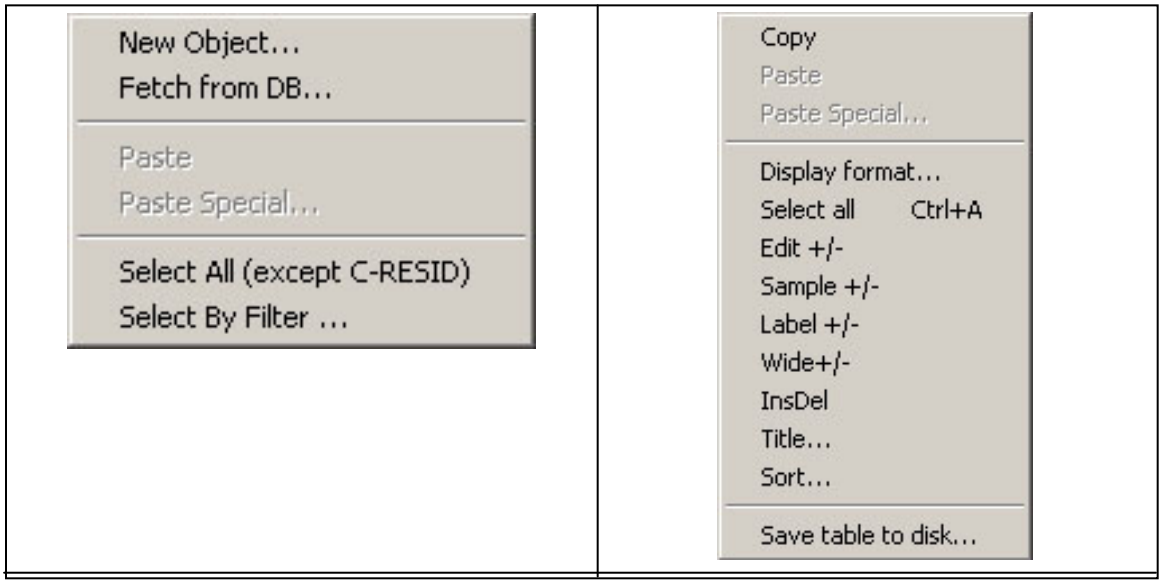


จากภาพข้างต้น จะเห็นได้ว่าโปรแกรม EViews จะมีหน้าตาคล้ายกับชุดโปรแกรมของ Microsoft ซึ่งสามารถใช้งานเหมือนกับ Window feature คือ มี Main menu, มี Minimize window, Maximize window, Closed เหมือนกัน และใช้เป็น Pull down menu เหมือนกับโปรแกรมของ Microsoft โดยทั่วไป



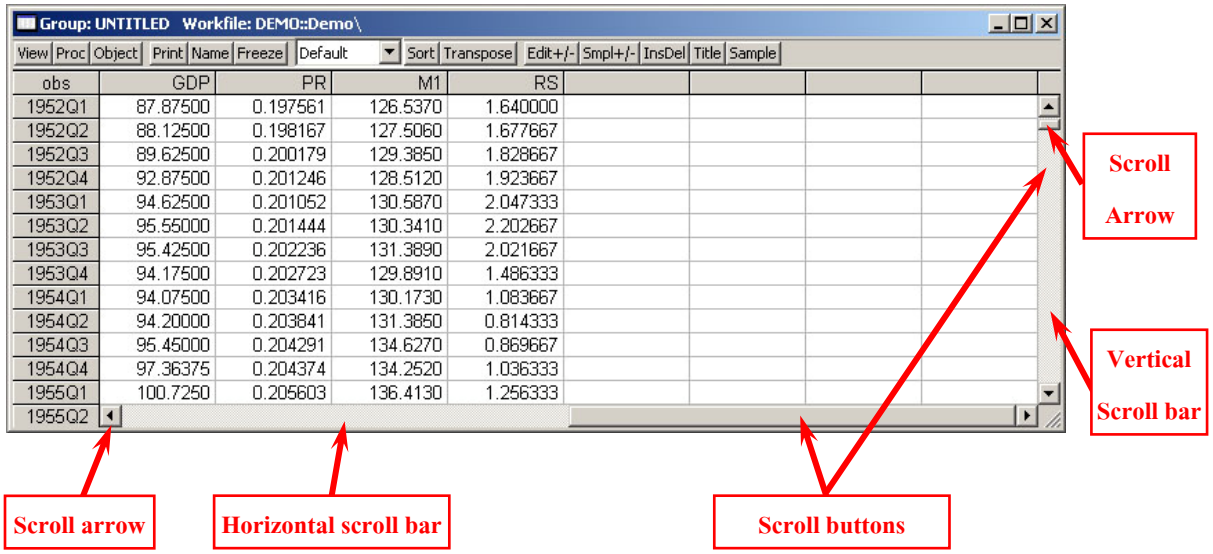
นอกจากนี้ยังมี Tip ต่างๆ ในการใช้หน้าต่าง แถบเครื่องมือ บนโปรแกรม EViews ดังนี้

- การใช้ Mouse โดยถ้ากดปุ่มทางซ้ายมือ ก็จะเป็นการเปิดหน้าต่างที่ต้องการ และถ้า Double-click ที่ Objects ก็จะเป็นการเปิดหน้าต่าง Objects แต่ถ้ากดปุ่มทางขวามือที่หน้าต่างหลัก (บริเวณพื้นที่ทำงาน), Workfile และ Objects ก็จะปรากฏหน้าต่างให้เลือกทำงานดังนี้



คลิกขวาที่ Workfile **คลิกขวาที่ Objects**

- ในหน้าต่างของแต่ละ Objects ของโปรแกรม EViews จะมี Scroll bar ให้ทั้งในแนวนอนและแนวตั้ง เหมือนกับโปรแกรมต่างๆ ของ Microsoft ดังนั้นผู้ใช้งานจึงสามารถเลื่อนดูข้อมูลหรือรายละเอียดต่างๆ ภายในหน้าต่าง Objects ได้ตลอดทั้งแนวนอนและแนวตั้ง ดังนี้



- ผู้ใช้สามารถเลือกเปิด Objects ต่างๆ ได้พร้อมกัน โดยกดปุ่ม **CTRL** ได้ตลอดเวลาของการเลือก และถ้าต้องการยกเลิกรายการที่เลือกไว้ ให้กดปุ่ม **CTRL** และ Objects ที่ยกเลิก
- ถ้าต้องการยกเลิกการเปิดเมนู หรือหน้าต่างใดๆ บน EViews ให้กดปุ่ม **ESC** เช่นเดียวกัน ถ้าต้องการหยุด/ยกเลิกการทำงานของโปรแกรม EViews ให้กดปุ่ม **ESC**
- ถ้าต้องการปิดหน้าต่างของ Objects ต่างๆ หรือต้องการออกจากโปรแกรม EViews ให้กดปุ่ม **ALT** และ **F4** พร้อมๆ กันได้

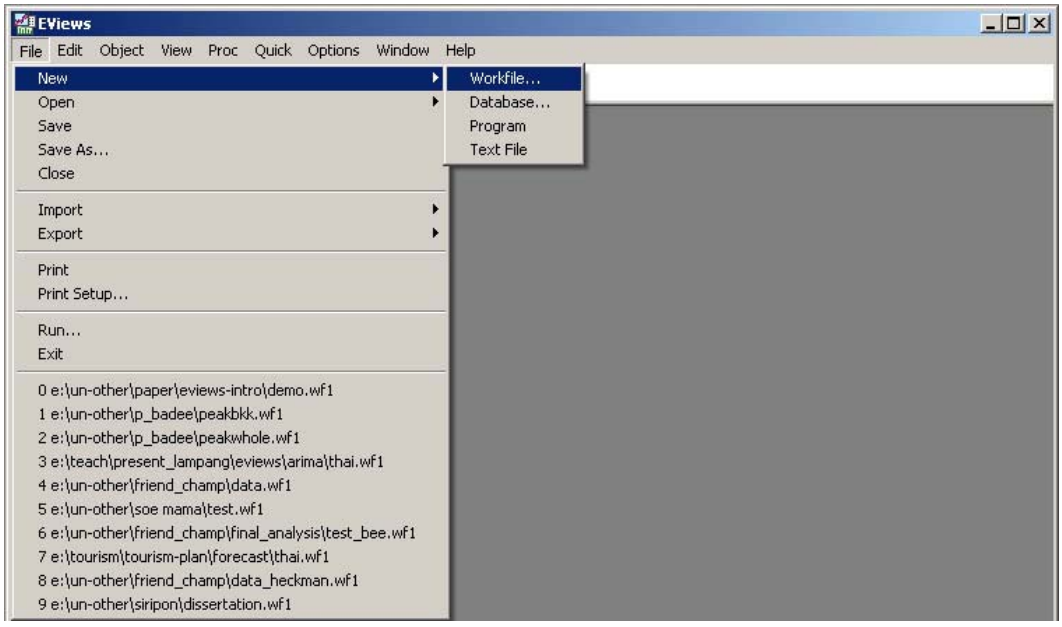
3. การใช้งานโปรแกรม EViews เบื้องต้น

เพื่อให้ผู้ใช้มีความเข้าใจในการใช้โปรแกรม EViews ดังนั้นในที่นี่จะมีการสมมติตัวอย่าง ที่ให้มาโดยโปรแกรม EViews ซึ่งสามารถเปิดดูได้ใน c:\programfile\EViews5\Example\data\demo.wf1 หรือ demo.xls ในตัวอย่างนี้จะเป็นข้อมูลอนุกรมเวลารายไตรมาสตั้งแต่ ปี ค.ศ. 1952 ไตรมาสที่ 1 ถึงปี ค.ศ. 1996 ไตรมาสที่ 4 โดยประกอบด้วยตัวแปรต่างๆ จำนวน 4 ตัวแปร คือ ตัวแปรอุปทานของเงิน (M1), ผลิตภัณฑ์มวลรวมภายในประเทศ (GDP), อัตราดอกเบี้ยระยะสั้น (RS) และระดับราคาสินค้าที่เป็นตัวที่แสดงถึงภาวะเงินเฟ้อ (PR) ส่วนเนื้อหาในหัวข้อนี้จะบรรยายถึงการสร้าง Workfile ใหม่, การสร้าง Objects Series ของข้อมูล, การสร้างกราฟ และการวิเคราะห์ข้อมูลด้วยสถิติพื้นฐาน ดังมีรายละเอียดในแต่ละหัวข้อดังนี้

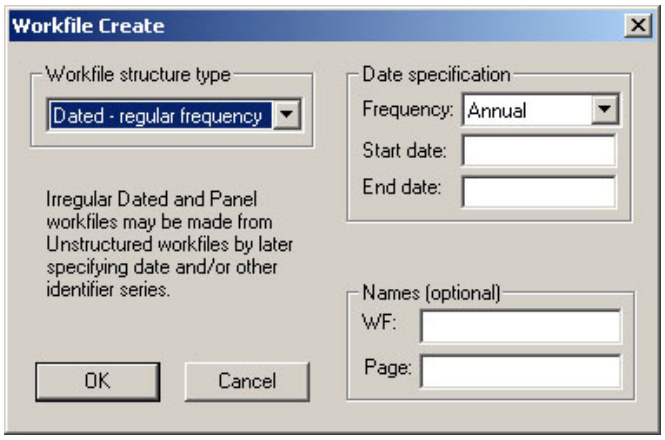
3.1 การสร้าง Workfile ใหม่

หลังจากเปิดโปรแกรม EViews แล้ว ก็จะปรากฏหน้าต่างเมนูหลัก [ดูได้จากหัวข้อที่ 2 หน้าที่ 2 ของคู่มือฉบับนี้] ต่อมาถ้าต้องการทำงานจะเปิดหรือสร้าง Workfile ใหม่ สำหรับใช้ทำงาน ถ้าในกรณีที่เป็นการทำงานครั้งแรกผู้ใช้จะต้องสร้าง Workfile ใหม่ก่อน โดยปฏิบัติตามขั้นตอนดังต่อไปนี้

ขั้นตอนที่ 1 เลือก **File/New/Workfile** ในเมนูหลัก (**Main menu**) ของ EViews ดังรูป



ขั้นตอนที่ 2 หลังจากนั้นจะปรากฏหน้าต่าง **Workfile Create** เพื่อให้กำหนดคุณลักษณะของข้อมูลที่ใช้ในการวิเคราะห์ และกำหนดค่าเริ่มต้นและค่าสิ้นสุดในช่องว่างด้านล่างของหน้าต่าง Workfile Create



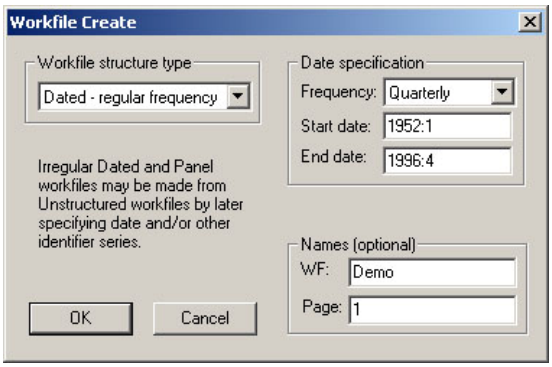
สำหรับรายละเอียดของรูปแบบข้อมูล ใน EViews 5.1 ได้แบ่งชนิดโครงสร้างของข้อมูลออกเป็น 3 แบบ ดังมีรายละเอียดต่างๆ ดังนี้

ก. Dated – regular frequency (**Default**) เป็นรูปแบบของข้อมูลประเภท Time series ผู้ใช้จะต้องกำหนดลักษณะและรูปแบบของข้อมูล รวมทั้งการกำหนดค่าเริ่มต้น (Start) และค่าสุดท้าย (End) ของข้อมูลที่จะใช้ดังมีรายละเอียดต่อไปนี้

รูปแบบของข้อมูล	**การกำหนดค่า Start date และ End date ของข้อมูลที่จะใช้**
Annual (ข้อมูลรายปี)	YYYY
Semi Annual (ข้อมูลรายครึ่งปี)	YYYY:S
Quarterly (ข้อมูลรายไตรมาส)	YYYY:Q
Monthly (ข้อมูลรายเดือน)	YYYY:M
Weekly (ข้อมูลรายสัปดาห์)	M/DD/YYYY
Daily (5 day weeks) (ข้อมูลทุก 5 วัน)	M/ DD/YYYY
Daily (7 day weeks) (ข้อมูลทุก 7 วัน)	M/ DD/YYYY
Integer date (กำหนดเป็นรูปแบบใดก็ได้)	กำหนดตามเงื่อนไขข้างต้น

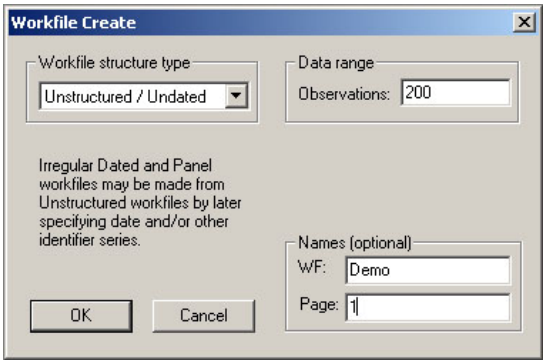
หมายเหตุ : YYYY= ปี พ.ศ./ค.ศ., S = ครึ่งปี, Q = ไตรมาส, M = เดือน, และ D = วันที่

ในที่นี้กำหนดให้ตัวอย่างที่ใช้เป็นข้อมูลรายไตรมาส ตั้งแต่ปี ค.ศ. 1952 ไตรมาสที่ 1 ถึงปี ค.ศ. 1996 ไตรมาสที่ 4 ดังนั้นจึงสามารถกำหนดค่าต่างๆ ดังรูป



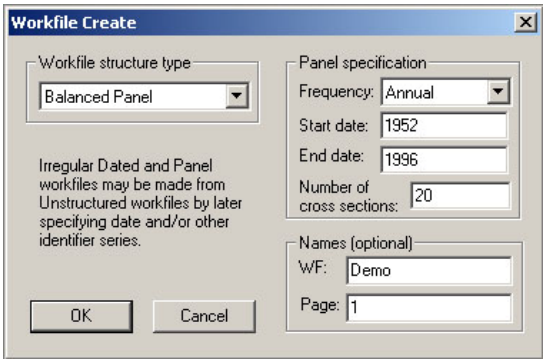
ข. Unstructured/Undated เป็นรูปแบบของข้อมูลประเภท Cross section ผู้ใช้จะต้องกำหนดจำนวนตัวอย่างของข้อมูลให้เป็นจำนวนเต็ม โดยใส่ในช่อง **Data range/Observations:** ดังรูป

ในที่นี้กำหนดให้มีจำนวนตัวอย่างจำนวน 200 ตัวอย่าง โดยการเติมเลขจำนวนเต็ม 200 ในช่อง **Data range/Observations:** ดังรูป

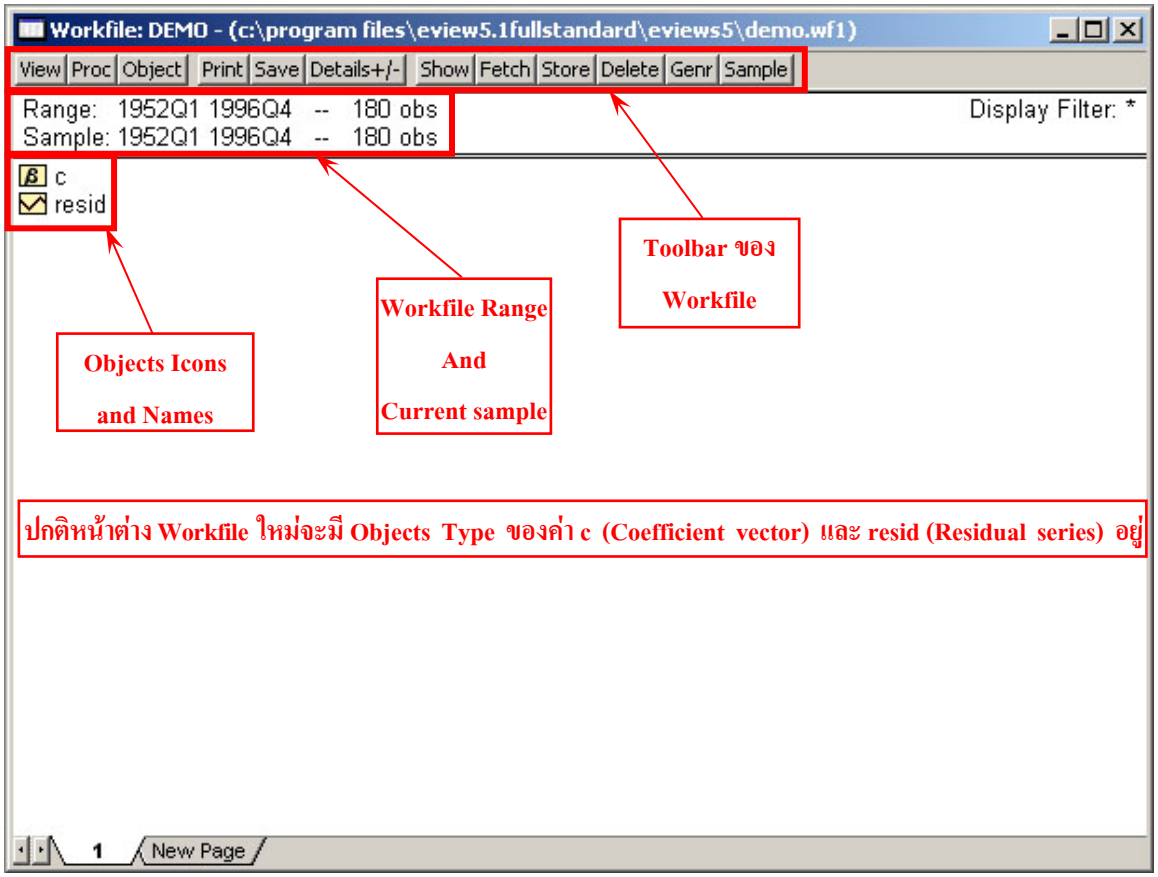


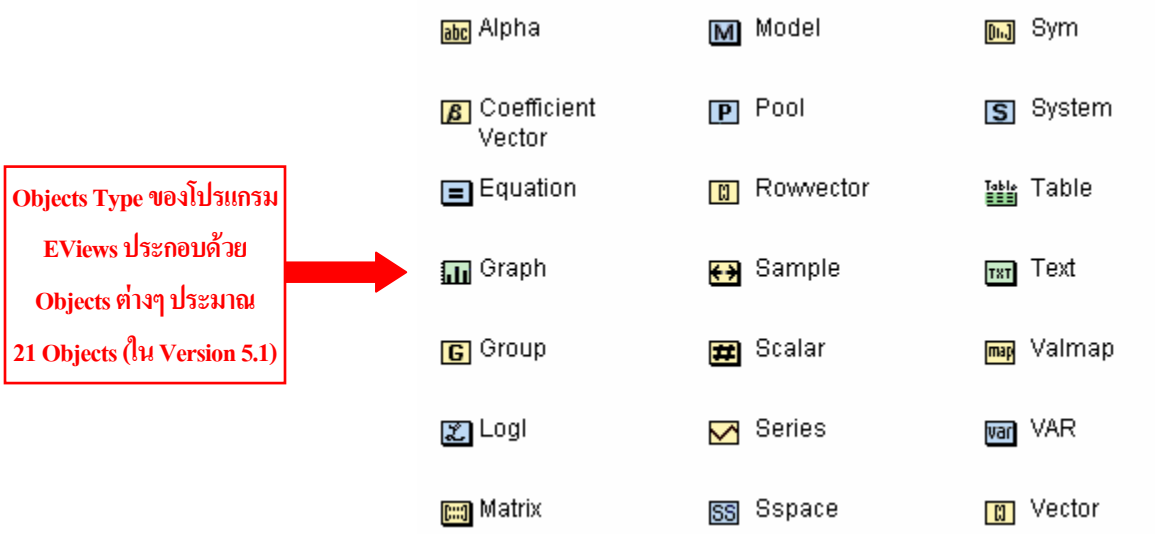
ค. Balanced Panel เป็นรูปแบบของข้อมูลประเภท Panel/Pool ผู้ใช้ต้องกำหนดรูปแบบของข้อมูลประเภท Time series (ตามรายละเอียดในข้อ ก.) และรูปแบบของข้อมูลประเภท Cross section (ตามรายละเอียดในข้อ ข.) ดังรูป

ในที่นี้จะกำหนดให้ตัวอย่างที่ใช้เป็นข้อมูลรายปี ระหว่างปี 1952 – 1996 และในแต่ละปีมีจำนวนตัวอย่างจำนวน 20 ตัวอย่าง โดยเติมตัวเลขต่างๆ ดังรูป



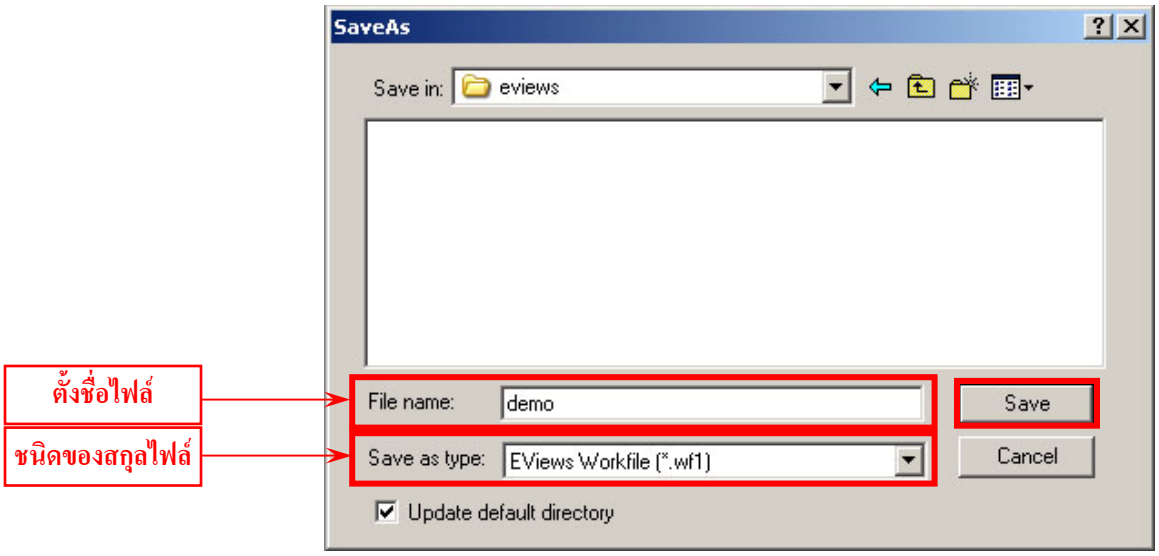
ขั้นตอนที่ 3 ***[กรณีตัวอย่าง: ในขั้นตอนที่ 2 เพื่อความเข้าใจ กำหนดให้ข้อมูลที่จะวิเคราะห์เป็นข้อมูลรายไตรมาส ตั้งแต่ปี ค.ศ. 1952 ไตรมาสที่ 1 ถึงปี ค.ศ. 1996 ไตรมาสที่ 4 และให้ Workfile ชื่อว่า "Demo" และกำหนดให้หน้าที่เก็บข้อมูลชื่อว่า "1"]*** ภายหลังจากปฏิบัติตามขั้นตอนที่ 1 – 2 เสร็จแล้วให้กดปุ่ม OK ก็จะปรากฏหน้าต่างใหม่ที่เป็น Workfile ชื่อ "**Demo**" โดย EViews จะแสดงหน้าต่างดังกล่าวในพื้นที่ทำงานของหน้าต่างหลัก และภายในหน้าต่าง Workfile จะมีข้อมูลของช่วง (Range) และตัวอย่าง (Sample) ไว้ที่มุมบนทางด้านซ้ายมือ ผู้ใช้สามารถเปลี่ยนขนาดของช่วงและตัวอย่างได้ตลอดเวลาตามที่ต้องการ นอกจากนี้โปรแกรม EViews จะสร้าง Objects ให้ 2 Objects คือ Objects "**c**" [Coefficient vector] ที่มีสมาชิกทุกตัวมีค่าเท่ากับ 0 และ "**resid**" [Residual series] ที่มีสมาชิกทุกตัวมีค่าเท่ากับ NA ดังรูป





ในการสร้าง Objects ใหม่ ทำได้โดยการเลือก **Objects/New Object** จาก **Main menu** หรือ Workfile

ขั้นตอนที่ 4 ทำการจัดเก็บ Workfile) ที่สร้างขึ้น โดยเลือกปุ่ม Save ในเมนูของ Workfile หรือ **File/Save** หรือ **File/Save As** ในเมนูของหน้าต่างหลัก ซึ่งภายหลังจากกดปุ่มดังกล่าวแล้วจะปรากฏหน้าต่างดังนี้



ให้ตั้งชื่อไฟล์ในช่อง File Name และเลือก Save as type เป็น Workfile ดังรูป แล้วกดปุ่ม Save

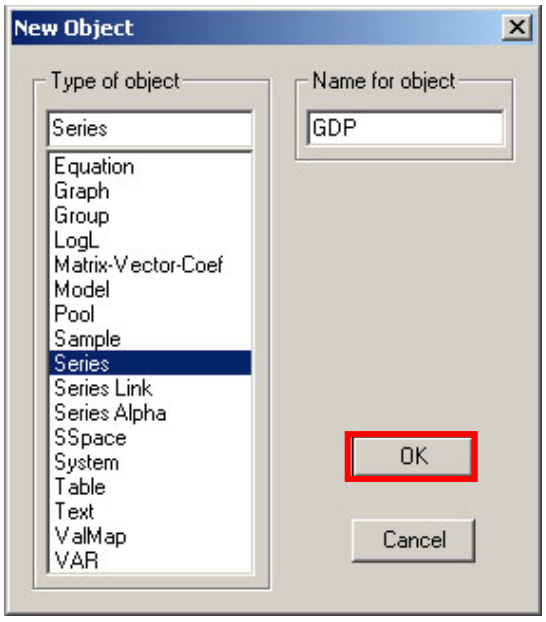
หมายเหตุ : สำหรับโปรแกรม EViews จะเก็บ Workfile ไว้ในนามสกุล *.wf1 และเก็บคำสั่งของโปรแกรมไว้ในนามสกุล *.prg และเก็บ Database ไว้ในนามสกุล *.edb

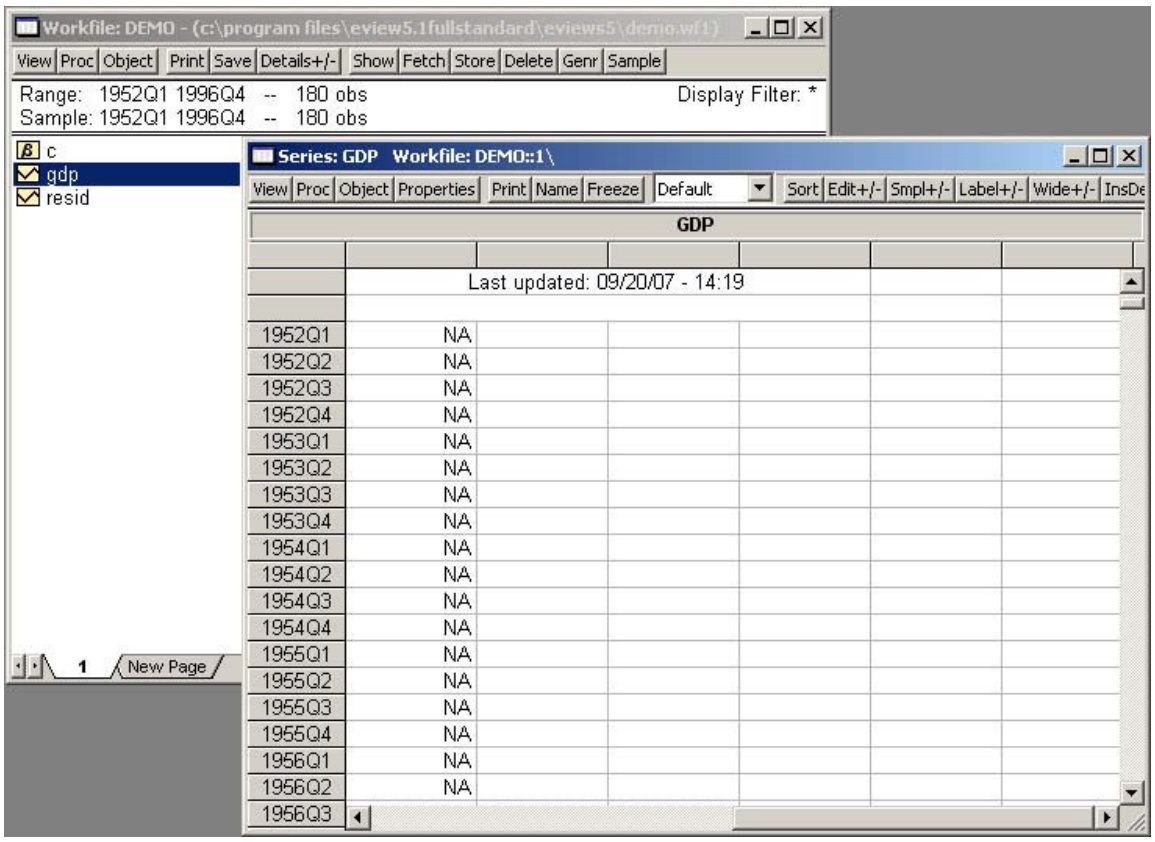
3.2 การสร้าง Objects Series ของข้อมูล

การสร้าง Objects Series ของข้อมูล มีวิธีการสร้างได้หลายๆ วิธี โดยเริ่มตั้งแต่การป้อนข้อมูลเข้าโดยตรง การนำข้อมูลเข้าจากโปรแกรม Spreadsheet และการ Copy และ Paste ข้อมูลจาก Spreadsheet สู่โปรแกรม EViews โดยวิธีการต่างๆ สามารถสรุปรายละเอียดได้ดังนี้

➢ ***การป้อนข้อมูลเข้าโดยตรง***

ขั้นตอนที่ 1 สร้าง Series ของข้อมูลตามตัวแปรที่ต้องการ โดยเลือก **Object/New Object/Series** จากเมนูหลัก หรือเมนู Workfile แล้วใส่ชื่อตัวแปรที่ต้องการในช่อง Name for object แล้วกดปุ่ม OK

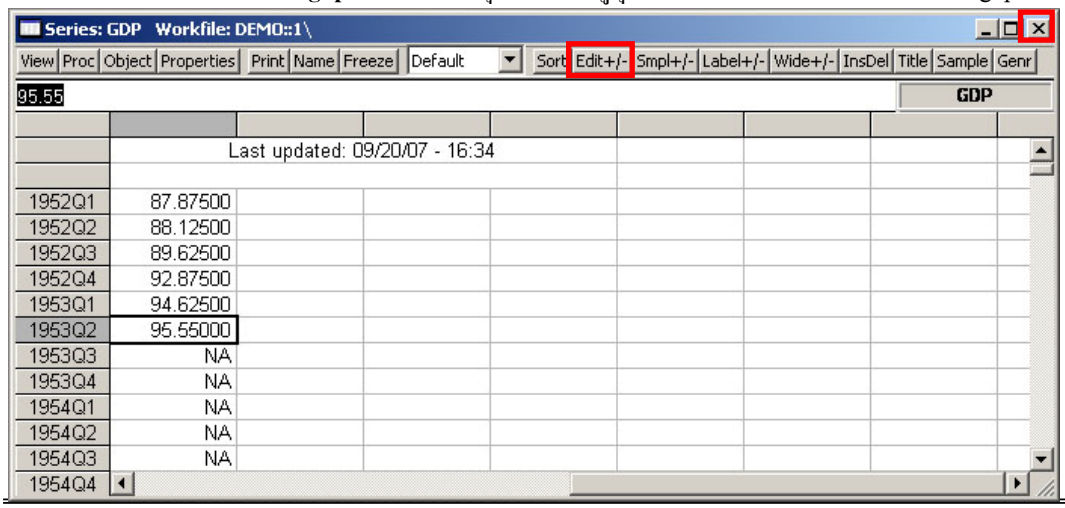




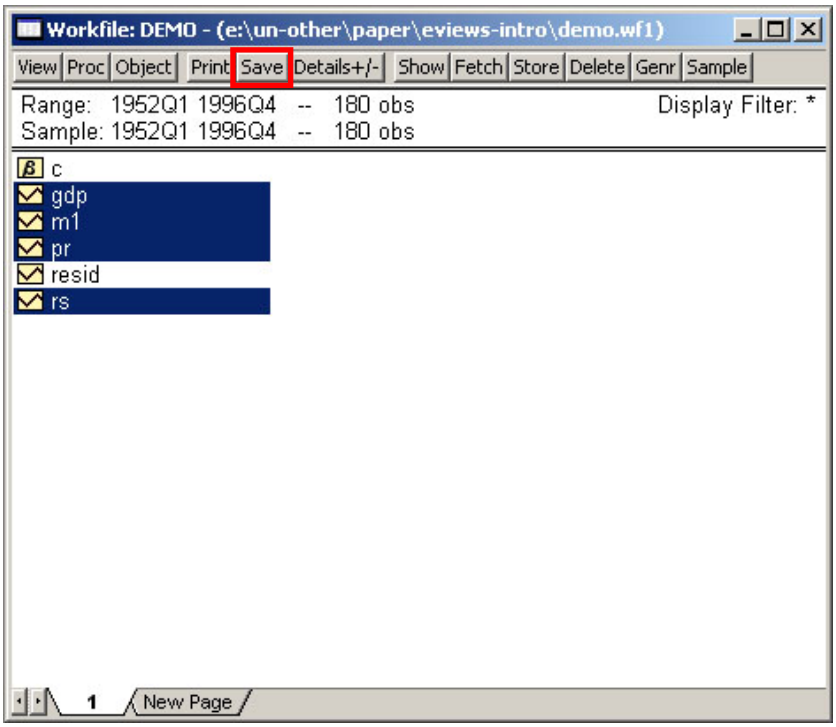
หลังจากกดปุ่ม OK แล้ว โปรแกรม EViews จะไปสร้าง Series ไว้ที่หน้าต่าง **Workfile [Series gdp]** และเปิด Series gdp โดยการ Double-click ที่ Series gdp พบว่า สมาชิกทุกตัวใน Series นี้เป็น NA (Not available)

หมายเหตุ: นอกจากวิธีข้างต้นแล้ว ยังสามารถทำได้โดยเลือก **Quick/Empty Group (Edit Series)** บนแถบเครื่องมือของหน้าต่างหลัก (**Main menu**) แทนก็ได้

ขั้นตอนที่ 2 ใส่ข้อมูลที่ต้องการลงใน Series gdp โดยการกดปุ่ม Edit+/- ในหน้าต่าง **Series gdp** แล้วใส่ข้อมูลแทนที่ NA เมื่อทำเสร็จแล้วให้กดปุ่ม Edit+/- อีกครั้งหนึ่ง เพื่อป้องการแก้ไขข้อมูลโดยไม่ได้ตั้งใจ หลังจากนั้นก็ทำการปิดหน้าต่าง **Series gdp** โดยการคลิกปุ่ม ✕ ซึ่งอยู่มุมบนทางขวามือของหน้าต่าง Series gdp



ขั้นตอนที่ 3 ทำตามขั้นตอนที่ 1 และ 2 ซ้ำกับตัวแปรที่เหลือจนครบทุกตัวแปร ซึ่งในหน้าต่าง Workfile จะมี Series ของตัวแปรต่างๆ ครบทุกตัวแปร (ในตัวอย่างมีทั้งหมด 4 ตัวอย่าง คือ GDP, PR, M1 และ RS)



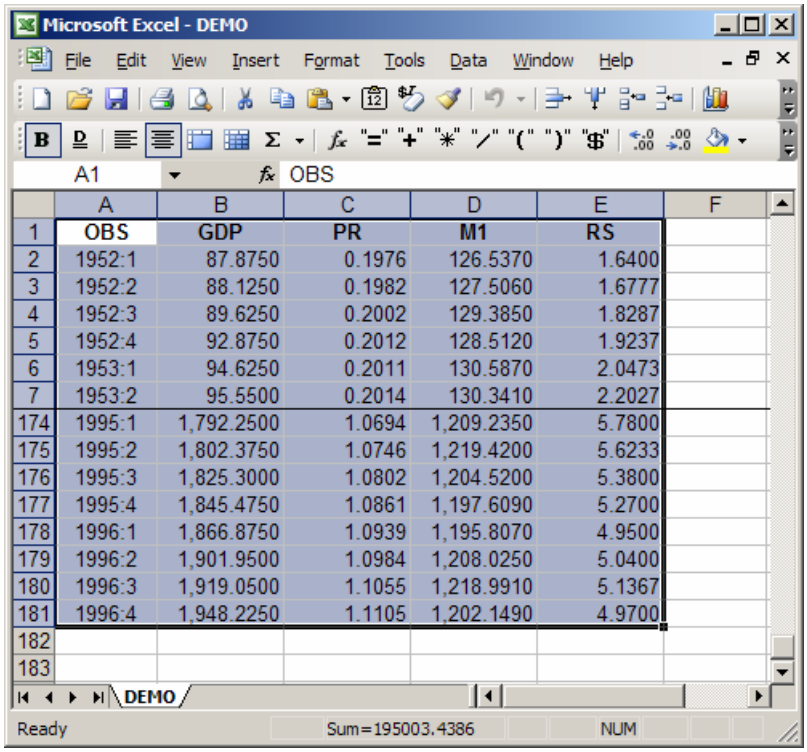
ขั้นตอนที่ 4 ทำการบันทึก Workfile โดยกดปุ่ม Save ในหน้าต่าง Workfile

หมายเหตุ : หากต้องการลบ Series ข้อมูลบาง Series ออก ให้ทำโดย เลือก Series ที่ต้องการลบ แล้วกดปุ่ม Delete ที่หน้าต่าง Workfile ซึ่งภายหลังจากกดปุ่มดังกล่าวแล้ว ก็จะปรากฏหน้าต่างดังรูป แล้วให้กดปุ่ม Yes แล้ว Series ข้อมูลดังกล่าวจะถูกลบออกจากหน้าต่าง Workfile

➢ ***การนำข้อมูลเข้าจากโปรแกรม Spreadsheet***

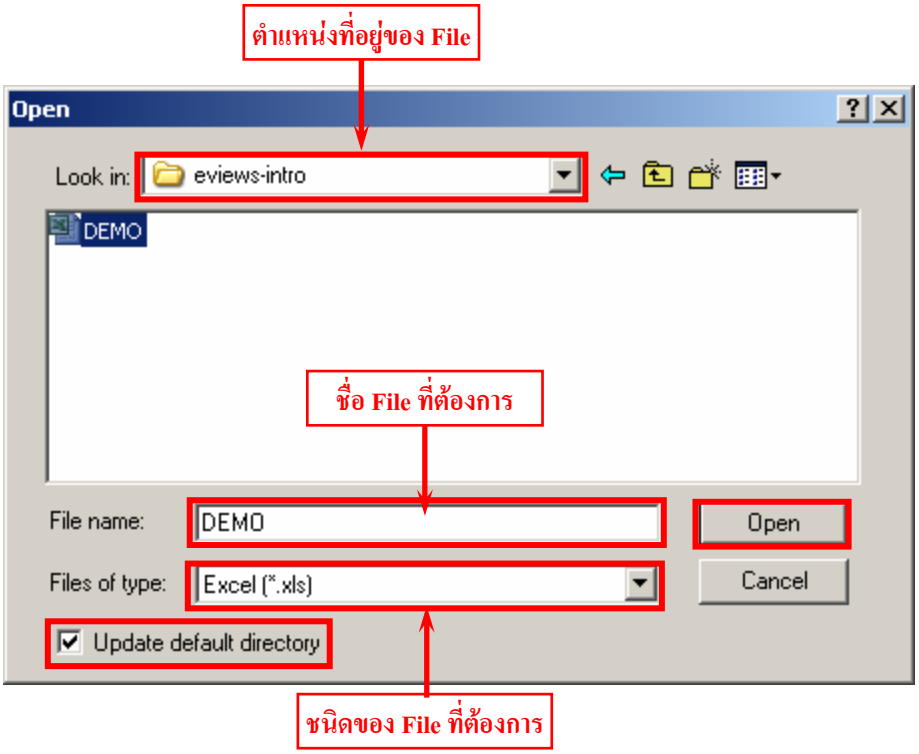
ก่อนที่จะนำข้อมูลจาก Excel เข้า จะต้องทำการสร้าง Workfile ตามขั้นตอนที่ได้กล่าวมาแล้ว และจะต้องทราบตำแหน่งที่อยู่ของ File Excel ที่ต้องการ พร้อมทั้งจำนวนตัวแปรที่ใช้ในการวิเคราะห์

ขั้นตอนที่ 1 เปิดแฟ้มงาน Excel เพื่อตรวจสอบความถูกต้องของข้อมูล และดูจำนวนตัวแปร จากรูปจะเห็นได้ว่าในแถวแรกจะเป็นชื่อของตัวแปร ในขณะที่แถวที่สองจะเป็นข้อมูล ดังนั้นเวลานำข้อมูลเข้าต้องระบุว่าข้อมูลเริ่มต้นที่ B2 และมีจำนวนตัวแปรทั้งหมด 4 ตัวแปร โดยทุกครั้งจะต้องระบุให้ Column แรก แสดงถึงจำนวนตัวอย่างที่ใช้ในการวิเคราะห์ เช่น ในตัวอย่างระบุว่า 1952:1 ถึง 1996:4 รวมทั้งหมด 180 ตัวอย่าง **<u>ภายหลังจากการตรวจสอบเรียบร้อยแล้วให้ปิด File Excel</u>**

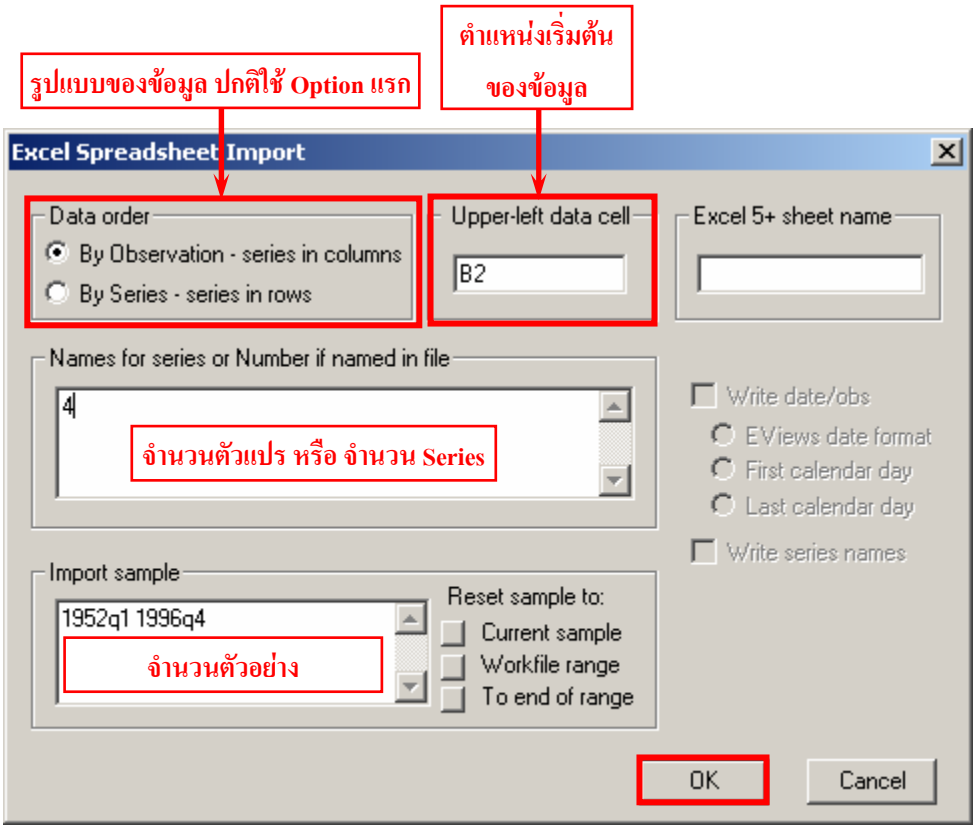


	A	B	C	D	E
1	OBS	GDP	PR	M1	RS
2	1952:1	87.8750	0.1976	126.5370	1.6400
3	1952:2	88.1250	0.1982	127.5060	1.6777
4	1952:3	89.6250	0.2002	129.3850	1.8287
5	1952:4	92.8750	0.2012	128.5120	1.9237
6	1953:1	94.6250	0.2011	130.5870	2.0473
7	1953:2	95.5500	0.2014	130.3410	2.2027
174	1995:1	1,792.2500	1.0694	1,209.2350	5.7800
175	1995:2	1,802.3750	1.0746	1,219.4200	5.6233
176	1995:3	1,825.3000	1.0802	1,204.5200	5.3800
177	1995:4	1,845.4750	1.0861	1,197.6090	5.2700
178	1996:1	1,866.8750	1.0939	1,195.8070	4.9500
179	1996:2	1,901.9500	1.0984	1,208.0250	5.0400
180	1996:3	1,919.0500	1.1055	1,218.9910	5.1367
181	1996:4	1,948.2250	1.1105	1,202.1490	4.9700
182					
183					

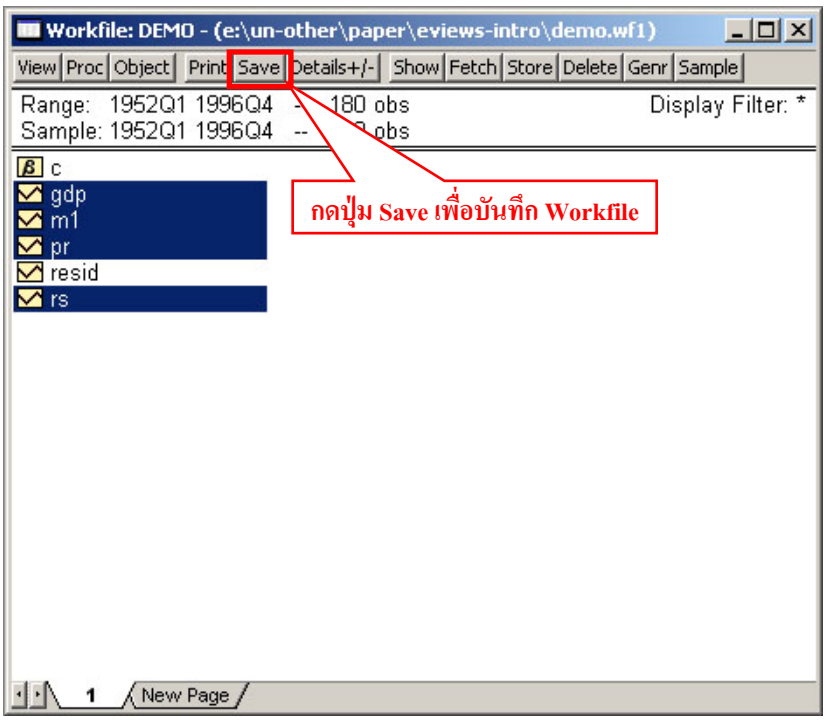
ขั้นตอนที่ 2 เปิดโปรแกรม EViews สร้าง Workfile ตามขั้นตอนที่ผ่านมา เลือก **Proc/Import/Read Text-Lotus-Excel** ในหน้าต่าง Workfile หรือ คลิก **File/Import/Read Text-Lotus-Excel** ในหน้าต่างหลัก ก็จะปรากฏหน้าต่างใหม่แล้วให้เลือก File Excel ที่ต้องการ แล้วกดปุ่ม Open



ขั้นตอนที่ 3 ภายหลังจากกดปุ่ม Open แล้ว จะปรากฏหน้าต่าง **Excel Spreadsheet Import** ซึ่งผู้ใช้จะต้องเลือก Option และเติมค่าต่างๆ ให้ครบตามที่ต้องการ แล้วกดปุ่ม OK

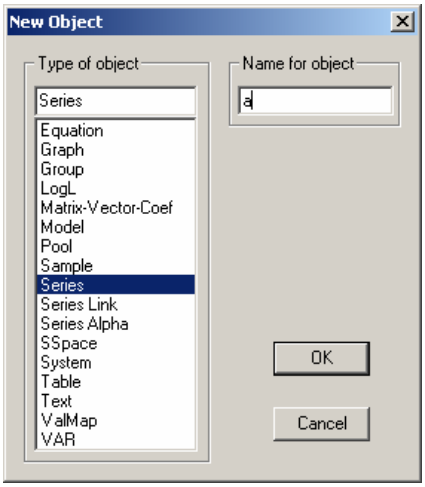


หลังจากกดปุ่ม OK โปรแกรม EViews จะนำข้อมูลเข้า และสร้าง Series ตามชื่อที่กำหนดไว้ใน Excel ก่อนที่จะทำการวิเคราะห์ใดๆ ให้ทำการ Save Workfile ก่อน โดยการกดปุ่ม Save ในหน้าต่าง Workfile

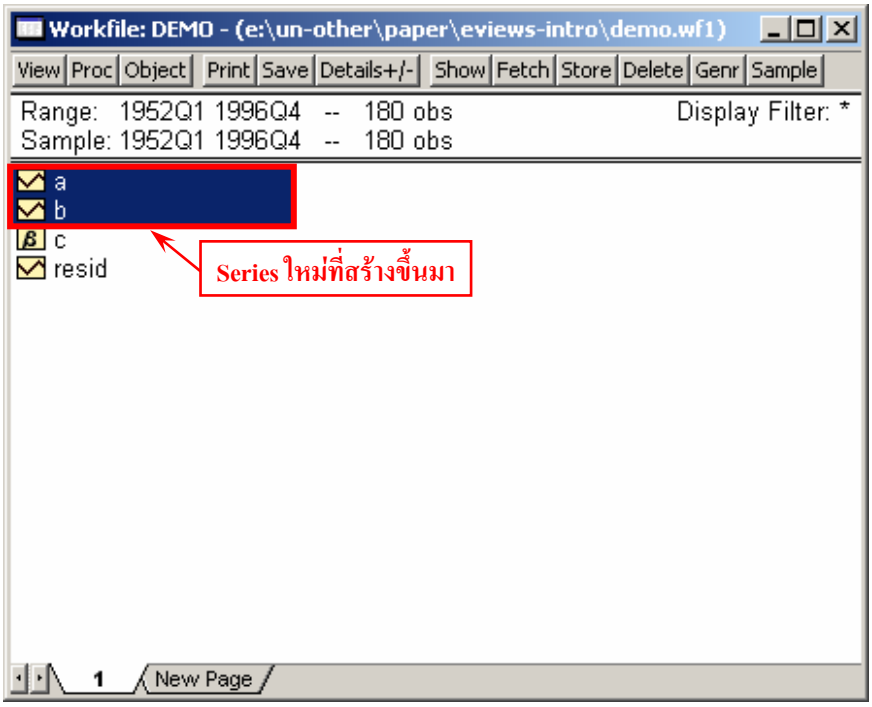


➢ ***การ Copy และ Paste ข้อมูลจาก Spreadsheet สู่โปรแกรม EViews***

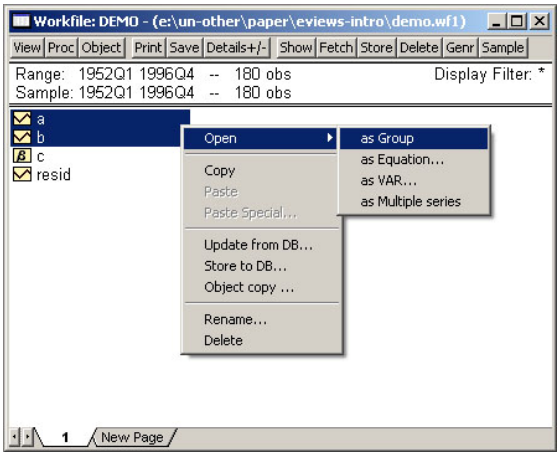
ขั้นตอนที่ 1 สร้างหน้าต่าง Workfile แล้วเลือก **Objects/New object** จากแถบเครื่องมือของหน้าต่าง Workfile จะปรากฏหน้าต่าง New Objects



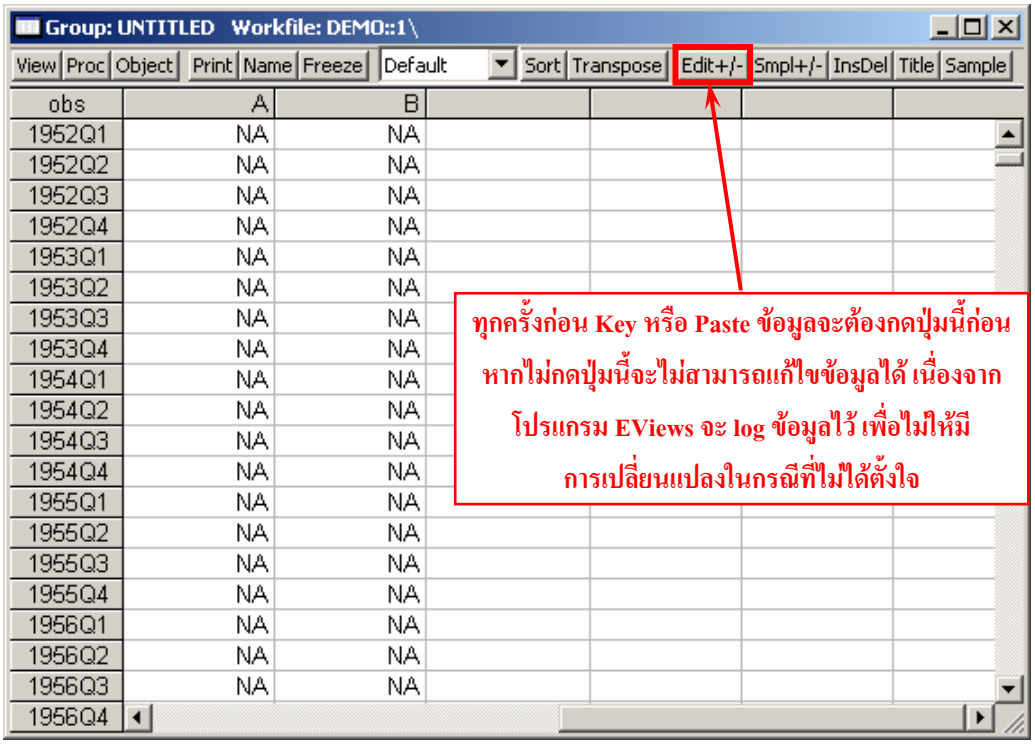
ขั้นตอนที่ 2 กำหนด Type of object เป็นประเภท Series แล้วตั้งชื่อที่ Name for object เป็นชื่ออะไรก็ได้ แต่ต้องเป็นภาษาอังกฤษ แล้วจึงกดปุ่ม OK ทำอย่างนี้ 2 ครั้ง จะได้ Series 2 Series ในที่นี้คือ a และ b ดังรูป



ขั้นตอนที่ 3 เลือก Series a และ b ที่สร้างขึ้นมาใหม่ แล้วให้คลิกขวาของ Mouse จะมีหน้าต่าง แล้วเลือก **Open/as Group** ดังรูป

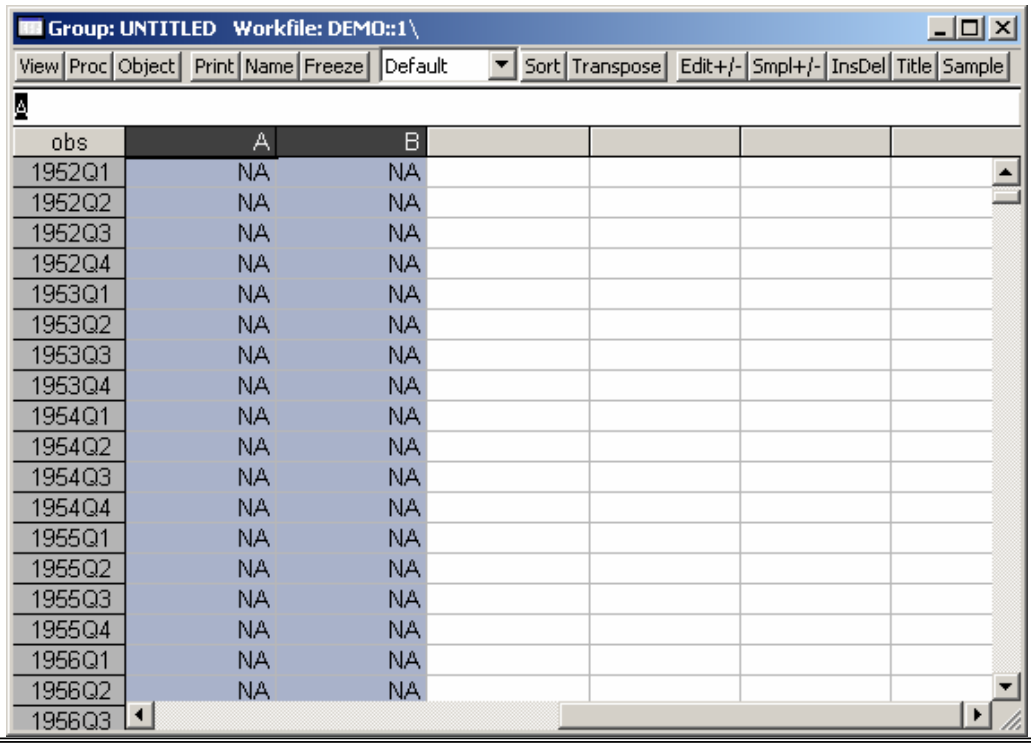


ขั้นตอนที่ 4 จากขั้นตอนที่ 3 จะได้หน้าต่าง Group ขึ้นมาดังรูป ซึ่งในหน้าต่างนี้ จะเห็นว่าคอลัมน์ A และ B จะมีค่า NA แสดงว่าในสองคอลัมน์ดังกล่าวไม่มีข้อมูล ก่อนที่จะนำข้อมูลจาก Excel มา Paste ลงในตารางนี้ ให้กดปุ่ม Edit+/- ก่อน



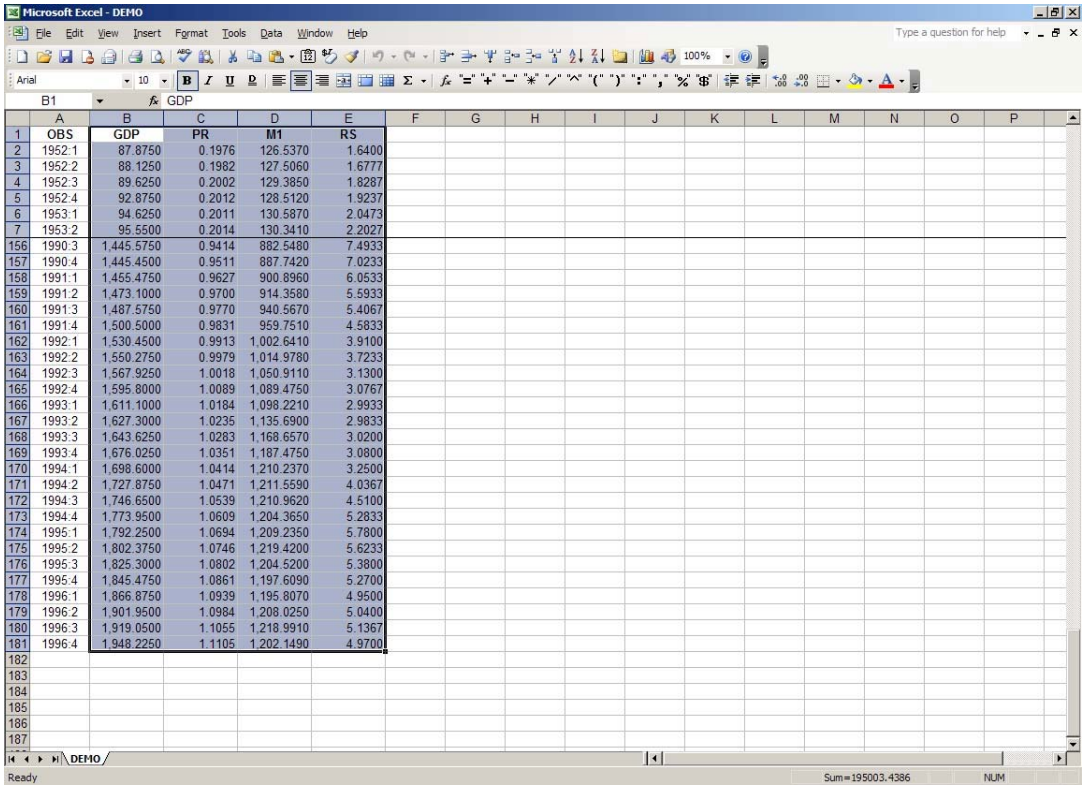
obs	A	B
1952Q1	NA	NA
1952Q2	NA	NA
1952Q3	NA	NA
1952Q4	NA	NA
1953Q1	NA	NA
1953Q2	NA	NA
1953Q3	NA	NA
1953Q4	NA	NA
1954Q1	NA	NA
1954Q2	NA	NA
1954Q3	NA	NA
1954Q4	NA	NA
1955Q1	NA	NA
1955Q2	NA	NA
1955Q3	NA	NA
1955Q4	NA	NA
1956Q1	NA	NA
1956Q2	NA	NA
1956Q3	NA	NA
1956Q4		

ขั้นตอนที่ 5 ภายหลังจากกดปุ่ม Edit+/- แล้วให้ลาก Mouse เพื่อเลือกคอลัมน์ A และ B ดังรูป



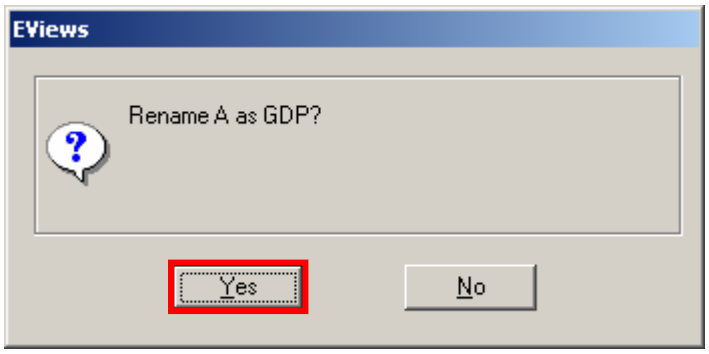
obs	A	B
1952Q1	NA	NA
1952Q2	NA	NA
1952Q3	NA	NA
1952Q4	NA	NA
1953Q1	NA	NA
1953Q2	NA	NA
1953Q3	NA	NA
1953Q4	NA	NA
1954Q1	NA	NA
1954Q2	NA	NA
1954Q3	NA	NA
1954Q4	NA	NA
1955Q1	NA	NA
1955Q2	NA	NA
1955Q3	NA	NA
1955Q4	NA	NA
1956Q1	NA	NA
1956Q2	NA	NA
1956Q3		

ขั้นตอนที่ 6 ต่อมาให้ไป Copy ข้อมูลจากตาราง Excel พร้อมทั้งชื่อตัวแปรดังรูป

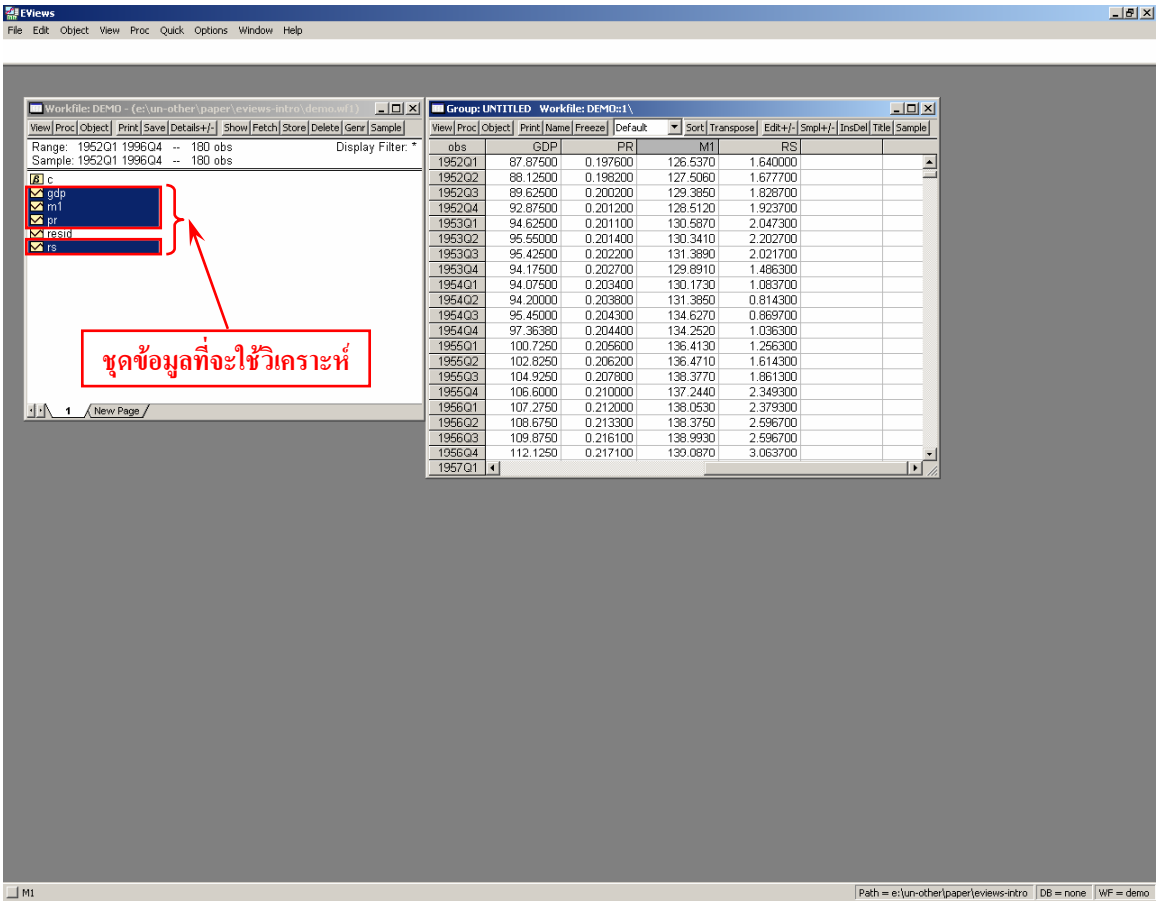


	A	B	C	D	E
1	OBS	GDP	PR	M1	RS
2	1952:1	87.8750	0.1976	126.5370	1.6400
3	1952:2	88.1250	0.1982	127.5060	1.6777
4	1952:3	89.6250	0.2002	129.3850	1.8287
5	1952:4	92.8750	0.2012	128.5120	1.9237
6	1953:1	94.6250	0.2011	130.5870	2.0473
7	1953:2	95.5500	0.2014	130.3410	2.2027
156	1990:3	1,445.5750	0.9414	882.5480	7.4933
157	1990:4	1,445.4500	0.9511	887.7420	7.0233
158	1991:1	1,455.4750	0.9627	900.8960	6.0533
159	1991:2	1,473.1000	0.9700	914.3580	5.5933
160	1991:3	1,487.5750	0.9770	940.5670	5.4067
161	1991:4	1,500.5000	0.9831	959.7510	4.5833
162	1992:1	1,530.4500	0.9913	1,002.6410	3.9100
163	1992:2	1,550.2750	0.9979	1,014.9780	3.7233
164	1992:3	1,567.9250	1.0018	1,050.9110	3.1300
165	1992:4	1,595.8000	1.0089	1,089.4750	3.0767
166	1993:1	1,611.1000	1.0184	1,098.2210	2.9933
167	1993:2	1,627.3000	1.0235	1,135.6900	2.9833
168	1993:3	1,643.6250	1.0283	1,168.6570	3.0200
169	1993:4	1,676.0250	1.0351	1,187.4750	3.0800
170	1994:1	1,698.6000	1.0414	1,210.2370	3.2500
171	1994:2	1,727.8750	1.0471	1,211.5590	4.0367
172	1994:3	1,746.6500	1.0539	1,210.9620	4.5100
173	1994:4	1,773.9500	1.0609	1,204.3650	5.2833
174	1995:1	1,792.2500	1.0694	1,209.2350	5.7800
175	1995:2	1,802.3750	1.0746	1,219.4200	5.6233
176	1995:3	1,825.3000	1.0802	1,204.5200	5.3800
177	1995:4	1,845.4750	1.0861	1,197.6090	5.2700
178	1996:1	1,866.8750	1.0939	1,195.8070	4.9500
179	1996:2	1,901.9500	1.0984	1,208.0250	5.0400
180	1996:3	1,919.0500	1.1055	1,218.9910	5.1367
181	1996:4	1,948.2250	1.1105	1,202.1490	4.9700

ขั้นตอนที่ 7 นำข้อมูลที่ได้ Copy จาก Excel ไป Paste ในหน้าต่าง Group ที่ได้เตรียมไว้ ต้องดำเนินการเปลี่ยนชื่อตัวแปร A และ B ให้เป็นชื่อตัวแปร GDP และ PR เมื่อเปลี่ยนชื่อตัวแปรทั้งสอง จะปรากฏหน้าต่างดังรูป แล้วให้กดปุ่ม Yes เพื่อตกลงให้โปรแกรมเปลี่ยนชื่อตัวแปรจาก A เป็น GDP



ขั้นตอนที่ 8 หลังจากนั้นให้กดปุ่ม Edit+/- อีกครั้ง เพื่อปิดการ Edit ข้อมูล ในที่สุดเราก็จะมีชุดข้อมูล (Series) พร้อมที่จะวิเคราะห์ในครั้งนี้ คือ มี Series ของตัวแปร GDP M1 PR และ RS ดังรูป



ขั้นตอนที่ 9 ภายหลังจากนำข้อมูลเข้า EViews แล้ว ให้ปิดหน้าต่าง Group ซึ่งจะ Save ชื่อ Group หรือไม่ก็ได้ หลังจากปิดหน้าต่าง Group แล้ว ให้บันทึกข้อมูล โดยกดปุ่ม Save ที่หน้าต่าง **Workfile** แล้วเลือก Path Directory ที่ต้องการจะเก็บ Workfile นี้ไว้ พร้อมทั้งตั้งชื่อ File ที่ต้องการ นามสกุล File ของ EViews จะมีนามสกุล .wf1 ซึ่งต่อไปก็สามารถนำ File ดังกล่าวกลับมาวิเคราะห์ใหม่ได้อีก

➢ *เครื่องหมายทางคณิตศาสตร์ในโปรแกรม EViews*

โปรแกรม EViews ได้บรรจุเครื่องหมาย ฟังก์ชันทางคณิตศาสตร์ และสถิติไว้จำนวนมากมาย ซึ่งผู้ใช้สามารถเปิดดูได้จาก Function Reference ซึ่งอยู่ภายในเมนู Help ของโปรแกรม EViews นอกจากนี้ภายในโปรแกรม EViews ยังมีคำสั่งสำเร็จรูปที่เป็นลักษณะของ Command สำหรับการเขียนโปรแกรมในการวิเคราะห์ทางเศรษฐมิติ สำหรับผู้ที่ต้องการประยุกต์ใช้การวิเคราะห์ทางเศรษฐมิติในขั้นสูง โดยเฉพาะการวิเคราะห์ข้อมูลอนุกรมเวลา

โปรแกรม EViews จะทำงานตามคำสั่งจากซ้ายไปขวา และให้ความสำคัญกับเครื่องหมายทางคณิตศาสตร์ตามลำดับความสำคัญจากมากไปหาน้อยดังนี้

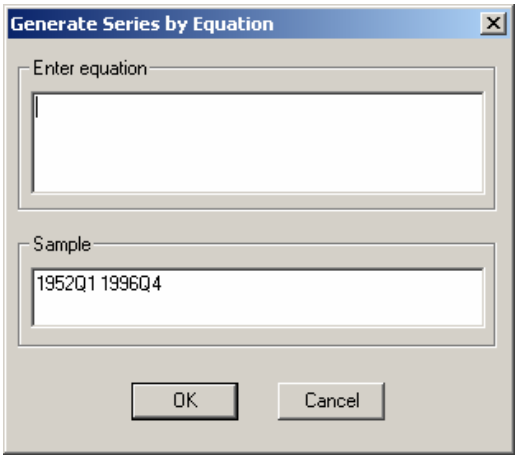
- ยกกำลัง (^)
- คูณ (*), หาร (/)
- น้อยกว่า (<), มากกว่า (>), น้อยกว่าหรือเท่ากับ (<=), มากกว่าหรือเท่ากับ (>=), เท่ากับ (=)
- and, or

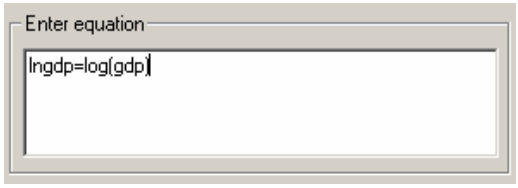
สำหรับรายละเอียดของการใช้เครื่องหมายทางคณิตศาสตร์ต่างๆ เหล่านี้ สามารถแสดงได้ดังตารางข้างล่างนี้

เครื่องหมาย	ความหมาย	วิธีการใช้และคำอธิบาย
+	บวก	X+Y : บวก X และ Y
-	ลบ	X-Y : ลบ Y จาก X
*	คูณ	X*Y : คูณ X โดย Y
/	หาร	X/Y : หาร X โดย Y
^	ยกกำลัง	X^Y : X ยกกำลัง Y
>	มากกว่า	X>Y : ให้ค่า 1 ถ้า X มากกว่า Y, ให้ค่า 0 กรณีอื่นๆ
<	น้อยกว่า	X<Y : ให้ค่า 1 ถ้า Y มากกว่า X, ให้ค่า 0 กรณีอื่นๆ
=	เท่ากับ	X=Y : ให้ค่า 1 ถ้า X และ Y เท่ากัน, ให้ค่า 0 กรณีอื่นๆ
<>	ไม่เท่ากับ	X<>Y : ให้ค่า 1 ถ้า X และ Y ไม่เท่ากัน, ให้ค่า 0 กรณีถ้าทั้งสองเท่ากัน
<=	น้อยกว่าหรือเท่ากับ	X<=Y : ให้ค่า 1 ถ้า X ไม่มากกว่า Y, ให้ค่า 0 กรณีอื่นๆ
>=	มากกว่าหรือเท่ากับ	X>=Y : ให้ค่า 1 ถ้า Y ไม่มากกว่า X, ให้ค่า 0 กรณีอื่นๆ
And	และ	X and Y : ให้ค่า 1 ถ้าทั้ง X และ Y ไม่เป็นศูนย์, ให้ค่า 0 กรณีอื่นๆ
Or	หรือ	X or Y : ให้ค่า 1 ถ้า X หรือ Y ไม่เป็นศูนย์, ให้ค่า 0 กรณีอื่นๆ

➢ *คำสั่งพื้นฐานในการ Transform ข้อมูล*

ขั้นตอนที่ 1 นำเข้าข้อมูลและทำการบันทึกข้อมูลให้เรียบร้อย ถ้าหากผู้ใช้มี Workfile เก่า สามารถที่จะเปิด Workdfile ดังกล่าวมาใช้ได้ โดยเลือก **File/Open/Workfile** ที่แถบเครื่องมือของหน้าต่างหลัก (**Main menu**) แล้วเลือก Workfile ที่ต้องการเปิด นอกจากนี้ในบางครั้งต้องการสร้าง/แปลง (Generate) ตัวแปรก็สามารถทำได้ โดยการเข้าไปที่ **Quick/Generate Series** ที่แถบเครื่องมือของหน้าต่างหลัก (**Main menu**) หรือกดปุ่ม Genr ที่แถบเครื่องมือของหน้าต่าง **Workfile** ภายหลังเลือกคำสั่งดังกล่าวแล้วจะปรากฏหน้าต่าง **Generate Series by Equation** ดังรูป





ขั้นตอนที่ 2 ใส่สมการหรือสูตรการคำนวณตามที่ต้องการในช่อง Enter equation ส่วนในช่อง Sample ให้ระบุขอบเขตกลุ่มตัวอย่างที่ต้องการ ในที่นี้จะนำเสนอสูตรหรือคำสั่งที่นิยมใช้กันโดยทั่วไปในการแปลงตัวแปร ดังนี้

1. การบวก (+) ลบ (-) คูณ (*) หาร (/) เช่น X1 + X2, X1 – X2, X1* X2 หรือ X1 / X2
2. การยกกำลัง ให้ใช้เครื่องหมาย ^ เช่น XX = X^2 เป็นต้น ดังนั้นถ้าต้องการยกกำลัง n หรือ X คูณกัน n ครั้ง ก็ให้ใช้คำสั่ง X^n
3. การถอดราก ถ้าเป็นการถอดรากที่สองของตัวแปรให้ใช้คำสั่ง SQR (X) หรือ X^(1/2) ถ้าต้องการถอดรากที่สาม หรือมากกว่าให้คำสั่งว่า X^(1/3) [รากที่สาม] เป็นต้น ดังนั้นถ้าต้องการถอดรากที่ n ก็ใช้คำสั่ง X^(1/n)
4. การแปลงตัวแปรโดยการใช้ Natural logarithm ให้ใช้คำสั่ง log (….) เช่น lnX = log(X) เป็นต้น และถ้าต้องการ Exponential X ให้ใช้คำสั่ง exp (....) เช่น X = exp(lnX) เป็นต้น
5. การสร้าง Lag variable ให้ใช้เครื่องหมาย (- n) [n คือ จำนวน Lag ที่ต้องการ] เช่น ต้องการสร้าง X_{t-1} ให้ใช้คำสั่งว่า X(-1) ซึ่งหมายความว่า X ณ เวลาที่ t – 1 ดังนั้นถ้าต้องการสร้าง X_{t-n} ก็ต้องคำสั่งว่า X(-n) ซึ่งหมายความว่า X ณ เวลาที่ t – n
6. การสร้าง Difference variable ให้ใช้คำสั่ง d(Variable, n) (Variable คือ ตัวแปรที่ต้องการ Difference ส่วน n คือ จำนวนการ Difference เช่น ต้องการสร้าง $\Delta X_t = X_t - X_{t-1}$ ให้เขียนคำสั่งว่า d(X, 1) และถ้าต้องการ Difference จำนวน n ครั้ง $\Delta X_t^n = \Delta X_t^{n-1} - \Delta X_{t-1}^{n-1}$

3.3 การสร้างกราฟ

ในการศึกษาหาความสัมพันธ์ หรือการหารูปแบบฟังก์ชันที่เหมาะสมในสมการ หรือการตรวจสอบรูปแบบของข้อมูลในเบื้องต้น นักวิจัยมักจะพิจารณาจากค่าสถิติพื้นฐานก่อน เช่น ดูจากค่าเฉลี่ย, ค่าสูงสุด, ค่าต่ำสุด, ค่ามัธยฐาน ค่าเบี่ยงเบนมาตรฐาน เป็นต้น นอกจากนี้ยังต้องพิจารณาจากกราฟ ซึ่งการพิจารณาจากกราฟจะเป็นวิธีการที่ง่ายที่สุดในการตรวจสอบความสัมพันธ์หรือตรวจสอบข้อมูลที่นำมาใช้ โปรแกรม EViews ได้มี Option ในการสร้างกราฟหลายๆ แบบ และมี Option ในการจัดการกราฟอยู่มากมายและสามารถทำได้ง่าย โดยมีขั้นตอนในการปฏิบัติดังนี้

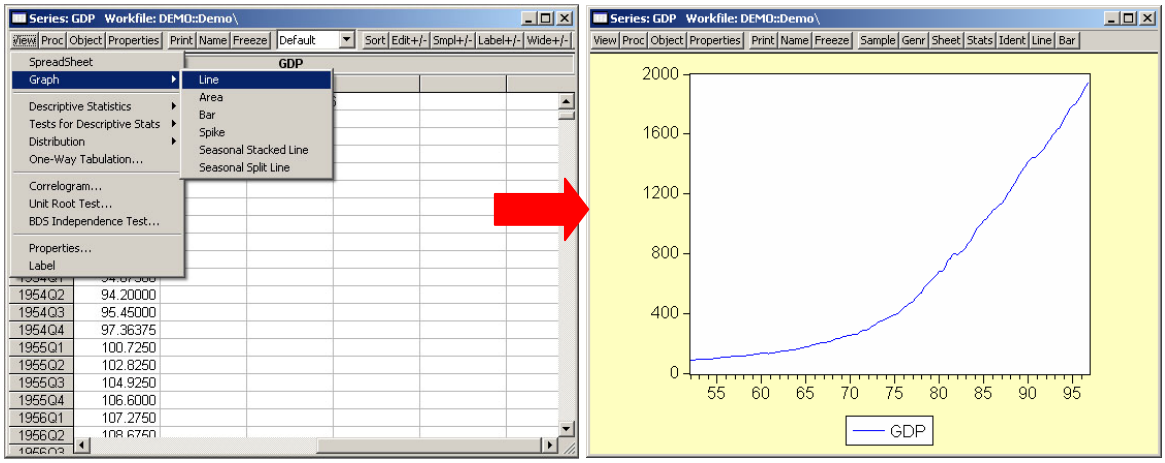
ขั้นตอนที่ 1 เปิด Workfile ที่ต้องการใช้ ในตัวอย่างนี้คือ “demo.wf1”

ขั้นตอนที่ 2 เลือกตัวแปรที่ต้องการสร้างกราฟ ซึ่งสามารถทำได้ตั้งแต่ 1 ตัวแปร ถึง n ตัวแปร โดยปกติมักจะใช้เพียง 2 ตัวแปร หรือสร้างกราฟแบบสองมิติเท่านั้น เช่น ในตัวอย่างต้องการพิจารณากราฟของ GDP (ผลิตภัณฑ์มวลรวมภายในประเทศ) และต้องการพิจารณาดูความสัมพันธ์ ระหว่าง M1 (อุปทานของเงิน) กับ GDP (ผลิตภัณฑ์มวลรวมภายในประเทศ)

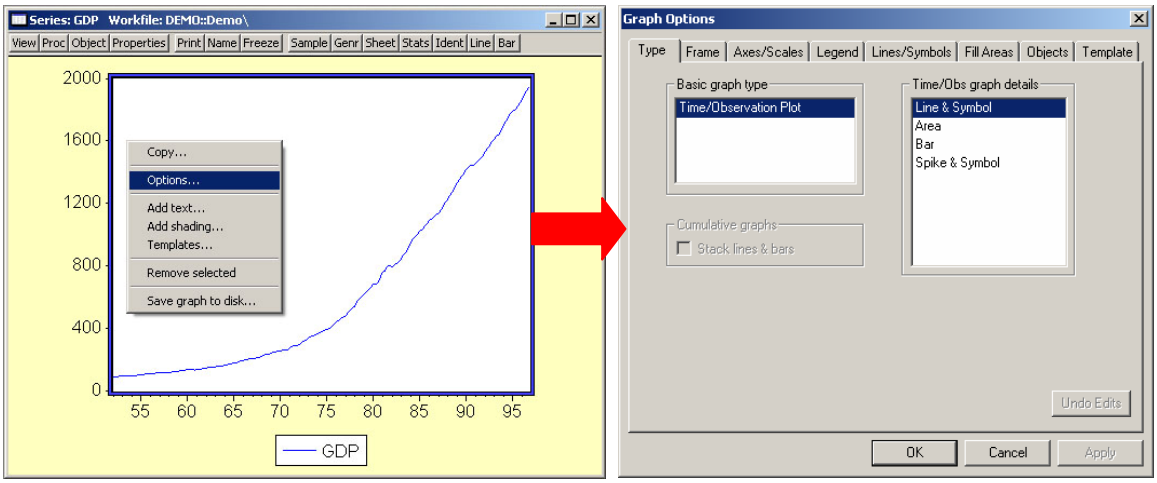
<u>การสร้างกราฟเส้นของ GDP</u>

ขั้นตอนที่ 1 เปิด Series GDP โดยการ **Double click** ที่ Series GDP หรือ คลิกขวา แล้วเลือก **Open** แล้วโปรแกรมก็จะเปิดหน้าต่าง **Series GDP**

ขั้นตอนที่ 2 เลือก **View/Graph/Line** โปรแกรมก็จะสร้างกราฟเส้นให้ตามที่ต้องการ



ขั้นตอนที่ 3 หากต้องการจัดการกราฟ หรือแก้ไข หรือต้องการเปลี่ยนแปลงรายละเอียดของกราฟบางอย่าง ก็สามารถทำได้โดยการคลิกด้านขวาของ Mouse ในรูปกราฟที่สร้างขึ้น แล้วเลือก Options หลังจากนั้นโปรแกรมก็จะเปิดหน้าต่าง **Graph Options** มาให้ดังรูป

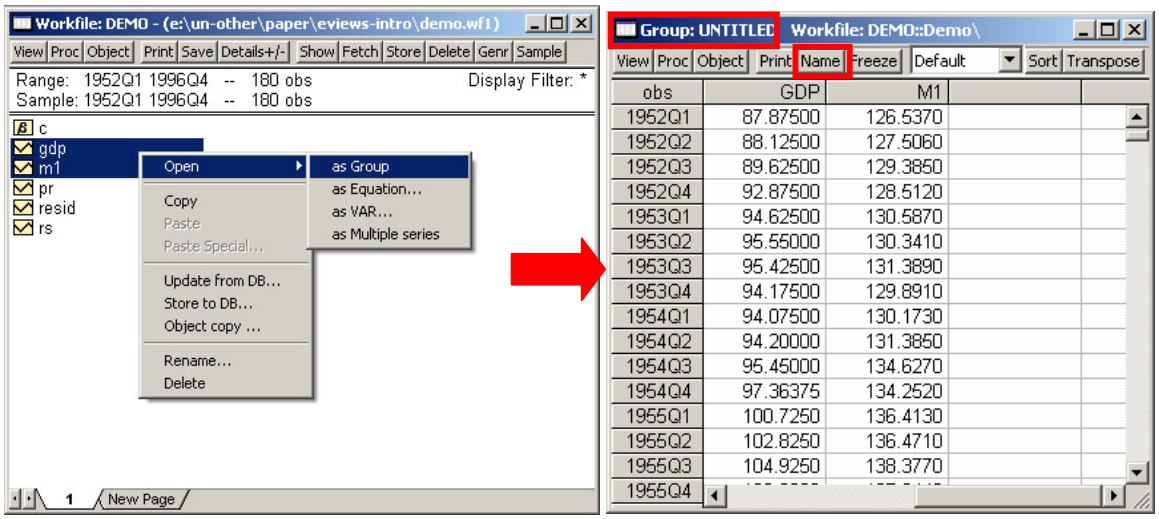


ภายใน Graph Options ผู้ใช้สามารถเปลี่ยนแปลง แก้ไขรายการ (Options) ต่างๆ ของกราฟได้มากมาย เช่น การเปลี่ยนแปลงชนิดของกราฟ ขนาดและลักษณะต่างๆ ของกราฟ การเปลี่ยนแปลงหน่วยหรือ Scale ของแกนของกราฟ การเปลี่ยนแปลงสี สัญลักษณ์ต่างๆ ของกราฟ เป็นต้น

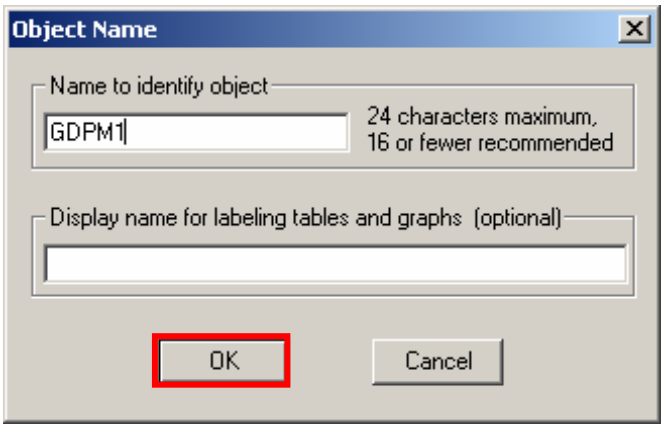
<u>การสร้างกราฟที่แสดงความสัมพันธ์ระหว่าง M1 กับ GDP</u>

ตามทฤษฎีที่เรียนมา บอกให้ทราบว่า เมื่อผลิตภัณฑ์มวลรวมภายในประเทศ (GDP) เพิ่มขึ้นจะทำให้อุปทานของเงินภายในประเทศ (M1) เพิ่มขึ้นตามไปด้วย หรืออาจกล่าวได้ว่า GDP กับ M1 มีความสัมพันธ์กันในทางบวก ซึ่งในเบื้องต้นจะทำการตรวจสอบความสัมพันธ์ดังกล่าวด้วยกราฟ

ขั้นตอนที่ 1 สร้าง Group ของ Series M1 และ GDP โดยการเลือก Series M1 และ GDP แล้วเลือก **Open/ as Group** แล้วโปรแกรมก็จะเปิดหน้าต่าง **Group: UNTITLED** มาให้

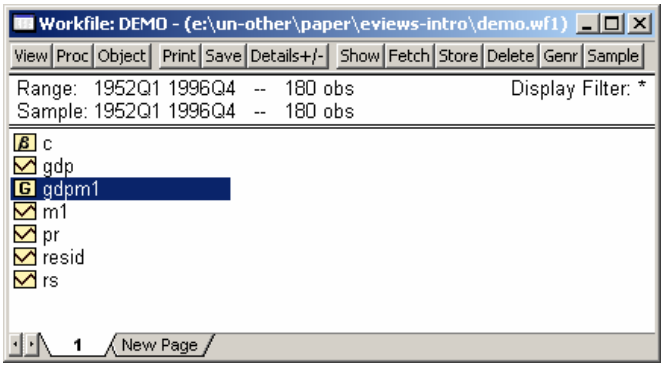


UNTITLED คือ ชื่อ Default ที่โปรแกรมตั้งให้ และถ้าต้องการตั้งชื่อ Group ด้วยตัวเองให้กดปุ่ม Name ที่แถบเครื่องมือของหน้าต่าง **Group** ก็จะปรากฏหน้าต่าง **Object Name** ดังรูป

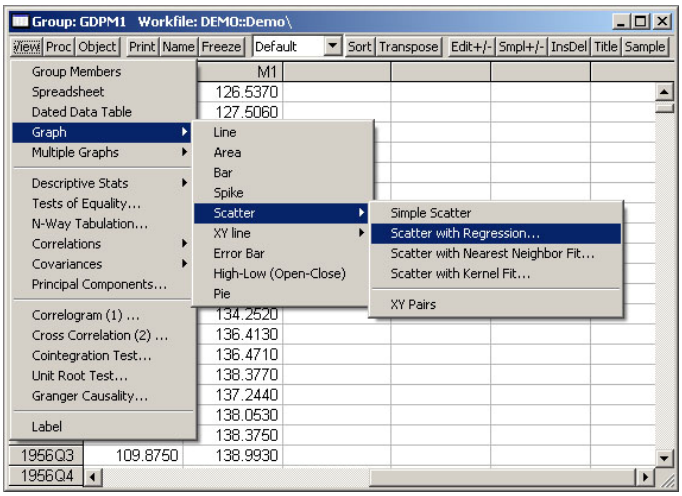


ระบุชื่อของ Group ที่ต้องการในช่อง Name to identify object ซึ่งจะต้องไม่เกิน 24 ตัวอักษร แต่โปรแกรมแนะนำมาว่าควรระบุที่ 16 หรือประมาณ 2 – 3 ตัวอักษร

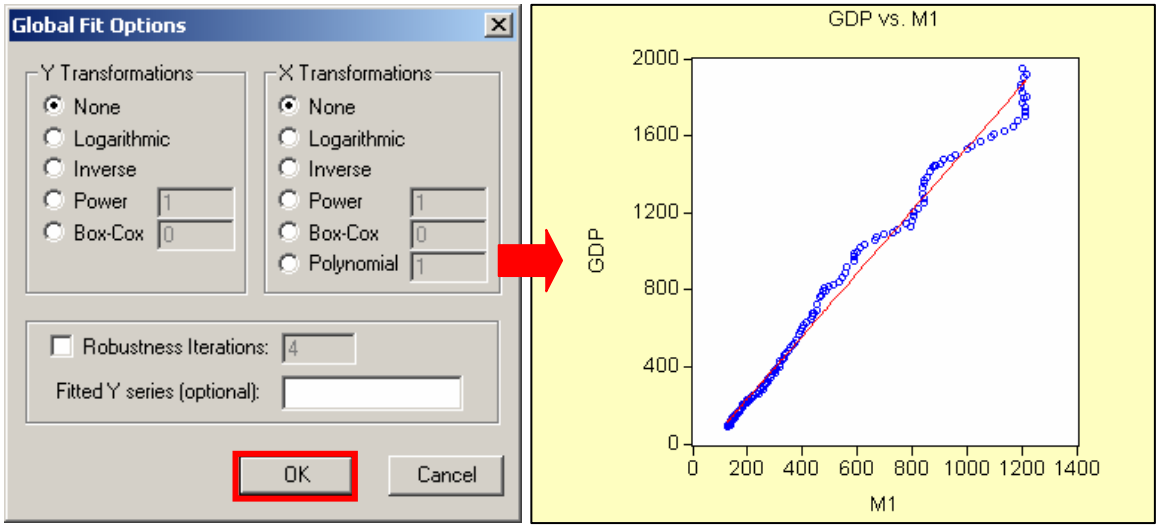
ส่วนช่องด้านล่างไม่ต้องเติมอะไร แต่ถ้าต้องการเติมอะไรไว้เตือนว่า Group ที่สร้างขึ้นมานี้เป็นมาอย่างไร ก็ให้ผู้ใช้พิมพ์สิ่งที่ต้องการลงในช่องดังกล่าวได้ตามความต้องการ หลังจากนั้นกดปุ่ม OK โปรแกรมจะสร้าง Objects Group ไว้ที่หน้าต่าง **Workfile** ตามชื่อที่สร้างขึ้นมาดังรูป ต่อมาก็ทำการบันทึก Workfile ดังนั้นเมื่อเปิด Workfile นี้ขึ้นมาใช้ ก็จะมี Objects Group gdpm1 ปรากฏทุกครั้ง



ขั้นตอนที่ 2 เลือก **View/Graph/Scatter/Scatter with Regression** ที่หน้าต่าง **Group: GDPM1** หมายความว่า ให้สร้างกราฟในรูปแบบ Scatter (การกระจาย) ระหว่าง GDP กับ M1 และให้สร้างเส้น Regression ที่แสดงถึงรูปแบบความสัมพันธ์ดังกล่าว โดยปกติ EViews จะใช้ Optimal – Linear Scaling เป็นค่า Default ในการประมาณค่าเส้น Regression และถ้าผู้ใช้ต้องการเปลี่ยนแปลงก็สามารถทำได้โดยไปแก้ไขที่ Graph Options



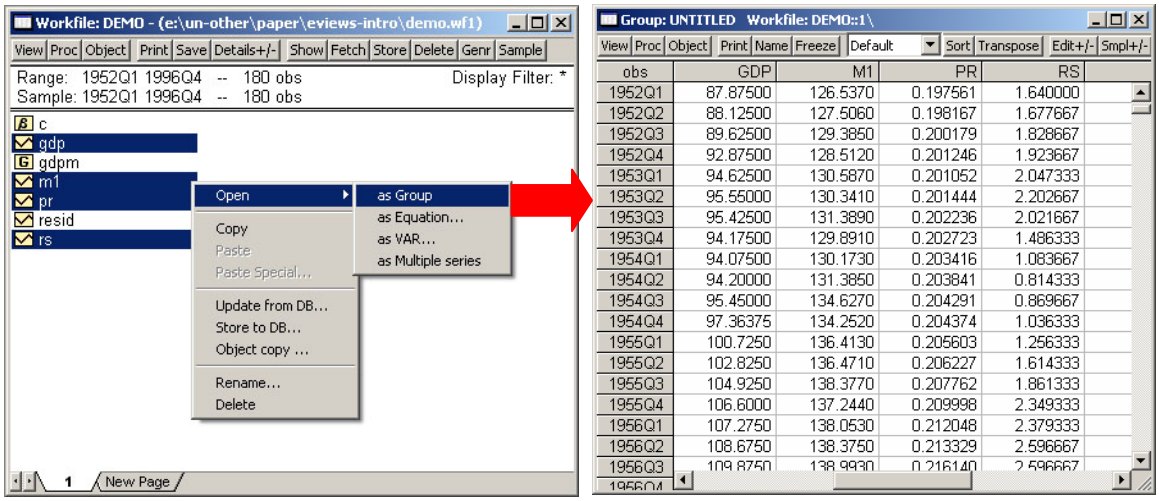
ขั้นตอนที่ 3 หลังจากการเลือกคำสั่งข้างต้นแล้วโปรแกรมEViews จะแสดงหน้าต่าง Global Fit Options ซึ่งเป็นหน้าต่างที่ให้ผู้ใช้ระบุว่าต้องการที่จะ Transformations ข้อมูลของตัวเองหรือไม่ ถ้าต้องการ Transformations ก็ระบุรูปแบบของการ Transformations ข้อมูล แล้วกดปุ่ม OK ถ้าไม่ต้องการ Transformations ข้อมูลก็เลือก **None** แล้วกดปุ่ม OK โปรแกรมก็จะสร้างกราฟที่เป็นการกระจาย (Scatter) และสร้างเส้นกราฟ Regression ที่เป็นความสัมพันธ์ระหว่าง GDP กับ M1



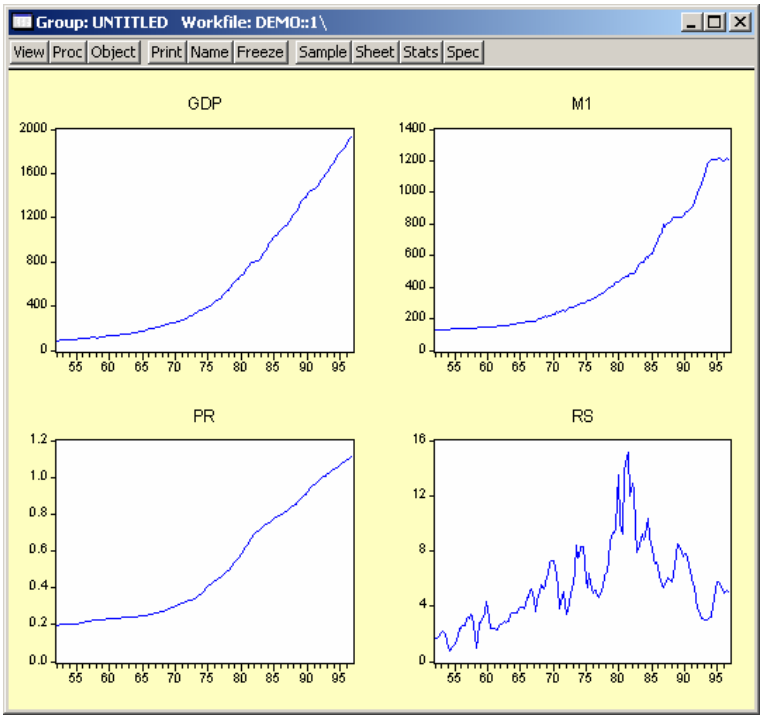
3.4 การตรวจสอบข้อมูล

ก่อนทำการวิเคราะห์ข้อมูลด้วยวิธีการทางเศรษฐมิติ ในเบื้องต้นต้องทำการตรวจสอบข้อมูลว่ามีคุณลักษณะอย่างไรบ้าง โดยเริ่มตั้งแต่การพิจารณาจากกราฟ การพิจารณาจากค่าสถิติพื้นฐาน การพิจารณาเมตริกสหสัมพันธ์ โดยโปรแกรม EViews ได้มีเครื่องมือสำหรับการตรวจสอบข้อมูลเบื้องต้น ดังนั้นในที่นี้จะขอนำเสนอขั้นตอนในการตรวจสอบข้อมูลเบื้องต้นต่างๆ ของข้อมูลตัวอย่างพอสังเขป ดังนี้

ขั้นตอนที่ 1 สร้าง Group ของ Series GDP M1 PR และ RS โดยการเลือก Series GDP M1 PR และ RS แล้วคลิก Mouse ทางด้านขวา ต่อมาให้เลือก **Open/as Group** โปรแกรมจะเปิดหน้าต่าง **Group: UNTITLED** มาให้ ซึ่งภายใน Group นี้จะประกอบด้วยสมาชิกของตัวแปรที่ต้องการดังรูป



ขั้นตอนที่ 2 <u>**การพิจารณาจากกราฟ**</u> ในกรณีนี้เป็นการพิจารณากราฟ หลายๆ ตัวแปรพร้อมๆ กันสามารถทำได้โดยการเลือก **View/Multiple Graphs/Line** จากแถบเครื่องมือที่หน้าต่าง **Group Objects** แล้วโปรแกรม EViews จะสร้างกราฟของแต่ละ Series ดังนี้



ขั้นตอนที่ 3 **<u>การพิจารณาจากค่าสถิติพื้นฐาน</u>** สามารถทำได้โดยเลือก **View/Descriptive Stats /Individual Sample** จากแถบเครื่องมือที่หน้าต่าง **Group Object** โปรแกรม EViews จะคำนวณและนำเสนอค่าสถิติของแต่ละ Series ค่าสถิติพื้นฐานที่โปรแกรม EViews คำนวนให้มีหลายตัว แต่ที่นิยมนำมาใช้พิจารณาประกอบด้วย ค่าเฉลี่ย (Mean), ค่ามัธยฐาน (Median), ค่าสูงสุด (Maximum), ค่าต่ำสุด (Minimum), ค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation), ค่าความเบ้ (Skewness), ค่าความโด่ง (Kurtosis) และค่าการกระจายตัวแบบปกติ (Jarque-Bera: JB)

Group: UNTITLED Workfile: DEMO::1\

View | Proc | Object | Print | Name | Freeze | Sample | Sheet | Stats | Spec

	GDP	M1	PR	RS
Mean	632.4190	445.0064	0.514106	5.412928
Median	374.3000	298.3990	0.383802	5.057500
Maximum	1948.225	1219.420	1.110511	15.08733
Minimum	87.87500	126.5370	0.197561	0.814333
Std. Dev.	564.2441	344.8315	0.303483	2.908939
Skewness	0.845880	0.997776	0.592712	0.986782
Kurtosis	2.345008	2.687096	1.829239	4.049883
Jarque-Bera	24.68300	30.60101	20.81933	37.47907
Probability	0.000004	0.000000	0.000030	0.000000
Sum	113835.4	80101.16	92.53909	974.3270
Sum Sq. Dev.	56988478	21284672	16.48625	1514.685
Observations	180	180	180	180

ขั้นตอนที่ 4 **<u>การพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation matrix)</u>** สามารถทำได้โดยการเลือก **View/ Correlations/Common Samples** จากแถบเครื่องมือที่หน้าต่าง **Group Object** ก็จะได้ตาราง Correlation matrix ดังรูป

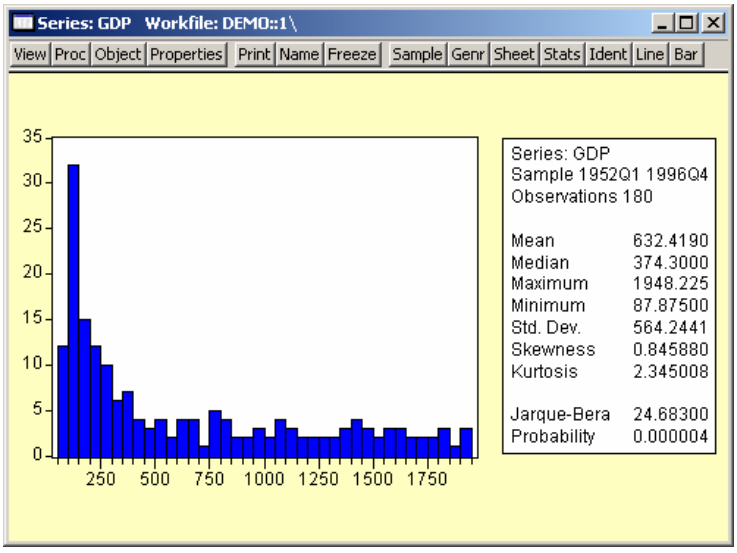
Group: UNTITLED Workfile: DEMO::1\

View | Proc | Object | Print | Name | Freeze | Sample | Sheet | Stats | Spec

Correlation Matrix

	GDP	M1	PR	RS
GDP	1.000000	0.995197	0.992475	0.333494
M1	0.995197	1.000000	0.980402	0.270059
PR	0.992475	0.980402	1.000000	0.412471
RS	0.333494	0.270059	0.412471	1.000000

นอกจากนี้ ยังสามารถแยกพิจารณาคุณลักษณะของข้อมูลเป็นราย Series ได้อีกด้วย โดยการเปิด Series ที่ต้องการแล้วเลือก **View/Descriptive Statistics /Histogram and Stats** หรือกดปุ่ม Stats จากแถบเครื่องมือของ Series ก็จะได้ผลการวิเคราะห์ดังรูป



<u>หมายเหตุ</u> สำหรับข้อมูลอนุกรมเวลา โปรแกรม EViews ได้พัฒนาเครื่องมือ และคำสั่งในการวิเคราะห์และทดสอบข้อมูลอนุกรมเวลา ไม่ว่าจะเป็นการทดสอบด้วย Correlogram การทดสอบ Unit root เป็นต้น ซึ่งผู้ใช้สามารถใช้ได้ที่ Views ที่หน้าต่าง **Series** โดยปกติแล้ว EViews จะมีขีดความสามารถในการวิเคราะห์ข้อมูลอนุกรมเวลาเหนือกว่าการวิเคราะห์ข้อมูลภาคตัดขวางหรือข้อมูล Panel /Poll

4. วิธีการประมาณค่าแบบกำลังสองน้อยที่สุด (Ordinary Least Square Method)

แนวความคิดของวิธีการประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีการประมาณค่าแบบกำลังสองน้อยที่สุด [เรียกสั้นๆ ว่า OLS] คือ การประมาณค่าเส้นการถดถอยที่สามารถหาได้ โดยการทำให้ผลบวกของกำลังสองของส่วนที่เบี่ยงเบนไปจากเส้นถดถอย [ค่าคลาดเคลื่อน: Error term (ส่วนใหญ่จะแทนด้วยสัญลักษณ์ e / ε)] ของค่าสังเกตของตัวแปรมีค่าน้อยที่สุด วิธีการประมาณค่า OLS เป็นวิธีการประมาณค่าที่สำคัญและมักถูกหยิบยกมาใช้บ่อยในทางเศรษฐมิติ [รายละเอียดทางทฤษฎีสามารถหาอ่านได้ในหนังสือเศรษฐมิติ สำหรับในเอกสารฉบับนี้จะไม่กล่าวถึงรายละเอียดของวิธีการ OLS] วิธีการนี้มีข้อสมมติบางประการดังนี้

4.1 ข้อสมมติพื้นฐานของวิธีการกำลังสองน้อยที่สุด

การใช้วิธีกำลังสองน้อยที่สุดในการประมาณค่าสัมประสิทธิ์ของแบบจำลอง ต้องอยู่ภายใต้ข้อสมมติพื้นฐานดังต่อไปนี้

ก. รูปแบบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระรวมถึงตัวคลาดเคลื่อนจะต้องเป็นแบบเส้นตรง [$y_i = \alpha + \beta x_i + \varepsilon_i$]

ข. ตัวแปรอิสระต้องไม่ใช่ตัวแปรสุ่ม กล่าวคือ มีค่าแน่นอน (ตัวแปรอิสระจะต้องเป็น Non – Stochastic Variable)

ค. ตัวแปรอิสระต้องไม่มีความสัมพันธ์เชิงเส้นอย่างสมบูรณ์ หรือตัวแปรอิสระไม่ควรมีความสัมพันธ์เชิงเส้นกันสูงมากเกินไป [$\text{Corr}(X_i, X_j \neq 1)$] [ถ้าดูจากค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficients) ควรมีค่าไม่ควรเกิน 0.8] หรือไม่ควรมีปัญหา Multicollinearity

ง. ตัวคลาดเคลื่อน (Error term) จะต้องมีการกระจายแบบปกติ โดยมีค่าเฉลี่ยเป็นศูนย์และมีค่าความแปรปรวนคงที่ [$\varepsilon_i \sim N(0, \sigma^2)$, โดยที่ $E(\varepsilon_i) = 0$ และ $E(\varepsilon_i^2) = \sigma^2$ หรือเป็น Homoskedasticity]

จ. ตัวคลาดเคลื่อนจะต้องไม่มีสหสัมพันธ์ระหว่างกันเอง หรือตัวคลาดเคลื่อนจะต้องมีการกระจายที่เป็นอิสระแก่กัน [$\text{Cov}(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i, \varepsilon_j) = 0$ สำหรับทุกค่าที่ $i \neq j$] นั่นก็คือ ไม่มีปัญหา Autocorrelation

ฉ. ตัวแปรอิสระจะต้องไม่มีความสัมพันธ์กับตัวคลาดเคลื่อน

4.2 รูปแบบฟังก์ชันที่ใช้ในสมการถดถอย

ในการประมาณค่าสัมประสิทธิ์ของตัวแปรต่างๆ มักจะมีแบบจำลองที่ใช้ในการประมาณค่าได้หลายๆ แบบจำลอง ขึ้นอยู่กับลักษณะของข้อมูล และทฤษฎี หรือแนวคิด ปกติโดยส่วนใหญ่มักจะใช้แบบจำลองหรือฟังก์ชันในรูปแบบเส้นตรง (Linear) เนื่องจากเป็นรูปแบบที่ง่ายต่อการทำความเข้าใจ สำหรับในที่นี้จะนำเสนอรูปแบบฟังก์ชันที่มักนิยมใช้ในการวิเคราะห์ ดังนี้

รูปแบบฟังก์ชัน	สมการ	คำสั่ง EViews
Linear	$Y = \alpha + \beta_1 X_1 + \beta_2 X_2$	Y C X_1 X_2
Double – Log	$\ln Y = \alpha + \beta_1 \ln X_1 + \beta_2 \ln X_2$	log(Y) C log(X_1) log(X_2)
Linear – Log	$Y = \alpha + \beta_1 \ln X_1 + \beta_2 X_2$	Y C log(X_1) X_2
Log – Linear	$\ln Y = \alpha + \beta_1 X_1 + \beta_2 X_2$	Log(Y) C X_1 X_2
Polynomial	$Y = \alpha + \beta_1 X_1 + \beta_2 X_1^2 + \beta_3 X_2$	Y C X_1 X_1^2 X_2
Inverse	$Y = \alpha + \beta_1 (1/X_1) + \beta_2 X_2$	Y C 1/X_1 X_2
Dummy	$Y = \alpha + \beta_1 X_1 + \beta_2 D_1$	Y C X_1 D_1
Dummy (Interaction with Variable)	$Y = \alpha + \beta_1 X_1 + \beta_2 D_1 + \beta_3 D_1 X_1$	Y C X_1 D_1 D_1*X_1

ที่มา : R.R. Johnson (2000).

การตัดสินใจเลือกว่าจะนำแบบจำลองหรือสมการไหนมาใช้ จะต้องพิจารณาจากค่าสถิติในการตัดสินใจ โดยทั่วไปจะใช้ค่า R^2, $\overline{R}^2$ (Adjusted R^2), Quasi R^2, F–statistic และ AÎC Akaike's Information Criterion โดยแต่ละค่ามีเงื่อนไขในการตัดสินใจดังนี้

ค่าสถิติในการตัดสินใจ	สูตรคำนวณ	เงื่อนไข
R^2	$R^2 = 1 - \frac{\hat{\varepsilon}'\hat{\varepsilon}}{(y-\bar{y})'(y-\bar{y})}$	ค่ายิ่งสูงยิ่งดี เพราะ แสดงว่าตัวแปรอิสระสามารถอธิบายตัวแปรตามได้ดี (แต่ควรใช้ Quasi R^2 ในการเปรียบเทียบ)
$\overline{R}^2$ (Adjusted R^2)	$\overline{R}^2 = 1 - (1-R^2)\frac{T-1}{T-k}$	ควรมีค่าใกล้เคียง R^2 ซึ่งแสดงให้เห็นว่าการเพิ่มตัวแปรอิสระไม่ได้ส่งผลต่อ R^2
F – statistic	$F = \frac{\sum(\hat{y}-\bar{y})^2/k}{\sum\varepsilon^2/(n-k-1)}$	ค่า F-statistic ต้องมีค่ามากพอที่จะทำให้ค่า P-value ของ F-statistic $< \alpha$ จึงจะแสดงว่าค่าสัมประสิทธิ์ของตัวแปรอิสระมีค่าแตกต่างไปจากศูนย์
AÎC (Akaike's information criterion)	$AIC = 2lT + 2k/T$	ค่ายิ่งต่ำยิ่งดี แสดงว่า ค่าคลาดเคลื่อนมีความแปรปรวนน้อย

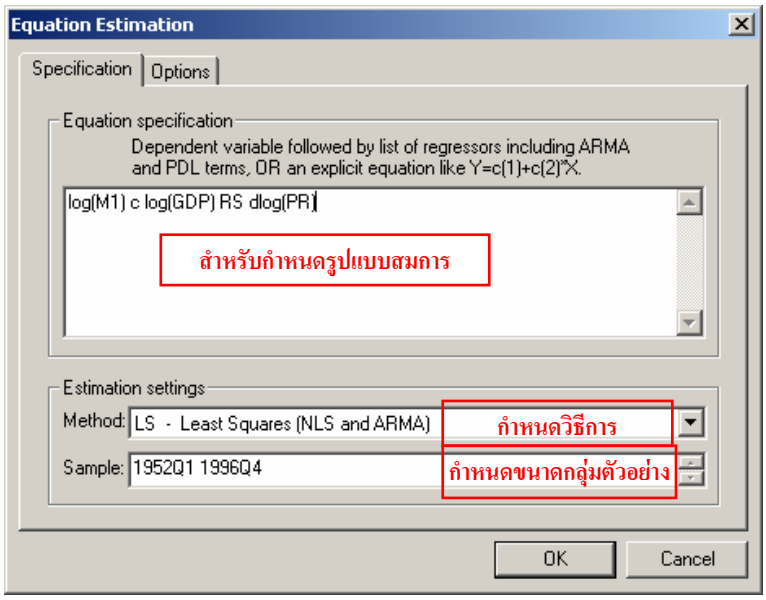
4.3 การใช้โปรแกรม EViews ในการประมาณค่าสัมประสิทธิ์ของสมการถดถอย

จากตัวอย่างที่บันทึกอยู่ใน File "demo.wf1" ซึ่งภายใน File ดังกล่าวเป็นข้อมูลอนุกรมเวลารายไตรมาสตั้งแต่ปี ค.ศ. 1952 ไตรมาสที่ 1 ถึงปี ค.ศ. 1996 ไตรมาสที่ 2 ประกอบด้วยตัวแปรทั้งหมด 4 ตัว คือ ตัวแปรอุปทานของเงิน (M1), ผลิตภัณฑ์มวลรวมภายในประเทศ (GDP), อัตราดอกเบี้ยระยะสั้น (RS) และระดับราคาสินค้าที่เป็นตัวแสดงถึงภาวะเงินเฟ้อ (PR) สำหรับในหัวข้อนี้ต้องการประมาณค่าสมการถดถอย โดยมีรูปแบบฟังก์ชันที่ใช้ในกรณีตัวอย่างดังนี้

ขั้นตอนที่ 1 เปิด Workfile “demo.wf1”

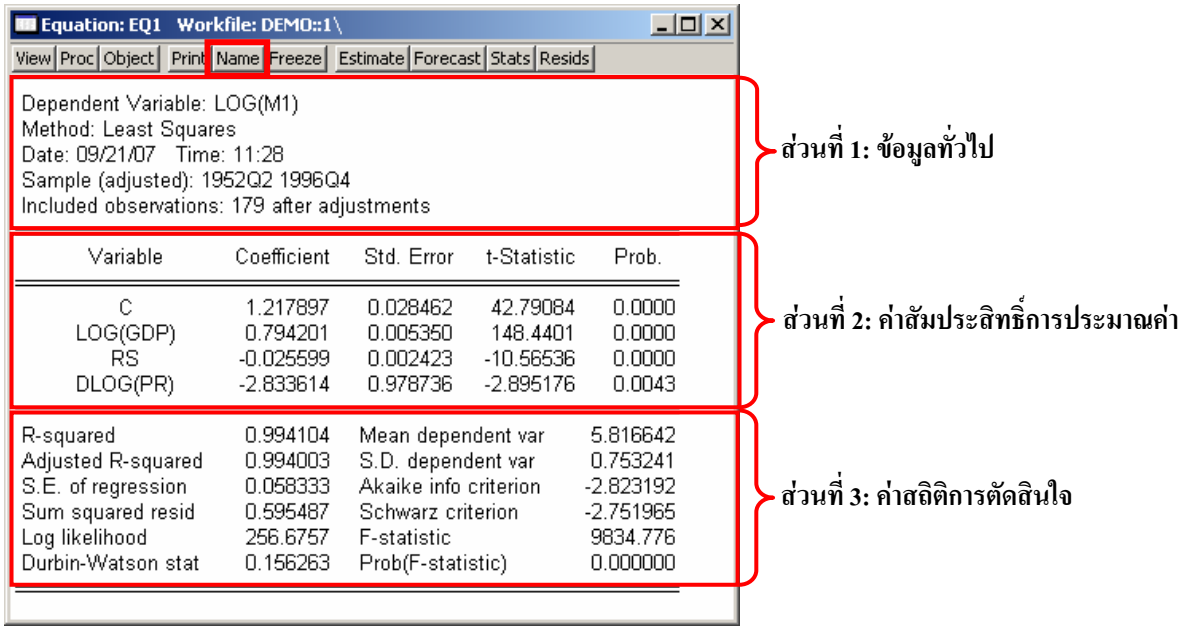
ขั้นตอนที่ 2 เลือก **Objects/New Object/Equation** จากเมนูในหน้าต่าง **Workfile** แล้วตั้งชื่อ Equation หรือเลือก **Quick/Estimate Equation** จากเมนูหลัก (**Main menu**) โปรแกรม EViews เปิดหน้าต่าง **Equation Specification** ดังรูป

ขั้นตอนที่ 3 ใส่ตัวแปรลงในช่อง Equation specification โดยให้ตัวแปรตามอยู่ทางซ้ายมือ แล้วตามด้วยตัวแปรอิสระ และถ้าต้องการให้มีค่าคงที่ให้ใส่ตัวอักษร “**C**” ทุกๆ ครั้งที่ป้อนตัวแปรในช่อง Equation specification ต้องให้ตัวแปรตัวแรกเป็นตัวแปรตามเสมอ



ขั้นตอนที่ 4 เลือกวิธีการประมาณค่าในช่อง “Method” และกำหนดขนาดตัวอย่างในช่อง “Sample” แล้วกดปุ่ม OK จะได้ผลลัพธ์ ดังนี้

ขั้นตอนที่ 5 บันทึก Objects Equation ด้วยการกดปุ่ม Name ที่แถบเครื่องมือของหน้าต่าง **Equation** แล้วตั้งชื่อให้ Objects นี้ เช่น “EQ1” ที่หน้าต่าง Workfile จะปรากฏ Objects eq1 [eq1] หลังจากนั้นให้ Save Workfile นี้ ดังนั้นถ้าหากเปิด Workfile นี้ขึ้นมาทำงานใหม่ก็จะพบว่ามี Objects eq1 แสดงอยู่ตลอด



เมื่อโปรแกรม EViews ประมาณค่าสัมประสิทธิ์ของสมการเสร็จแล้ว ก็จะรายงานผลการประมาณค่าสัมประสิทธิ์ดังกล่าว ใน Objects Equation ดังรูป ซึ่งภายใน Objects ดังกล่าว ประกอบด้วยส่วนย่อยๆ 3 ส่วน ดังนี้

<u>ส่วนที่ 1: ข้อมูลทั่วไป</u> ประกอบด้วย

- ชื่อของตัวแปรตาม (Dependent Variable: ……)
- วิธีการประมาณค่าสัมประสิทธิ์ (Method: …….)
- วัน เดือน ปี และเวลาที่ทำงาน (Date: …… Time: …..)
- ช่วงของตัวอย่างที่ใช้ (Sample: ……)
- จำนวนตัวอย่างทั้งหมด (Included observations: …..)

<u>ส่วนที่ 2: ค่าสัมประสิทธิ์การประมาณค่า</u> ประกอบด้วย

- คอลัมน์ที่ 1 เป็นชื่อของตัวแปร (Variable)
- คอลัมน์ที่ 2 เป็นค่าสัมประสิทธิ์ที่ได้จากการประมาณค่า (Coefficient)
- คอลัมน์ที่ 3 เป็นค่าความคลาดเคลื่อนมาตรฐานของตัวประมาณค่าสัมประสิทธิ์ (Std. Error)
- คอลัมน์ที่ 4 เป็นค่าสถิติ t (t–statistic) ได้มาจากการนำค่าในคอลัมน์ที่ 2 หาร คอลัมน์ที่ 3 (Coefficient / Std. Error)
- คอลัมน์ที่ 5 เป็นค่า Probability value (Prob.) [P-value] จะมีค่าอยู่ระหว่าง 0 – 1 ค่านี้เป็นค่าความน่าจะเป็นที่ค่า t-statistic ที่คำนวณได้มีค่ามากกว่าค่า t-statistic มาตรฐาน โดยจะเปรียบเทียบค่า P-value ที่ค่าคำนวณได้กับค่า α (ระดับนัยสำคัญที่กำหนด) ถ้าค่า P-value $< \alpha$ แสดงว่า ปฏิเสธสมมติฐานหลัก

<u>ส่วนที่ 3: ค่าสถิติการตัดสินใจ</u> โปรแกรม EViews จะนำเสนอค่าสถิติต่างๆ จำนวน 12 ค่า ดังนี้

- R – Squared (R^2) : Coefficient of Determination ค่า R^2 จะมีค่าอยู่ระหว่าง 0 – 1 การพิจารณาค่า R^2 ให้พิจารณาว่าค่า R^2 ยิ่งวิ่งเข้าใกล้หนึ่งยิ่งดี โดยปกติถ้าเป็นข้อมูลภาคตัดขวาง R^2 จะมีค่าอยู่ระหว่าง 0.4 – 0.6 แต่ถ้าเป็นข้อมูลอนุกรมเวลา ค่า R^2 จะมีค่าอยู่ระหว่าง 0.7 – 0.9
- Adjusted R – Squared ($\overline{R}^2$): เป็นค่า R^2 ที่มีการปรับค่าเพื่อไม่ให้เกิดปัญหาในกรณีที่มีการเพิ่มตัวแปรอิสระเข้าไปในสมการ โดยปกติถ้าแบบจำลองมี 2 ตัวแปร ค่า R^2 กับ $\overline{R}^2$ จะมีค่าใกล้เคียงกัน แต่ถ้าแบบจำลองมีตัวแปรมากกว่า 2 ตัว ค่า R^2 จะมีค่ามากกว่า $\overline{R}^2$ และถ้าค่า R^2 มีค่าน้อยมาก จะทำให้ $\overline{R}^2$ มีค่าติดลบได้
- S.E. of regression: คือ ค่าความคลาดเคลื่อนมาตรฐานของ Regression หรือค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า
- Sum squared resid: คือ ผลรวมกำลังสองของค่าคลาดเคลื่อน โดยวิธีการ OLS จะพยายามเลือกค่าสัมประสิทธิ์ที่ทำให้ค่านี้มีค่าน้อยที่สุด
- Log likelihood: นำไปใช้ประโยชน์ในการทดสอบสมมติฐาน
- Durbin – Watson stat: เป็นค่าที่ใช้ในการพิจารณาว่าแบบจำลองมีปัญหา Autocorrelation หรือไม่ โดยถ้าค่า D.W. มีค่าใกล้เคียง 2 แสดงว่าแบบจำลองที่กำลังพิจารณาไม่มีปัญหา Autocorrelation
- Mean dependent var: คือ ค่าเฉลี่ยของตัวแปรตาม
- S.D. dependent var: คือ ค่าเบี่ยงเบนมาตรฐานของตัวแปรตาม

- Akaike info criterion: คือ ค่าที่ใช้ในการพิจารณาเพื่อเลือกแบบจำลอง
- Schwarz criterion: คือ ค่าที่ใช้ในการพิจารณาเพื่อเลือกแบบจำลอง
- F–statistic: เป็นค่าที่ใช้ในการทดสอบสมมติฐานว่าค่าสัมประสิทธิ์ที่ประมาณค่าได้มีค่าแตกต่างไปจากศูนย์หรือไม่ ค่านี้ยิ่งสูงยิ่งดี
- Prob(F – statistic): เป็นค่า P–value ของ F–statistic ค่า P-value จะต้องน้อยกว่า α ที่กำหนดไว้ ถึงจะสามารถปฏิเสธสมมติฐานหลักที่ว่าค่าสัมประสิทธิ์ที่ประมาณค่าได้มีค่าไม่แตกต่างไปจากศูนย์

4.4 การอ่านผลการประมาณค่าสัมประสิทธิ์ที่ได้จากโปรแกรม EViews

ผลการประมาณค่าสัมประสิทธิ์ด้วยวิธีการ OLS ที่ได้จากโปรแกรม EViews สามารถเขียนเป็นสมการได้ดังนี้

$$\ln(M1_t) = 1.218 + 0.794(\ln GDP_t) - 0.026RS_t - 2.834\Delta\ln(PR_t)$$

(t-statistic) (42.791)(148.441) (-10.565) (-2.895)

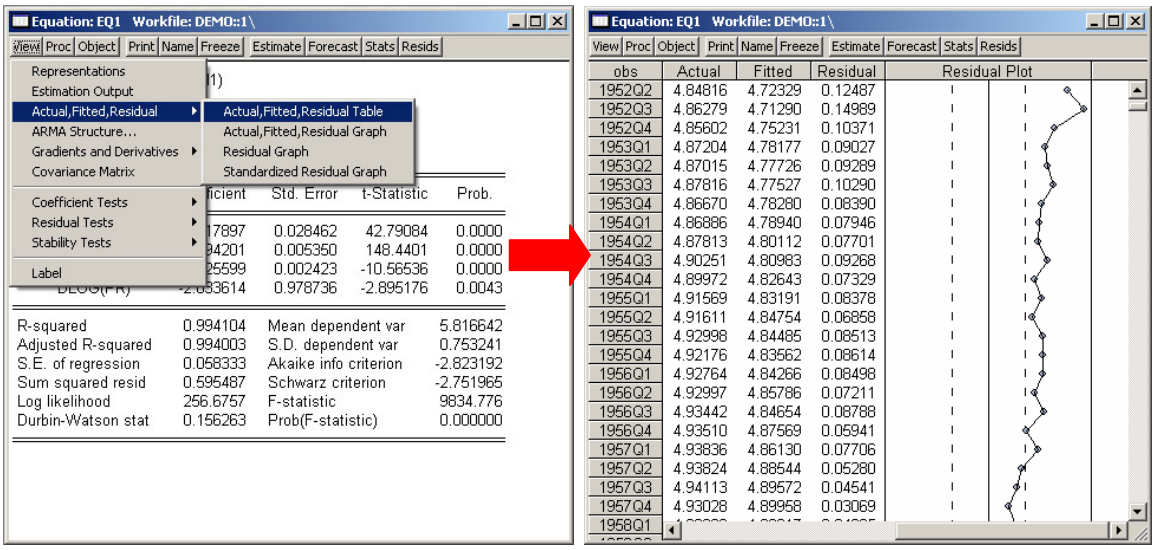
$R^2 = 0.9941$ $\overline{R}^2 = 0.9940$ D.W. = 0.1563 F – statistic = 9,834 (Prob. = 0.0000)

จากสมการข้างต้นซึ่งเป็นสมการอุปทานของเงิน จะเห็นได้ว่าสมการนี้ตัวแปรอิสระสามารถอธิบายตัวแปรตามได้ถึงร้อยละ 99 [พิจารณาจากค่า R^2] และจากการพิจารณาค่า F-statistic ที่เป็นการทดสอบสมมติฐานที่ว่า ค่าสัมประสิทธิ์ที่คำนวณทั้งหมดมีค่าเท่ากับศูนย์หรือไม่ ปรากฏว่า ค่าสัมประสิทธิ์ที่คำนวณได้มีค่าแตกต่างไปจากศูนย์ ณ ระดับความเชื่อมั่นที่ 99 เปอร์เซ็นต์ [พิจารณาจากค่า F-statistic ซึ่งมากกว่าค่าวิกฤติ ณ ระดับความเชื่อมั่นที่ 99% หรือ Prob. $< \alpha$] แต่เป็นที่น่าสังเกตว่าค่า D.W. มีค่าต่ำกว่า D_L และ D_U [ปกติค่า D.W. ที่ดีควรมีค่าใกล้เคียง 2] แสดงว่า สมการนี้มีปัญหา Autocorrelation อย่างแน่นอน [ตัวคลาดเคลื่อนมีสหสัมพันธ์กัน] ดังนั้นจึงยังไม่ควรนำสมการนี้ไปใช้ในการพยากรณ์หรืออธิบาย ควรดำเนินการแก้ไขปัญหา Autocorrelation ก่อน แต่ถ้าสมมติให้สมการนี้ไม่มีปัญหา Autocorrelation ก็สามารถอธิบายเบื้องต้นได้ว่า ผลิตภัณฑ์มวลรวมภายในประเทศ (GDP) มีความสัมพันธ์กับอุปทานของเงินในทิศทางเดียวกัน ในขณะที่อัตราดอกเบี้ยระยะสั้น (RS) และภาวะเงินเฟ้อ (PR) จะมีความสัมพันธ์กับอุปทานของเงินในทิศทางตรงกันข้าม ซึ่งผลที่ได้สอดคล้องกับทฤษฎีทางด้านเศรษฐศาสตร์ที่ว่า อุปทานของเงินจะแปรผันตรงกับผลิตภัณฑ์มวลรวมภายในประเทศ และแปรผกผันกับอัตราดอกเบี้ยระยะสั้น และภาวะเงินเฟ้อ

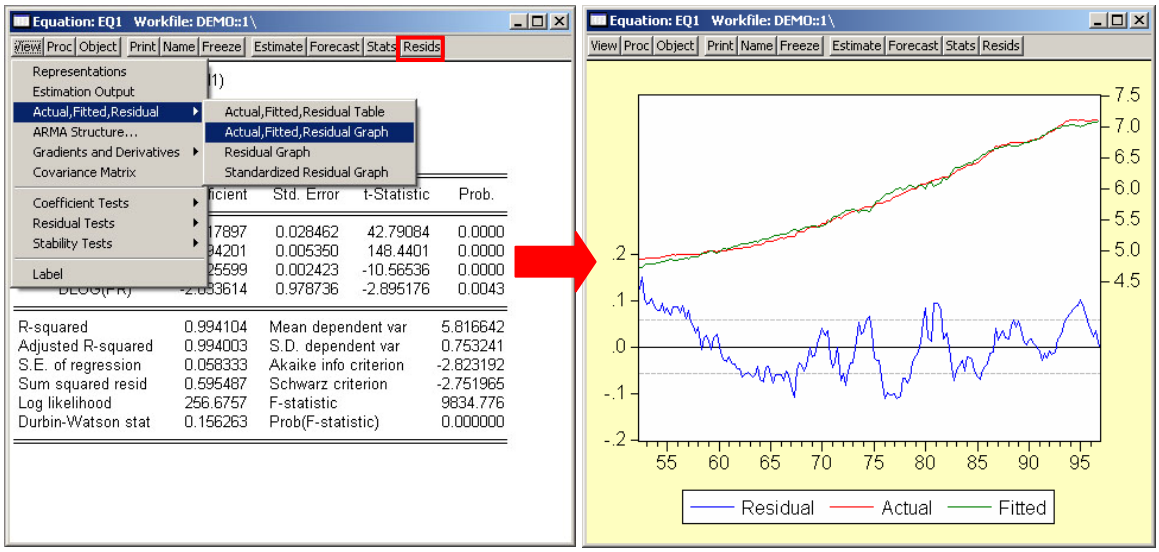
4.5 การแสดงผลค่าจริง ค่าพยากรณ์ และค่าความคลาดเคลื่อน

หลังจากการประมาณค่าสัมประสิทธิ์เสร็จแล้ว ถ้าหากผู้ใช้ต้องการดู/ต้องการเปรียบเทียบค่าจริง (ใน EViews เรียกว่า “Actual”) กับค่าพยากรณ์ (ใน EViews เรียกว่า “Fitted”) รวมทั้งต้องการดูค่าความคลาดเคลื่อน (ใน EViews เรียกว่า “Residual”) ก็สามารถทำได้ 2 วิธีการ ดังต่อไปนี้

วิธีการที่ 1 ต้องการดูตารางที่แสดงค่า Actual, Fitted และ Residual สามารถทำได้โดยการเลือก **View/Actual, Fitted,Residual/Actual,Fitted,Residual Table** จากแถบเครื่องมือของหน้าต่าง **Equation** ก็จะได้ตารางที่แสดงค่า Actual, Fitted และ Residual พร้อมทั้ง Residual Plot ดังรูป



วิธีการที่ 2 ต้องการดูกราฟที่แสดงค่า Actual, Fitted และ Residual สามารถทำได้โดยการเลือก **View/Actual, Fitted,Residual/Actual,Fitted,Residual Graph** หรือ กดปุ่ม Resids จากแถบเครื่องมือของหน้าต่าง **Equation** ก็จะได้กราฟที่แสดงค่า Actual, Fitted และ Residual ดังรูป

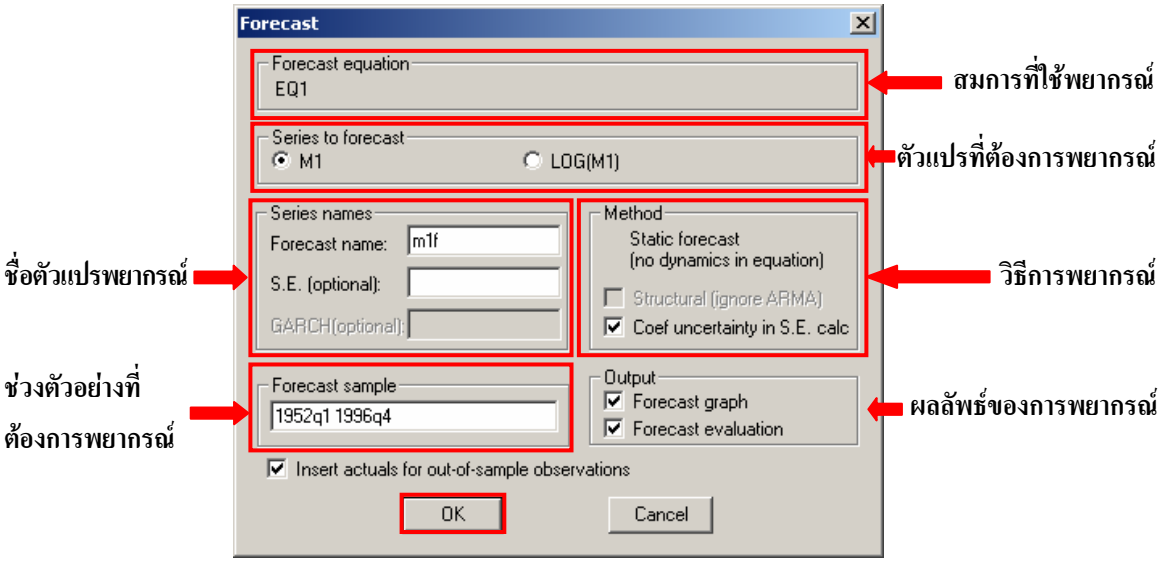


4.6 การสร้างค่าพยากรณ์ และค่าความคลาดเคลื่อน

ถ้าผู้ใช้ต้องการค่าพยากรณ์ และค่าความคลาดเคลื่อนที่ได้มาจากการประมาณค่าสมการถดถอย ก็สามารถทำได้ด้วยโปรแกรม EViews หรือถ้าในการวิเคราะห์บางแบบจำลอง อย่างเช่น Error Correction Model (ตามวิธีการของ Engle and Granger) ผู้ใช้จำเป็นต้องมีค่าความคลาดเคลื่อนจากการประมาณค่าสมการถดถอยมาใช้ในแบบจำลองดังกล่าว ในกรณีการสร้างค่าพยากรณ์ และค่าความคลาดเคลื่อนภายหลังจากการประมาณค่าสมการถดถอยด้วยวิธีการ OLS สามารถทำได้ด้วยโปรแกรม EViews ตามขั้นตอนต่อไปนี้

ก. <u>การสร้างค่าพยากรณ์</u>

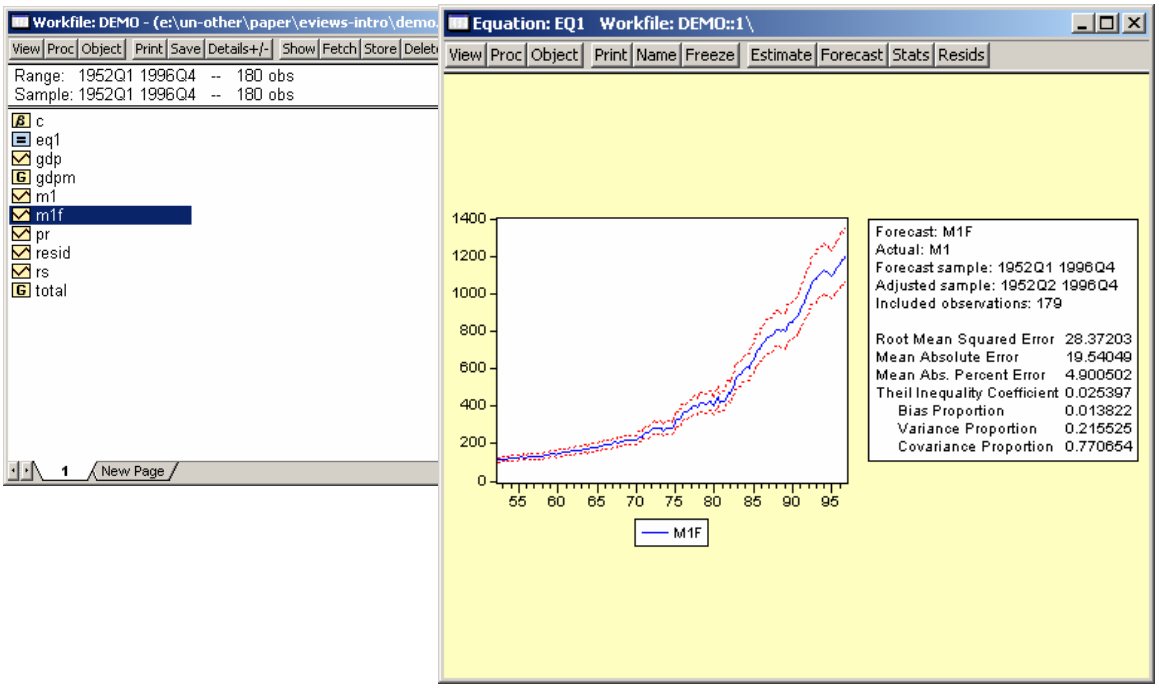
ขั้นตอนที่ 1 สามารถทำได้โดยการเลือก **Proc/Forecast…** หรือ กดปุ่ม Forecast จากแถบเครื่องมือของหน้าต่าง **Equation** จะปรากฏหน้าต่าง **Forecast** ดังนี้



ขั้นตอนที่ 2 กำหนดค่าต่างๆ ตามที่ต้องการ เสร็จแล้วให้กดปุ่ม OK โปรแกรม EViews ก็จะคำนวณค่าพยากรณ์ แล้วบันทึกไว้ใน Series m1f พร้อมทั้งแสดงกราฟของค่าพยากรณ์ นอกจากนี้ถ้าหากผู้ใช้กำหนดให้โปรแกรม EViews เสนอผลการประเมินค่าการพยากรณ์ โปรแกรม EViews ก็จะนำเสนอสถิติของการประเมินค่าการพยากรณ์ ซึ่งประกอบด้วยค่าสถิติ 4 ตัว ดังนี้

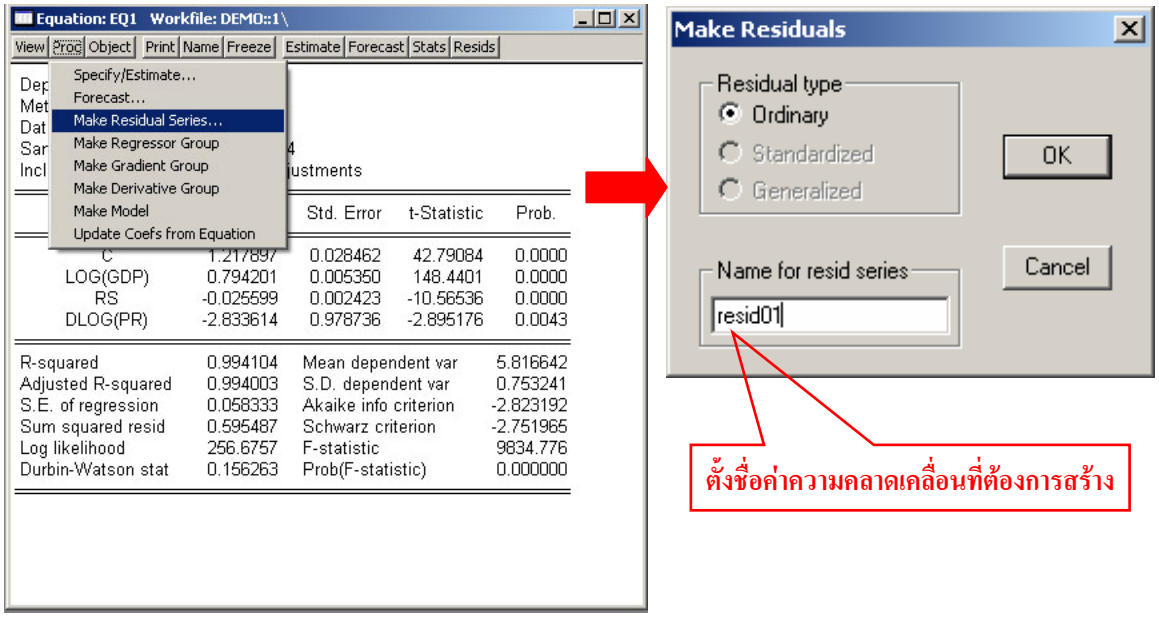
- Root Mean Squared Error $$\sqrt{\sum_{t=T+1}^{T+h}\left(\hat{Y}_t - Y_t\right)^2 / h}$$
- Mean Absolute Error $$\sum_{t=T+1}^{T+h}\left|\hat{Y}_t - Y_t\right| / h$$
- Mean Absolute Percentage Error $$100\sum_{t=T+1}^{T+h}\left|\frac{\hat{Y}_t - Y_t}{Y_t}\right| / h$$
- Theil Inequality Coefficient $$\frac{\sqrt{\sum_{t=T+1}^{T+h}\left(\hat{Y}_t - Y_t\right)^2 / h}}{\sqrt{\sum_{t=T+1}^{T+h}\hat{Y}_t^2 / h} + \sqrt{\sum_{t=T+1}^{T+h}Y_t^2 / h}}$$

ค่าสถิติสองตัวแรกเป็นค่าสถิติที่ใช้ในการพิจารณาความคลาดเคลื่อนของการพยากรณ์ ดังนั้นค่าสถิติทั้งสองตัวควรมีค่ายิ่งน้อยยิ่งดี ส่วนค่าสถิติอีก 2 ตัวที่เหลือ ก็คือ ค่าที่บอกถึงความไม่แปรเปลี่ยน (Invariant) โดยเฉพาะค่า Theil Inequality Coefficient โดยปกติมีค่าระหว่าง 0 – 1 และถ้าค่า Theil Inequality Coefficient มีค่าเข้าใกล้ 0 เท่าไรก็ยิ่งดีเท่านั้น

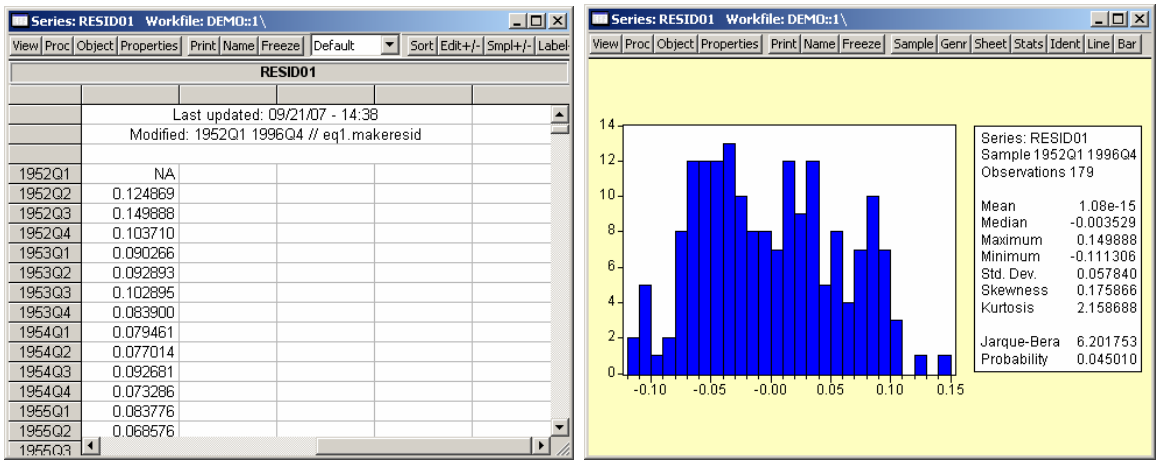


ข. *การสร้างค่าความคลาดเคลื่อน*

ขั้นตอนที่ 1 สามารถทำได้โดยการเลือก **Proc/Make Residual Series…** จากแถบเครื่องมือของหน้าต่าง **Equation** จะปรากฏหน้าต่าง **Make Residuals** ดังนี้



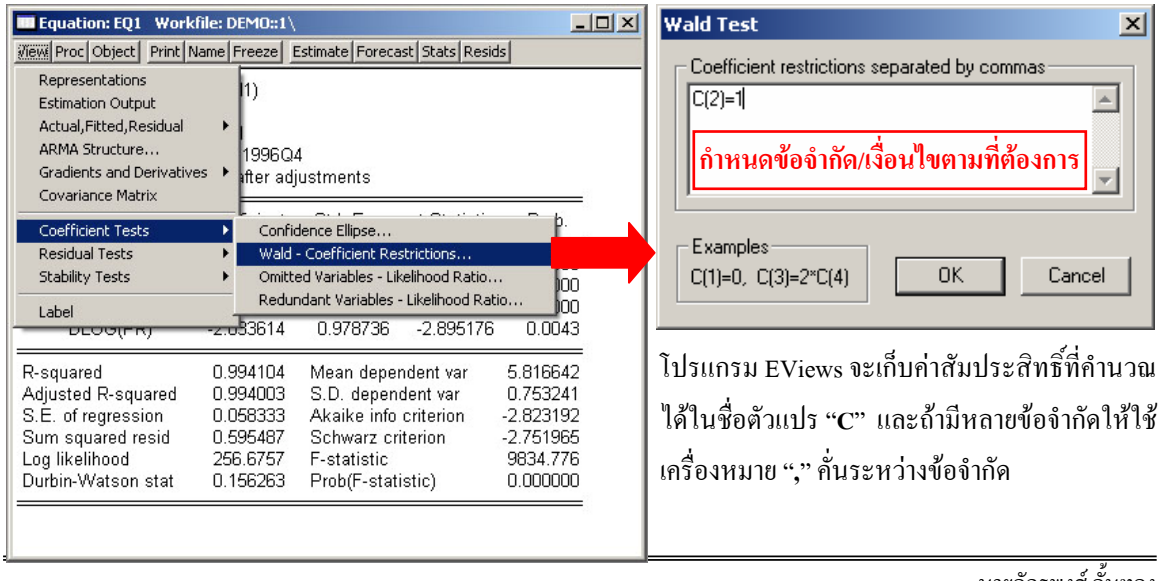
ขั้นตอนที่ 2 ภายหลังจากตั้งชื่อค่าความคลาดเคลื่อนเสร็จแล้ว ให้กดปุ่ม OK โปรแกรม EViews ก็จะสร้าง Series resid01 พร้อมทั้งแสดงตารางของค่าดังกล่าว (ดังรูป) ถ้าผู้ใช้ต้องการดูค่าสถิติพื้นฐานของค่าความคลาดเคลื่อนที่คำนวณมาได้ ให้ผู้ใช้กดปุ่ม Stats ที่หน้าต่าง **RESID01** ก็จะได้กราฟ และค่าสถิติพื้นฐานต่างๆ ดังรูป



4.7 การทดสอบสมมติฐานข้อจำกัดของค่าสัมประสิทธิ์ด้วย Wald Test

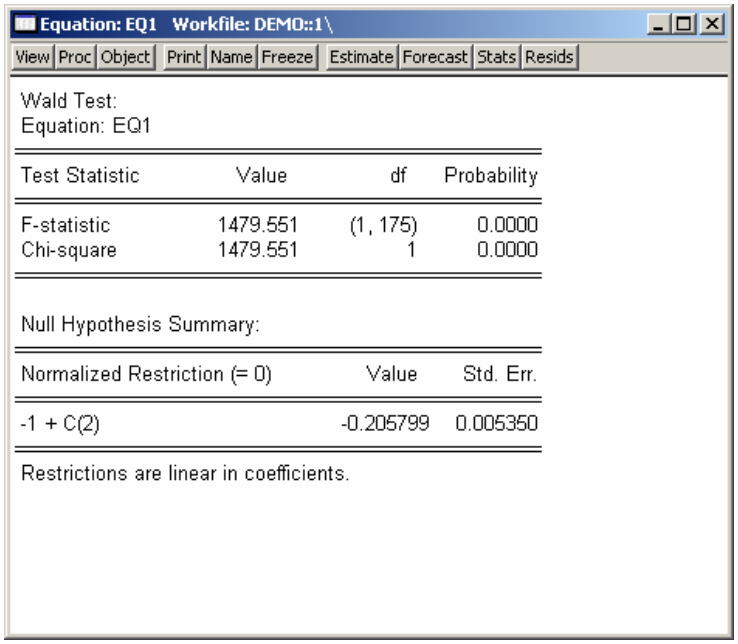
ในบางกรณีผู้ใช้ต้องการทดสอบสมมติฐานข้อจำกัดของค่าสัมประสิทธิ์ ด้วยวิธีการที่เรียกว่า Wald Test ในโปรแกรม EViews มีคำสั่งสำเร็จรูปที่ใช้ในการทดสอบดังกล่าว ตัวอย่างเช่น ในกรณีตัวอย่างถ้าต้องการทดสอบว่าค่าสัมประสิทธิ์ของตัวแปร ln(GDP) [0.79420] มีค่าเท่ากับ 1 หรือไม่ [H_0: $\beta_1 = 1$] สามารถทำการทดสอบสมมติฐานดังกล่าว ตามขั้นตอนดังต่อไปนี้

ขั้นตอนที่ 1 ภายหลังจากการประมาณค่าสัมประสิทธิ์ของตัวแปรต่างๆ แล้ว ผู้ใช้สามารถทำการทดสอบสมมติฐานข้อจำกัดของค่าสัมประสิทธิ์ด้วย Wald Test โดยการเลือก **View/Coefficient Tests /Wald – Coefficient Resctictions…** จากแถบเครื่องมือของหน้าต่าง **Equation** แล้วจะปรากฏหน้าต่าง **Wald Test** ดังนี้



โปรแกรม EViews จะเก็บค่าสัมประสิทธิ์ที่คำนวณได้ในชื่อตัวแปร “**C**” และถ้ามีหลายข้อจำกัดให้ใช้เครื่องหมาย “,” คั่นระหว่างข้อจำกัด

ขั้นตอนที่ 2 หลังจากกำหนดข้อจำกัดตามที่ต้องการแล้ว [ในกรณีตัวอย่าง ต้องกำหนดว่า C(2) =1] ให้กดปุ่ม OK โปรแกรม EViews จะทำการประมวลผลและนำเสนอผลการทดสอบดังรูป



จากผลการทดสอบสมมติฐานหลัก [H_0: $\beta_1 = 1$] ปรากฏว่าค่า F-statistic ที่คำนวณได้มีค่าสูงกว่าค่าวิกฤติ ณ ระดับความเชื่อมั่นที่ 99% [Prob.$< \alpha$] ทำให้ต้องปฏิเสธสมมติฐานหลัก แสดงว่า ค่าสัมประสิทธิ์ของตัวแปร ln(GDP) มีค่าแตกต่างไปจาก 1

กรณีที่มีข้อจำกัดที่ใช้ในการทดสอบเพียง 1 ข้อจำกัด ค่า F-statistic และ Chi-square ที่คำนวณได้จะมีค่าเท่ากัน แต่ถ้ามีข้อจำกัดมากกว่า 1 ข้อจำกัด ค่า F-statistic และ Chi-square ที่คำนวณได้จะมีค่าแตกต่างกัน

เพื่อความเข้าใจเรื่องการทดสอบสมมติฐานข้อจำกัดของค่าสัมประสิทธิ์ด้วย Wald Test ในที่นี้จะขออธิบายวิธีการ Wald Test พอสังเขป ดังนี้ [อ้างอิงมาจากคู่มือ Eviews 5.1]

จากแบบจำลอง Linear regression

$$Y = X\beta + \varepsilon$$

และสมมติฐานหลักของข้อจำกัดเชิงเส้น (Linear restriction)

$$H_0: R\beta - r = 0$$

โดยที่ R คือ เมทริกซ์ q×k

r คือ เวกเตอร์ q

q คือ จำนวนข้อจำกัด

ดังนั้น ค่า Wald statistic สามารถคำนวณได้จาก

$$W = (Rb - r)' \left(Rs^2 \left(X'x\right)^{-1} R'\right)^{-1} (Rb - r)$$

สถิติทดสอบนี้มีการแจกแจงเชิงเส้นกำกับ (Asymptotically distributed) เป็นแบบ χ^2 และเมื่อสมมติให้ค่าความคลาดเคลื่อน (error: ε) เป็นอิสระ และมีการกระจายแบบปกติ สามารถคำนวณค่า F-statistic ได้ดังนี้

$$F = \frac{W}{q} = \frac{\left(\tilde{\varepsilon}'\tilde{\varepsilon} - \varepsilon'\varepsilon\right)/q}{\left(\varepsilon'\varepsilon\right)/(T-k)}$$

โดยที่ $\tilde{\varepsilon}$ คือ ค่าความคลาดเคลื่อนจากการถดถอยแบบมีข้อจำกัด (Restricted regression)

ค่า F-statistic ที่คำนวณได้ จะเป็นการเปรียบเทียบระหว่างค่าผลรวมกำลังสองของค่าคลาดเคลื่อนที่คำนวณได้ในกรณีที่มีข้อจำกัดกับกรณีที่ไม่มีข้อจำกัด

5. Multicollinearity

5.1 ปัญหา Multicollinearity

เป็นปัญหาในกรณีที่ตัวแปรอิสระที่อยู่ในสมการถดถอยมีความสัมพันธ์ระหว่างกันสูง ซึ่งผิดข้อสมมติพื้นฐานของการประมาณค่าสัมประสิทธิ์ด้วยวิธีการ OLS ว่าตัวแปรอิสระแต่ละตัวต้องเป็นอิสระต่อกัน [$\text{Corr}(X_i, X_j) \neq 1$] ถ้าตัวแปรอิสระในแบบจำลองมีความสัมพันธ์กันอย่างสมบูรณ์ (Perfect multicollinearity) [$\text{Corr}(X_i, X_j) = 1$] จะไม่สามารถประมาณค่าสัมประสิทธิ์ของสมการถดถอยได้ โปรแกรม EViews จะแสดงข้อความว่า ***"Near singular matrix"*** และถ้าหากว่าตัวแปรอิสระเป็นอิสระต่อกัน (Orthogonal) [$\text{Corr}(X_i, X_j) = 0$] การประมาณค่าสัมประสิทธิ์ของสมการถดถอยอย่างง่าย (Simple regression) ก็เพียงพอที่จะสามารถใช้อธิบายความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามได้ ไม่จำเป็นที่จะใช้แบบจำลองสมการถดถอยพหุคูณ (Multiple regression)

แต่ในทางปฏิบัติมักพบว่า ตัวแปรอิสระที่นำมาใช้ในการศึกษามีความสัมพันธ์กันอยู่บ้าง ความสัมพันธ์ของตัวแปรอิสระเหล่านี้สามารถวัดได้จากค่าสหสัมพันธ์ (Correlation) ที่มีค่าอยู่ระหว่าง 0 – 1 โดยถ้าหากตัวแปรอิสระมีค่าสหสัมพันธ์กันสูงมาก [ไม่ควรมีค่าสหสัมพันธ์เกิน 0.80] ก็จะทำให้ตัวประมาณค่าสัมประสิทธิ์ของสมการถดถอยมีความแม่นยำและมีเสถียรภาพลดลง ดังนั้นปัญหา Multicollinearity จึงมิใช่เป็นเรื่องเกี่ยวกับการพิจารณาว่าตัวแปรอิสระมีความสัมพันธ์กันหรือไม่ แต่เป็นการพิจารณาในเรื่องของขนาด (Degree) ของความสัมพันธ์ ระหว่างตัวแปรอิสระ ซึ่งถ้าขนาดความสัมพันธ์มีค่าน้อยๆ ก็ยังถือว่าตัวประมาณค่าสัมประสิทธิ์ของสมการถดถอยมีความแม่นยำและมีเสถียรภาพในระดับที่น่าเชื่อถือได้

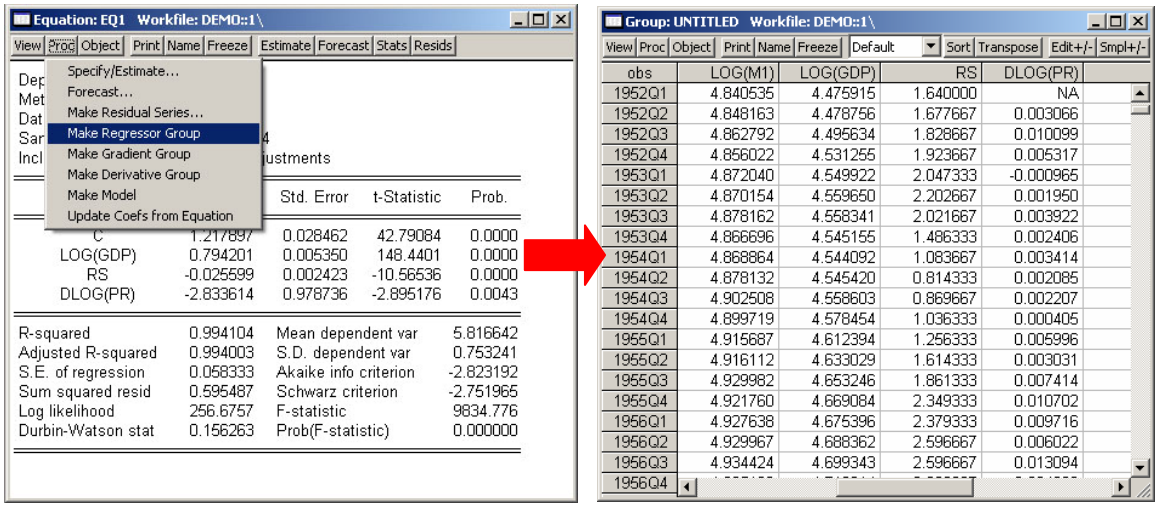
5.2 การตรวจสอบปัญหา Multicollinearity

วิธีการตรวจสอบปัญหา Multicollinearity ที่นิยมใช้ในปัจจุบันมีอยู่ด้วยกัน 2 วิธี คือ การตรวจสอบด้วย Simple Correlation Coefficients และ Variance Inflation Factors (VIF) สามารถทำการตรวจสอบโดยใช้โปรแกรม EViews ได้ดังนี้

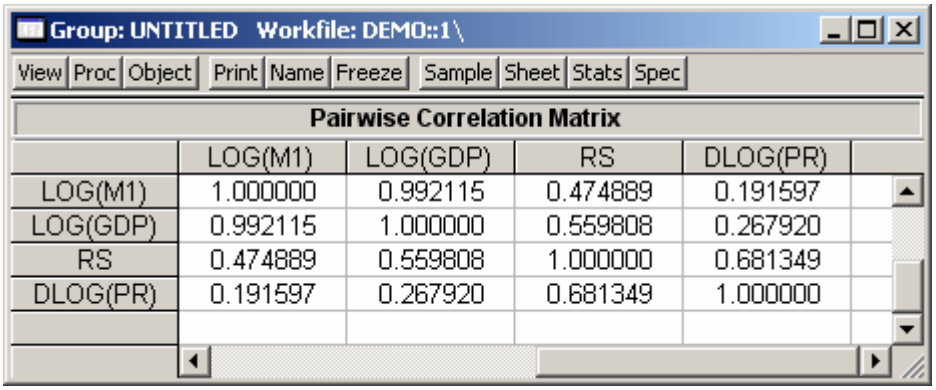
- ***การตรวจสอบด้วย Simple Correlation Coefficients***

จากกรณีตัวอย่างข้างต้น ต้องการตรวจสอบว่าตัวแปรอิสระที่นำมาใช้ในการวิเคราะห์ทั้งหมด 3 ตัว มีค่าสหสัมพันธ์สูง (High Simple Correlation Coefficients) จนก่อให้เกิดปัญหา Multicollinearity หรือไม่ ซึ่งผู้ใช้สามารถทำตามขั้นตอนได้ดังต่อไปนี้

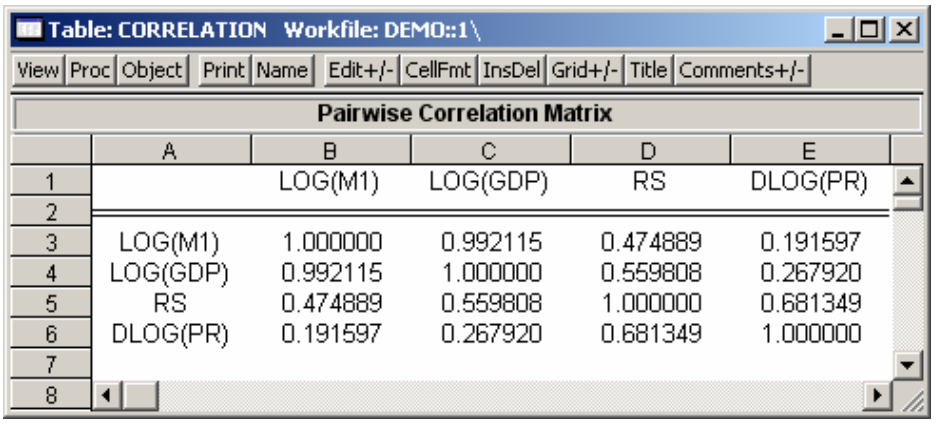
ขั้นตอนที่ 1 จากหน้าต่าง **Equation** (ที่ชื่อว่า EQ1) ให้เลือก **Proc/Make Regressor Group** จากแถบเครื่องมือของหน้าต่าง **Equation** แล้ว โปรแกรม EViews จะสร้างหน้าต่างของ **Regressor Group** ขึ้นมาให้ใหม่ ดังรูป



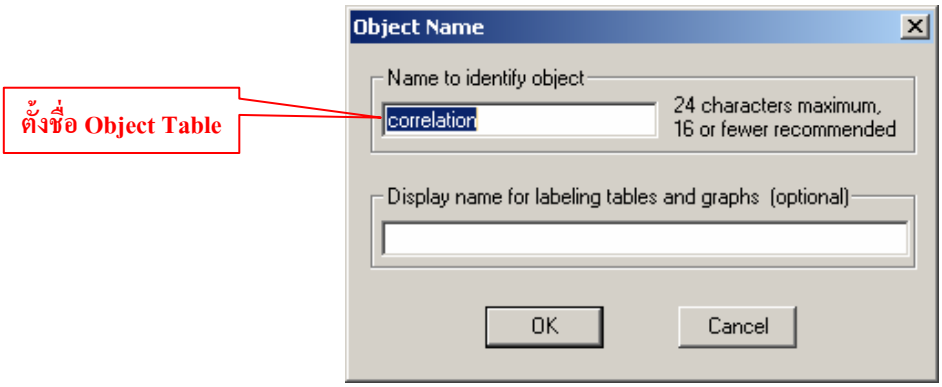
ขั้นตอนที่ 2 จากหน้าต่าง **Regressor Group** ที่สร้างขึ้นมาใหม่ ให้เลือก **View/Correlations/Pairwise Samples** จากแถบเครื่องมือของหน้าต่าง **Regressor Group** จะได้ตาราง Pairwise Correlation Matrix ดังรูป



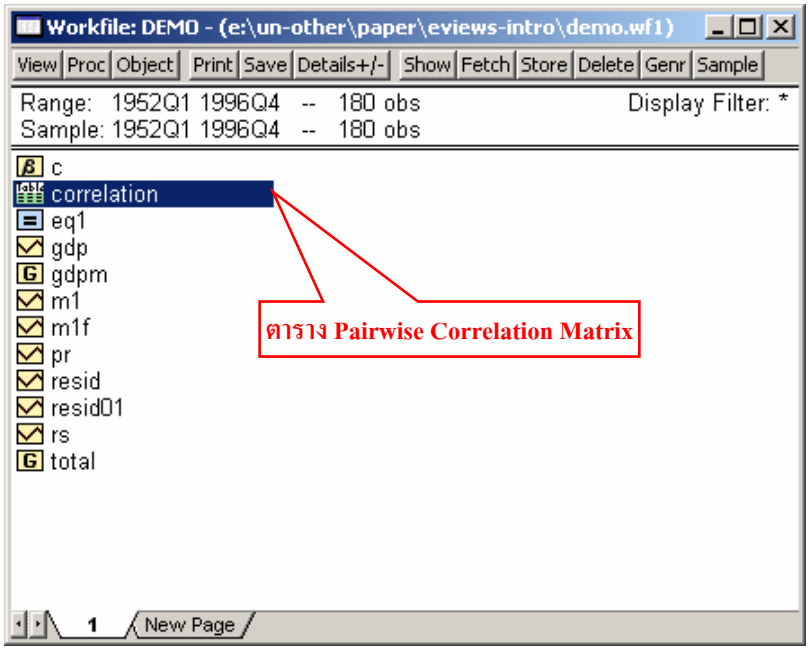
ขั้นตอนที่ 3 เมื่อต้องการเก็บผลการวิเคราะห์เอาไว้ ให้กดปุ่ม Freeze โปรแกรม EViews ก็จะสร้าง Objects Table ของตาราง Pairwise Correlation Matrix ขึ้นมาใหม่ ดังรูป



ขั้นตอนที่ 4 บันทึก Objects Table ของตาราง Pairwise Correlation Matrix โดยการกดปุ่ม Name จะปรากฏหน้าต่าง **Object Name** ดังรูป



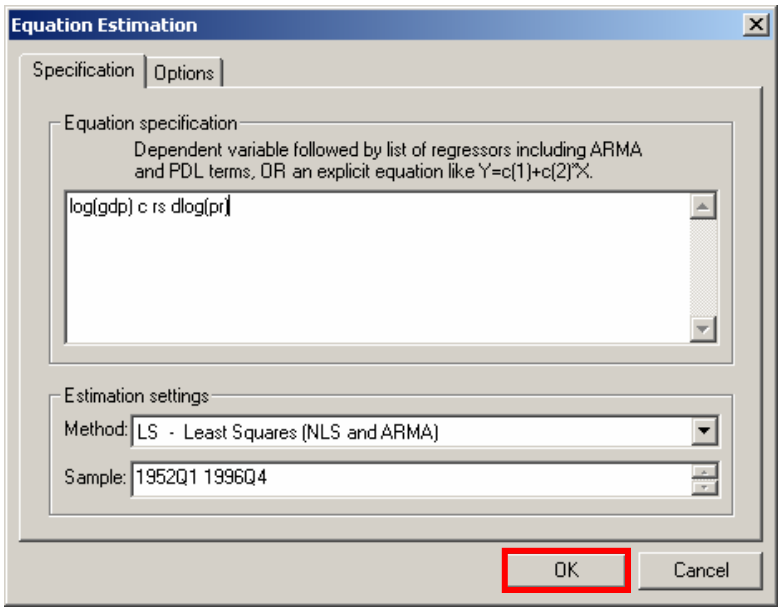
ขั้นตอนที่ 5 เมื่อตั้งชื่อ Objects Table ของตาราง Pairwise Correlation Matrix เสร็จแล้ว ให้กดปุ่ม OK จะปรากฏ Objects Table ของตาราง Pairwise Correlation Matrix ใน Workfile ดังรูป



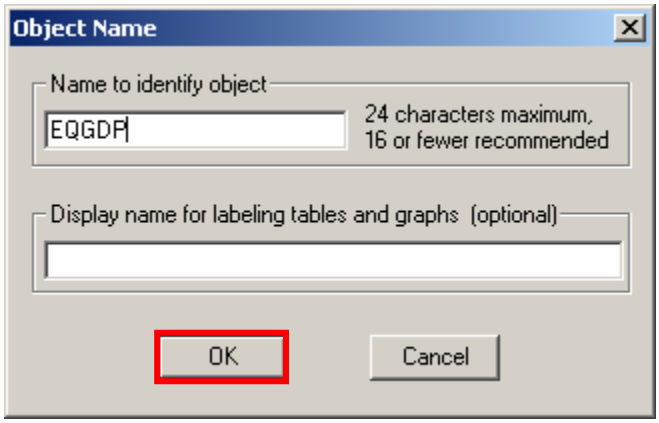
➢ ***การตรวจสอบด้วย Variance Inflation Factors (VIF)***

ค่าสถิติ VIF สามารถใช้วัดขนาดของความรุนแรงของปัญหา Multicollinearity ได้ โดยทั่วไปถ้าค่า VIF ที่คำนวณได้มีค่ามากกว่า 5 ขึ้นไป ก็ถือได้ว่าเกิดปัญหา Multicollinearity ที่รุนแรง (Studenmund 2006: 259) แต่มีนักเศรษฐมิติจำนวนมากใช้เกณฑ์มากกว่า 10 ขึ้นไป (ไพฑูรย์ ไกรพรศักดิ์, 2546) สำหรับวิธีการตรวจสอบด้วยค่า VIF สามารถทำได้ดังนี้

ขั้นตอนที่ 1 เลือก **Objects/New Object/Equation** จากแถบเครื่องมือหลักของหน้าต่าง **Workfile** แล้วพิมพ์คำสั่งว่า ***"LOG(GDP) C RS DLOG(PR)"*** ลงในหน้าต่าง **Equation Specification** เสร็จแล้วให้กดปุ่ม OK ดังรูป



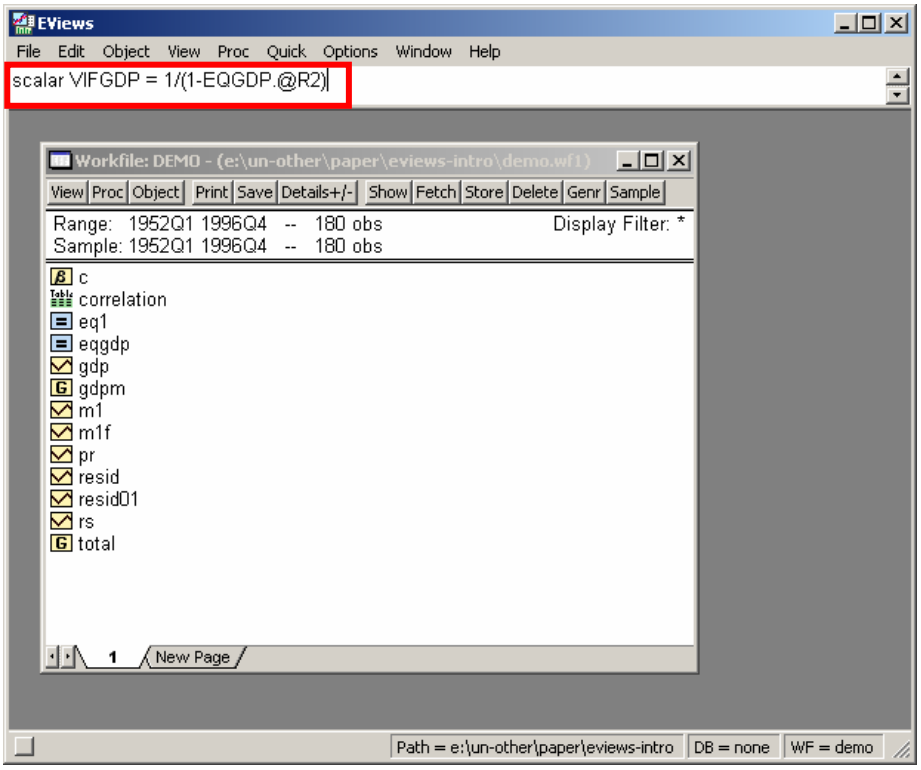
ขั้นตอนที่ 2 เมื่อกดปุ่ม OK โปรแกรม EViews ก็จะสร้าง Object Equation ใหม่ ขึ้นมาให้ ต่อมาให้ผู้ใช้กดปุ่ม Name ที่แถบเครื่องมือของหน้าต่าง **Objects Equation** เพื่อบันทึก Objects Equation ในชื่อ EQGDP เสร็จแล้วให้กดปุ่ม OK ดังรูป



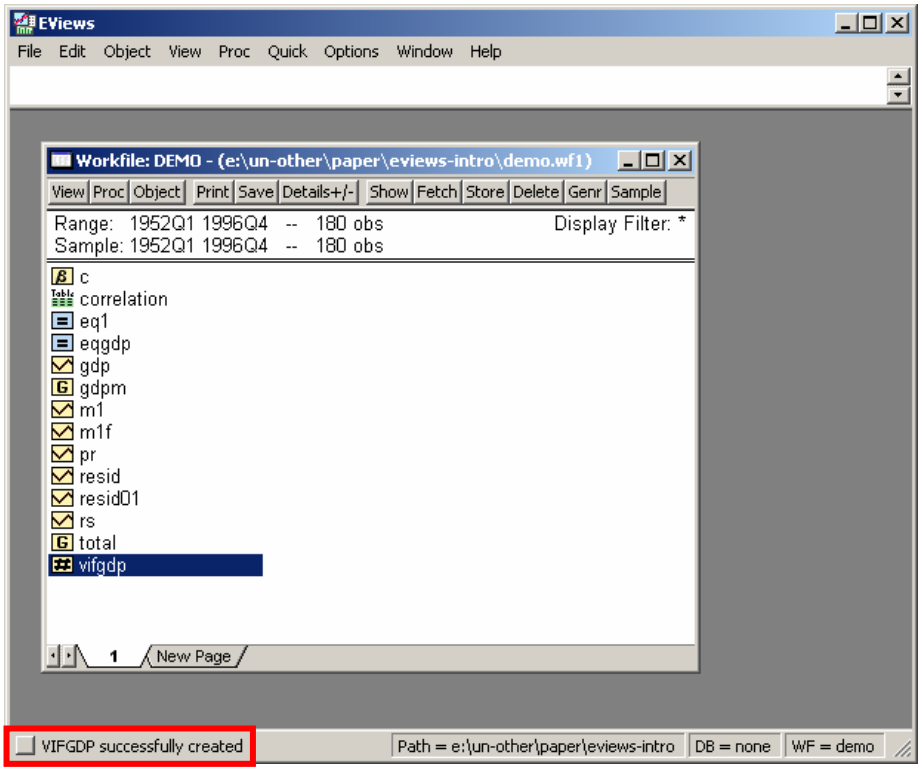
ขั้นตอนที่ 3 เมื่อได้ Objects Equation ที่ชื่อ EQGDP แล้ว และได้ผลลัพธ์จากการประมาณค่าสัมประสิทธิ์ของสมการถดถอยแล้ว ต่อมาจะคำนวณค่าสถิติ VIF โดยมีสูตรคำนวณดังนี้

$$\mathrm{VIF}\left(\hat{\beta}_i\right) = \frac{1}{\left(1-R_i^2\right)}$$

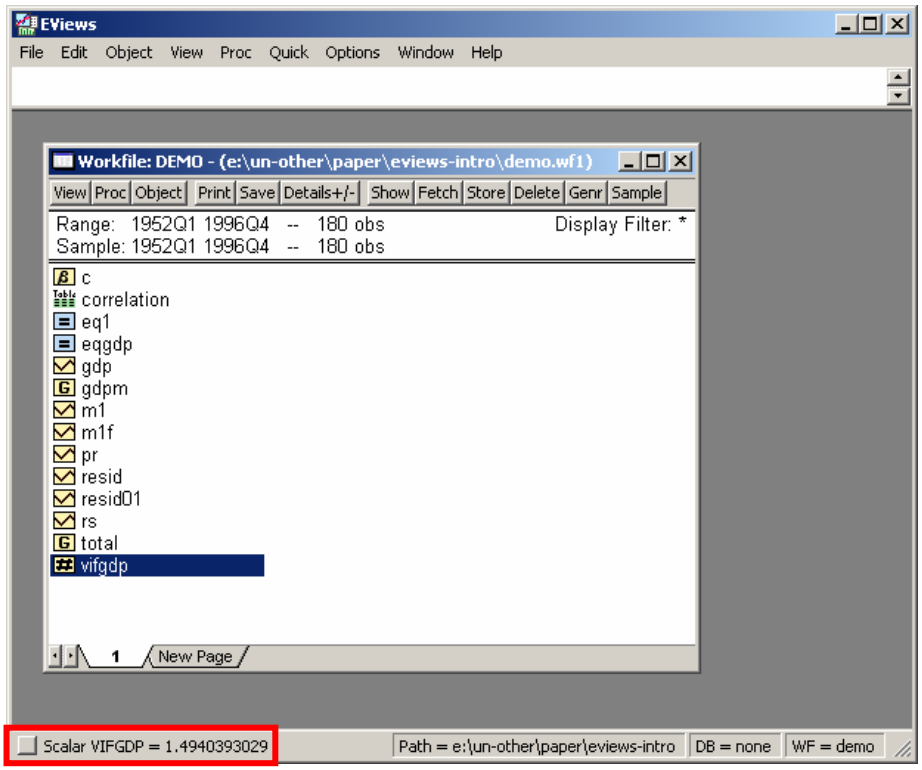
ให้เขียนสูตรดังกล่าวลงในหน้าต่าง **Command** โดยพิมพ์คำสั่งว่า ***"scalar VIFGDP = 1/(1-EQGDP.@R2)"*** แล้วกด **Enter**



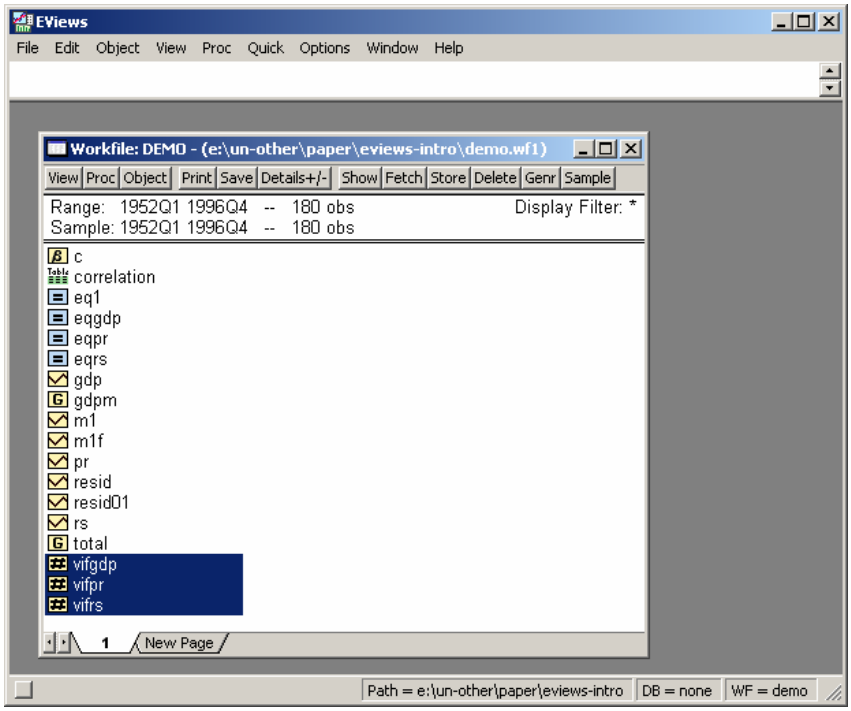
ขั้นตอนที่ 4 เมื่อกด **Enter** โปรแกรม EViews ก็คำนวณค่า VIFGDP เก็บไว้ที่ Objects Scalar และที่ Status Line [อยู่ด้านล่างของหน้าต่างหลัก] จะมีข้อความว่า ***"VIFGDP successfully created"*** ดังรูป



ขั้นตอนที่ 5 ถ้าหากต้องการดูค่า VIF ให้กด Double click ที่ Objects Scalar [**vifgdp**] โปรแกรม EViews ก็จะแสดงค่า VIFGDP ที่ Status Line [อยู่ด้านล่างของหน้าต่างหลัก] ดังรูป



ขั้นตอนที่ 6 ทำขั้นตอนที่ 1 – 5 สำหรับตัวแปรอิสระที่เหลืออีก 2 ตัว [RS และ Δln(PR)] ในที่สุดก็จะมี Objects Scalar ที่แสดงค่า VIF ของตัวแปร lnGDP [**vifgdp**], RS [**vifrs**] และ Δln(PR) [**vifpr**] ดังรูป



สามารถสรุปค่า VIF ของทั้ง 3 ตัวแปรได้ดังนี้

- ตัวแปร ln(GDP) มีค่า VIF = 1.494
- ตัวแปร RS มีค่า VIF = 2.588
- ตัวแปร Δln(PR) มีค่า VIF = 1.930

จากการตรวจสอบปัญหา Multicollinearity ด้วย Simple Correlation Coefficients และ Variance Inflation Factors (VIF) สามารถสรุปได้ว่า สมการถดถอยที่ใช้ในการประมาณค่าสัมประสิทธิ์ในครั้งนี้ไม่มีปัญหา Multicollinearity เนื่องจากค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรอิสระทั้ง 3 ตัว มีค่าต่ำกว่า 0.80 และค่าสถิติ VIF ของทั้ง 3 ตัวแปร ก็มีค่าต่ำกว่า 5

5.3 การแก้ไขกรณีเกิดปัญหา Multicollinearity

ก. ไม่ต้องดำเนินการแก้ไขปัญหา Multicollinearity เนื่องจากความพยายามในการแก้ไขปัญหาดังกล่าวอาจทำให้เกิดผลเสียหลายๆ ด้านต่อตัวประมาณค่าของสมการถดถอย นอกจากนี้ปัญหา Multicollinearity ไม่ได้ก่อให้เกิด Bias และในบางครั้งก็ไม่ได้ทำให้ t-statistic ที่คำนวณได้มีขนาดลดลงจนไม่มีนัยสำคัญทางสถิติ

ข. ตัดตัวแปรที่ก่อให้เกิดปัญหา Multicollinearity โดยเฉพาะตัวแปรที่มีความสัมพันธ์กับตัวแปรอิสระในระดับที่สูง แต่ต้องไม่ใช่ตัวแปรอิสระที่สนใจ หรือมีความสำคัญในแบบจำลอง

ค. เพิ่มขนาดของกลุ่มตัวอย่าง เนื่องจากการเพิ่มขนาดของกลุ่มตัวอย่างจะช่วยลดค่าความคลาดเคลื่อนมาตรฐานของตัวประมาณค่าสัมประสิทธิ์ทำให้ตัวประมาณค่าของสมการถดถอยมีความแม่นยำมากขึ้น

ง. การเปลี่ยนรูป (Transforming) ตัวแปรที่มีปัญหา Multicollinearity วิธีการนี้เหมาะสมในกรณีที่ผู้ใช้ไม่สามารถตัดตัวแปรอิสระตัวใดตัวหนึ่งออกจากแบบจำลอง รูปแบบของการเปลี่ยนรูปตัวแปรโดยใช้โปรแกรม EViews สามารถทำได้หลายวิธี เช่น

ชื่อฟังก์ชัน	รูปแบบของฟังก์ชัน	Eviews specification*
Linear combination	$Y_t = \beta_0 + \beta_1(X_t + Z_t)$	Y C X+Z
First difference	$Y_t = \beta_0 + \beta_1(X_t - X_{t-1})$	Y C d(X)
First difference of the logarithm	$Y_t = \beta_0 + \beta_1(\ln X_t - \ln X_{t-1})$	Y C dlog(X)
One-period % change (in decimal)	$Y_t = \beta_0 + \beta_1[(X_t - X_{t-1})/ X_t]$	Y C pch(X)

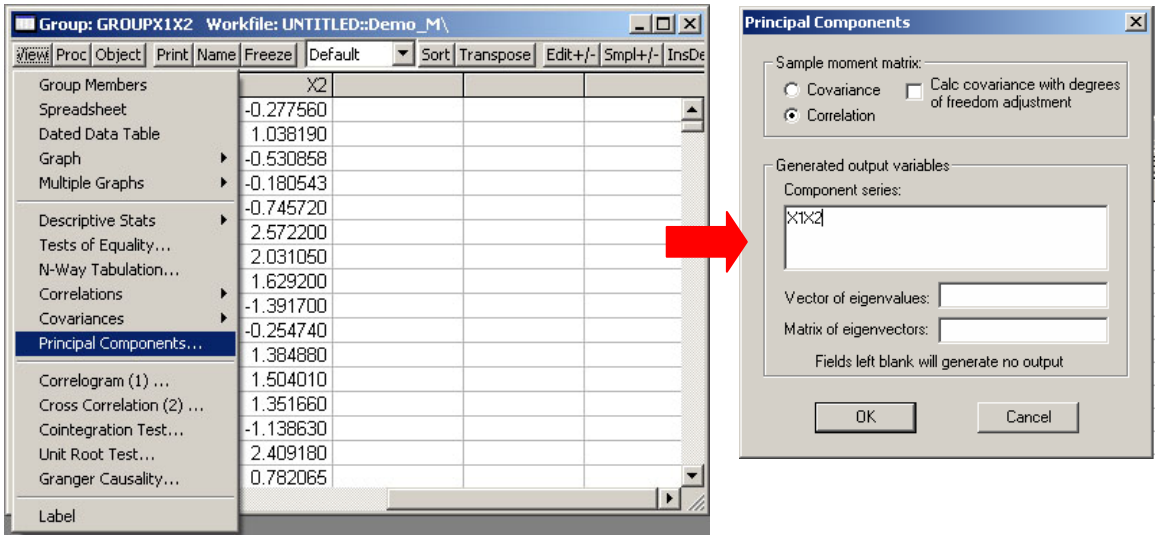
หมายเหตุ: * ผู้ใช้ระบุคำสั่งเหล่านี้ในช่อง Equation specification ที่หน้าต่าง **Equation Estimation**.

ที่มา : R.R. Johnson (2000).

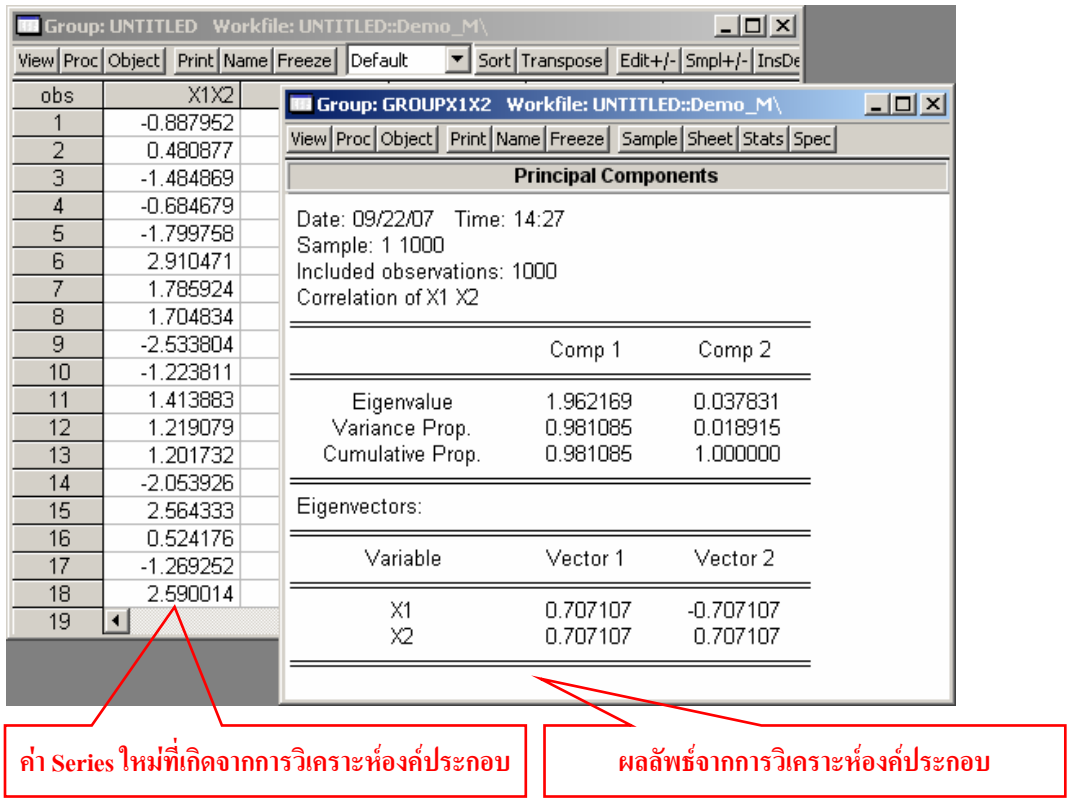
จ. การใช้ Factor Analysis หรือ Principal Components เพื่อสร้างตัวแปรอิสระใหม่จากตัวแปรอิสระที่มีความสัมพันธ์สูง แต่อย่างไรก็ตามผู้ใช้พึงระวังว่า ตัวแปรอิสระที่สร้างขึ้นมาใหม่ต้องสามารถเป็นตัวแปรทางด้านเศรษฐศาสตร์ได้ สำหรับใน Eviews 5.1 มีคำสั่งสำหรับการคำนวณ Principal Components โดยผู้ใช้สามารถดูได้จากหน้าต่าง **Objects Group** ของ Series แล้วไปที่ **View/Principal Components…** [ใน EViews 6.0 ได้เพิ่มคำสั่งสำหรับการคำนวณ Factor Analysis]

ดังนั้นเพื่อความเข้าใจในการใช้ Principal components ในที่นี้ขอยกตัวอย่าง โดยสมมติให้ในการประมาณค่าสมการถดถอยสมการหนึ่ง ซึ่งมีจำนวนตัวอย่าง 1,000 ตัวอย่าง แล้วปรากฏว่า X_1 และ X_2 มีความสัมพันธ์กันจนก่อให้เกิดปัญหา Multicollinearity ดังนั้นในที่นี้ต้องการใช้วิธีการ Principal components ในการ Group รวม X_1 และ X_2 แล้วสร้างเป็นตัวแปรใหม่ชื่อ “**Z**” สามารถใช้โปรแกรม EViews ในการคำนวณดังกล่าวได้ดังนี้

ขั้นตอนที่ 1 สร้าง Objects Group ของตัวแปร X_1 และ X_2 แล้วเลือก **View/Principal Components…** โปรแกรม EViews ก็จะเปิดหน้าต่าง **Principal Components** มาให้ ดังรูป



ขั้นตอนที่ 2 ให้ระบุว่าจะใช้ Matrix อะไรมาคำนวณ ระหว่าง Covariance และ Correlation [โปรแกรม EViews ตั้งค่า Default ที่ Correlation] สำหรับในช่อง Component series ให้ผู้ใช้กำหนดชื่อ Series ที่จะสร้างขึ้นมาใหม่ และถ้าต้องการผลลัพธ์ของ Vector of eigenvalues และ Matrix of eigenvectors ก็ให้ระบุชื่อในช่องทั้งสอง เสร็จแล้วให้กดปุ่ม OK โปรแกรม EViews ก็จะสร้าง Series ใหม่ชื่อ X1X2 [เป็น Series ที่เกิดจากการวิเคราะห์องค์ประกอบของตัวแปร X_1 และ X_2 โดยวิธี Principal components พร้อมทั้งนำเสนอผลการวิเคราะห์องค์ประกอบด้วยวิธี Principal components ดังรูป



จากผลการวิเคราะห์องค์ประกอบด้วยวิธีการ Principal components ได้แสดงให้เห็นว่า ตัวแปร X_1 และ X_2 สามารถนำมาสร้างองค์ประกอบได้ 1 องค์ประกอบ [ดูจากค่า Eigenvalue] องค์ประกอบที่สร้างขึ้นสามารถอธิบาย Variance ได้ 98% [ดูจากค่า Variance Prop.] โดยองค์ประกอบที่สร้างขึ้นมาใหม่มีน้ำหนักของตัวแปร X_1 และ X_2 เท่ากับ 0.707 เหมือนกัน เมื่อได้ Series X1X2 แล้ว ผู้ใช้สามารถนำตัวแปร (Series) ดังกล่าวไปใช้ในการประมาณค่าสมการถดถอย แต่อย่างไรก็ตามผู้ใช้พึงตระหนักอยู่เสมอว่า ตัวแปรใหม่ที่สร้างขึ้นมาต้องสามารถอธิบายในทางเศรษฐศาสตร์ได้

6. Heteroskedasticity

6.1 ปัญหา Heteroskedasticity

เป็นปัญหาที่เกี่ยวข้องกับตัวคลาดเคลื่อน (Error /Residuals: ε) โดยความแปรปรวนของตัวคลาดเคลื่อนที่ได้จากสมการประมาณค่ามีค่าไม่คงที่ [$E(\varepsilon_i^2) \neq \sigma^2$] ซึ่งผิดข้อสมมติพื้นฐานของวิธีการกำลังสองน้อยที่สุด [OLS] ที่ได้มีข้อสมมติพื้นฐานว่า ตัวคลาดเคลื่อนจะต้องมีค่าความแปรปรวนคงที่ [$E(\varepsilon_i^2) = \sigma^2$] การที่ความแปรปรวนของตัวคลาดเคลื่อนไม่คงที่เกิดจากสาเหตุ 2 ประการ คือ ก. เกิดจากการกำหนดรูปแบบหรือโครงสร้างของตัวแบบในสมการถดถอยไม่ถูกต้อง (Impure Heteroskedasticity) เช่น มีการละเลยตัวแปรอิสระบางตัว และ ข. เกิดขึ้นเอง (Pure Heteroskedasticity) โดยรูปแบบหรือโครงสร้างของตัวแบบในสมการถดถอยมีความถูกต้องทุกประการ ปกติแล้วการใช้ข้อมูลภาคตัดขวาง (Cross sectional data) มักจะมีโอกาสที่ค่าความคลาดเคลื่อนจะมีความแปรปรวนไม่คงที่

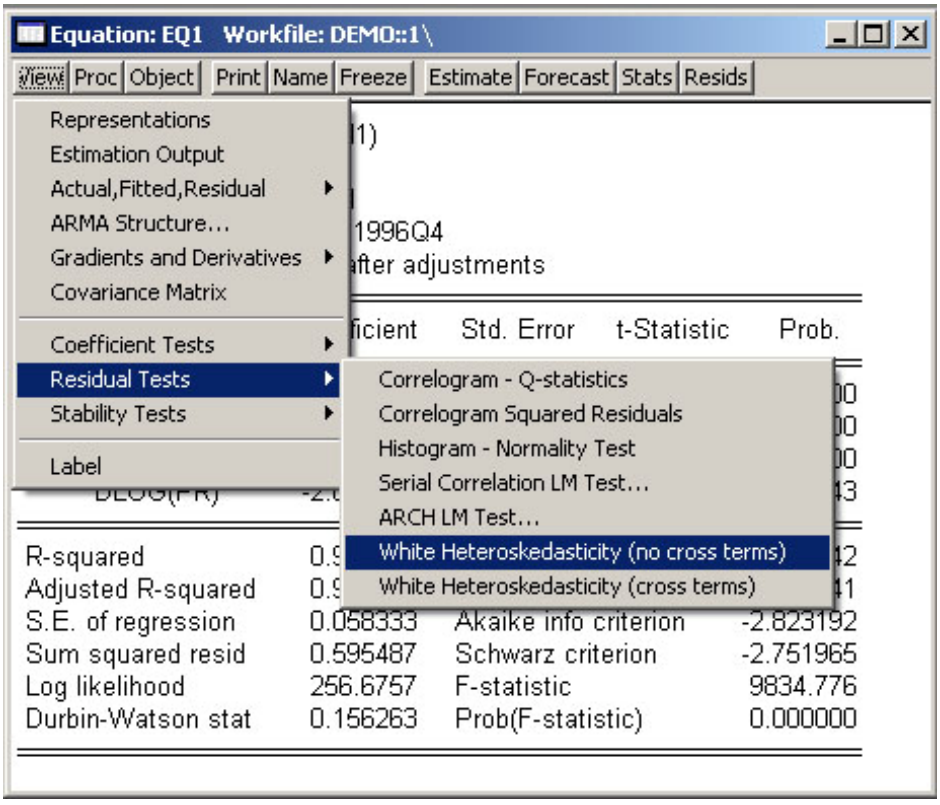
สูงกว่ากรณีที่ใช้ข้อมูลอนุกรมเวลา [Time series data] เนื่องจากค่าสังเกตของข้อมูลภาคตัดขวางจะมีความแตกต่างกันตามขนาดหรือลำดับ ในขณะที่ข้อมูลอนุกรมเวลาจะมีความแตกต่างในเรื่องดังกล่าวเพียงเล็กน้อย

การที่ตัวคลาดเคลื่อนมีความแปรปรวนไม่คงที่ หรือเกิดปัญหา Heteroskedasticity จะทำให้ตัวประมาณค่าสัมประสิทธิ์ของสมการถดถอยยังคงมีคุณสมบัติ Unbiased และ Consistency แต่จะสูญเสียคุณสมบัติ Efficiency นอกจากนี้การใช้วิธีการ OLS ในการประมาณค่าสัมประสิทธิ์ของสมการถดถอยเมื่อมีปัญหา Heteroskedasticity ก็จะทำให้ค่าความคลาดเคลื่อนมาตรฐานของตัวประมาณค่าสัมประสิทธิ์ของสมการถดถอยมีค่าแตกต่างไปจากความเป็นจริง ส่งผลให้ค่า t-statistic ที่คำนวณได้ ของค่าสัมประสิทธิ์แต่ละตัวไม่น่าเชื่อถือ ทำให้การทดสอบสมมติฐานของค่าสัมประสิทธิ์ในสมการถดถอยขาดความน่าเชื่อถือไปด้วย

6.2 การตรวจสอบปัญหา Heteroskedasticity

โปรแกรม Eviews 5.1 สามารถตรวจสอบปัญหา Heteroskedasticity ได้ โดยใน EViews 5.1 สามารถตรวจสอบปัญหาดังกล่าวด้วยวิธีการ White's Heteroskedasitcity test แต่ถ้าเป็น EViews 6.0 จะมีวิธีการตรวจสอบปัญหา Heteroskedasticity ให้เลือกหลายวิธี เช่น Breusch – Pagan – Godfrey, Harvey, Glejser เป็นต้น สำหรับในเอกสารฉบับนี้ได้ใช้โปรแกรม Eviews 5.1 เป็นตัวอย่างในการสาธิต ดังนั้นจึงจะขอสาธิตการตรวจสอบปัญหา Heteroskedasticity ด้วยวิธีการ White's Heteroskedasticity test เพียงวิธีการเดียว ดังมีรายละเอียดดังนี้

จากกรณีตัวอย่างเดิม ให้ผู้ใช้เปิด Objects Equation ที่บันทึกไว้ขึ้นมา [Double click ที่ Objects Equation [eq1] แล้วเลือก **View/Residual Tests/White Heteroskedasticity (no cross terms) หรือ (cross terms)** [ขึ้นอยู่กับผู้ใช้และขนาดของกลุ่มตัวอย่าง] ที่แถบเครื่องมือของหน้าต่าง **Objects Equation** ดังรูป



หลังจากนั้นจะปรากฏหน้าต่างดังนี้

Equation: EQ1 Workfile: DEMO::1\

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

White Heteroskedasticity Test:

F-statistic	6.700178	Prob. F(6,172)	0.000002
Obs*R-squared	33.91119	Prob. Chi-Square(6)	0.000007

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 09/22/07 Time: 17:51
Sample: 1952Q2 1996Q4
Included observations: 179

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.017909	0.020819	0.860185	0.3909
LOG(GDP)	-0.003569	0.007278	-0.490338	0.6245
(LOG(GDP))^2	0.000287	0.000582	0.493370	0.6224
RS	-0.001947	0.000497	-3.916420	0.0001
RS^2	0.000110	2.82E-05	3.885070	0.0001
DLOG(PR)	0.392949	0.160727	2.444831	0.0155
(DLOG(PR))^2	-8.764811	5.517560	-1.588530	0.1140

R-squared	0.189448	Mean dependent var	0.003327
Adjusted R-squared	0.161173	S.D. dependent var	0.003591
S.E. of regression	0.003289	Akaike info criterion	-8.558185
Sum squared resid	0.001861	Schwarz criterion	-8.433538
Log likelihood	772.9575	F-statistic	6.700178
Durbin-Watson stat	0.515464	Prob(F-statistic)	0.000002

White Heteroskedasticity (no cross terms)

Equation: EQ1 Workfile: DEMO::1\

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

White Heteroskedasticity Test:

F-statistic	4.856650	Prob. F(9,169)	0.000009
Obs*R-squared	36.78279	Prob. Chi-Square(9)	0.000029

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 09/22/07 Time: 17:52
Sample: 1952Q2 1996Q4
Included observations: 179

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.001512	0.024550	-0.061581	0.9510
LOG(GDP)	0.004432	0.008837	0.501530	0.6167
(LOG(GDP))^2	-0.000469	0.000735	-0.638033	0.5243
(LOG(GDP))*RS	0.000463	0.000277	1.667973	0.0972
(LOG(GDP))*(DLOG(PR))	-0.079382	0.123765	-0.641395	0.5221
RS	-0.004765	0.001752	-2.719960	0.0072
RS^2	8.38E-05	5.27E-05	1.590174	0.1137
RS*(DLOG(PR))	0.011808	0.042215	0.279710	0.7800
DLOG(PR)	0.773604	0.617379	1.253045	0.2119
(DLOG(PR))^2	-7.693573	8.655845	-0.888830	0.3754

R-squared	0.205490	Mean dependent var	0.003327
Adjusted R-squared	0.163179	S.D. dependent var	0.003591
S.E. of regression	0.003285	Akaike info criterion	-8.544656
Sum squared resid	0.001824	Schwarz criterion	-8.366589
Log likelihood	774.7467	F-statistic	4.856650
Durbin-Watson stat	0.529066	Prob(F-statistic)	0.000009

White Heteroskedasticity (cross terms)

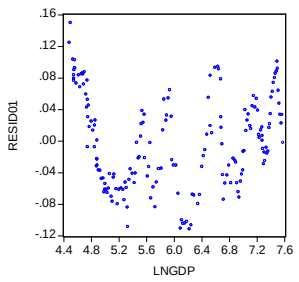
จากสมมติฐานที่ว่า

H_0: Homoscedasticity

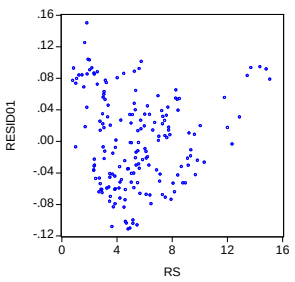
H_1: Heteroskedasticity

ผลการทดสอบ พบว่า ค่าสถิติ nR^2 ที่คำนวณได้มีค่าสูงกว่าค่าวิกฤต ณ ระดับความเชื่อมั่นที่ 99% [Prob. $< \alpha$] จึงปฏิเสธสมมติฐานหลัก แสดงว่า สมการถดถอยนี้มีปัญหา Heteroskedasticity และเมื่อพิจารณาในระดับรายละเอียดก็จะเห็นได้อย่างชัดเจนว่า ตัวแปร RS เป็นตัวแปรที่มีความสัมพันธ์กับค่า $Residual^2$ สูง [Prob. $< \alpha$] ซึ่งหมายความว่าความไม่คงที่ของค่าความแปรปรวนของค่าคลาดเคลื่อนในสมการถดถอยนี้อาจได้รับอิทธิพลมาจากตัวแปร RS

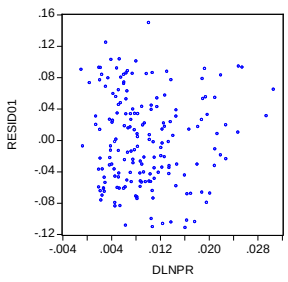
นอกจากการตรวจสอบด้วยวิธีการ White's Heteroskedasticity test แล้ว ในบางครั้งผู้ใช้สามารถตรวจสอบปัญหา Heteroskedasticity ในเบื้องต้น ด้วยการพิจารณากราฟการกระจายระหว่างค่าคลาดเคลื่อนกับตัวแปรอิสระ แต่ละตัว ถ้าแบบแผนการกระจายของค่าคลาดเคลื่อนขึ้นอยู่กับตัวแปรอิสระตัวแปรใดตัวแปรหนึ่ง แสดงว่า สมการถดถอยนี้มีโอกาสที่จะมีปัญหา Heteroskedasticity ดังรูปข้างล่าง



Scatter plot between Resid & lnGDP



Scatter plot between Resid & RS



Scatter plot between Resid & ΔlnPR

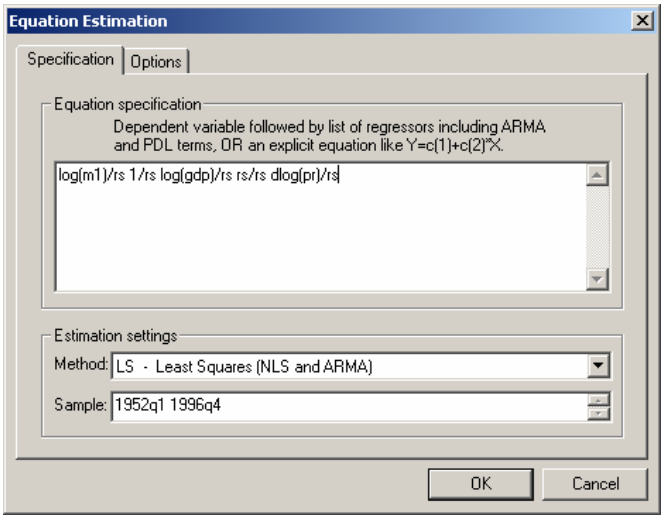
6.3 การแก้ไขกรณีเกิดปัญหา Heteroskedasticity

การแก้ไขในกรณีที่เกิดปัญหา Heteroskedasticity สามารถแก้ไขได้หลายวิธี สำหรับในโปรแกรม EViews มีวิธีการแก้ปัญหา Heterskedasticity ที่สำคัญอยู่ 2 วิธีการ คือ วิธีการ Weighted Least Square (WLS) และวิธีการ Heteroskedasticity-Corrected Standard Errors ดังมีรายละเอียดของขั้นตอนในการแก้ไขปัญหา Heteroskedasitcity ในแต่ละวิธีดังนี้

➢ ***Weighted Least Square (WLS)***

จากตัวอย่างที่ได้ทำการทดสอบปัญหา Heteroskedasticity แล้วพบว่า ตัวแปร RS เป็นตัวแปรที่มีความสัมพันธ์กับค่า Residual^2 สูง ดังนั้นในการแก้ปัญหาด้วยวิธีการ WLS จะใช้ตัวแปร RS เป็นตัว Weight ซึ่งในโปรแกรม EViews สามารถทำได้ 2 วิธีการดังนี้

<u>วิธีการที่ 1</u> เปิด Objects Equation ที่บันทึกไว้ขึ้นมา [Double click ที่ Objects Equation [eq1] แล้วกดปุ่ม Estimate หรือ เลือก **Proc/Specify/Estimate** ที่แถบเครื่องมือของหน้าต่าง **Object Equation** แล้วพิมพ์คำสั่งว่า "**log(m1)/rs 1/rs log(gdp)/rs rs/rs dlog(pr)/rs**" ในช่อง Equation specification ดังรูป



ภายหลังพิมพ์เสร็จแล้วให้กดปุ่ม OK ก็จะได้ผลลัพธ์ดังรูป

Equation: EQ1 Workfile: DEMO::1\

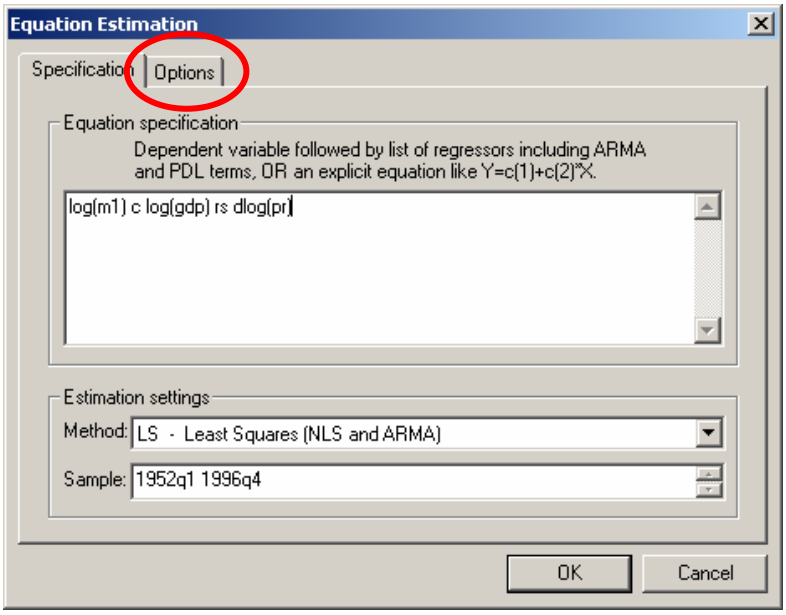
View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: LOG(M1)/RS
Method: Least Squares
Date: 09/22/07 Time: 20:48
Sample (adjusted): 1952Q2 1996Q4
Included observations: 179 after adjustments

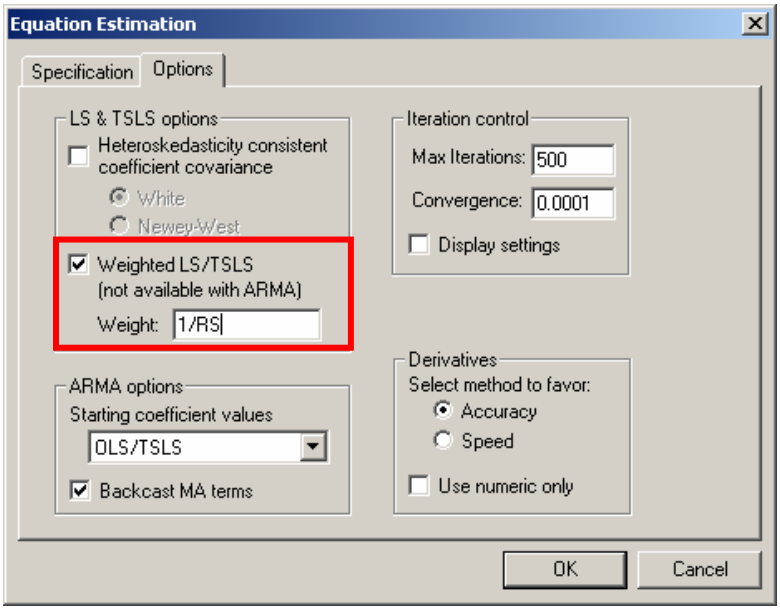
Variable	Coefficient	Std. Error	t-Statistic	Prob.
1/RS	1.293908	0.028216	45.85690	0.0000
LOG(GDP)/RS	0.793808	0.006421	123.6360	0.0000
RS/RS	-0.045287	0.003712	-12.20087	0.0000
DLOG(PR)/RS	-0.734142	1.246536	-0.588946	0.5567

R-squared	0.999641	Mean dependent var	1.420697
Adjusted R-squared	0.999634	S.D. dependent var	0.886402
S.E. of regression	0.016948	Akaike info criterion	-5.295186
Sum squared resid	0.050269	Schwarz criterion	-5.223959
Log likelihood	477.9191	Durbin-Watson stat	0.246860

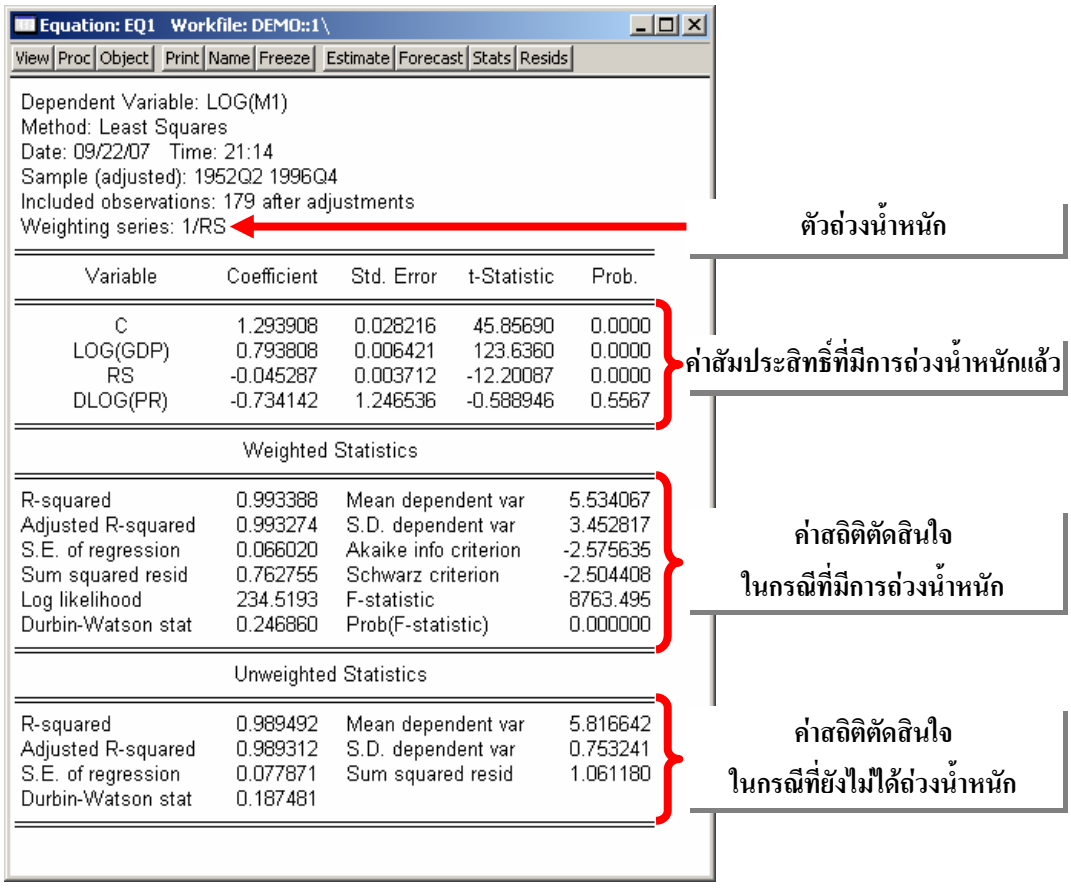
<u>วิธีการที่ 2</u> **ขั้นตอนที่ 1** เปิด Objects Equation ที่บันทึกไว้ขึ้นมา [Double click ที่ Objects Equation [eq1] แล้วกดปุ่ม Estimate หรือ เลือก **Proc/Specify/Estimate** ที่แถบเครื่องมือของหน้าต่าง **Object Equation** แล้วพิมพ์คำสั่งว่า "**log(m1) c log(gdp) rs dlog(pr)**" ในช่อง Equation specification ดังรูป



ขั้นตอนที่ 2 เลือก Options ที่หน้าต่าง **Equation Estimation** [ดังรูปข้างบน] จะปรากฏหน้าต่าง **Equation Options** ขึ้นมา ดังรูป แล้วให้คลิกที่ช่อง Weighted LS/TSLS เสร็จแล้ว ให้ใส่ตัวถ่วงน้ำหนัก (ในที่นี้ก็คือ "**1/RS**") ในช่อง Weight ดังรูป



ขั้นตอนที่ 3 กดปุ่ม OK โปรแกรม EViews ก็จะประมาณค่าสัมประสิทธิ์ของสมการถดถอยด้วยวิธีการ WLS โดยมี "**1/RS**" เป็นตัวถ่วงน้ำหนัก เสร็จแล้วก็จะนำเสนอผลลัพธ์ดังรูป

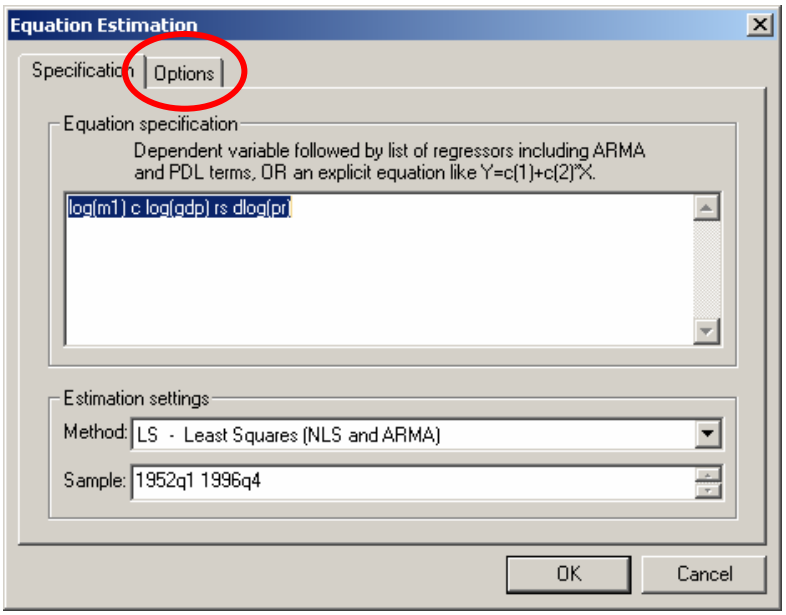


จากข้างต้นจะเห็นได้ว่า วิธีการที่ 1 และวิธีการที่ 2 จะให้ค่าสัมประสิทธิ์ของสมการถดถอยเหมือนกัน แต่มีค่าสถิติการตัดสินใจที่แตกต่างกัน เนื่องจากวิธีการที่ 1 เป็นวิธีการถ่วงน้ำหนักแบบทางอ้อม (Indirect Weighted Least Squares) จะใช้ในกรณีที่ไม่สามารถใช้วิธีการ WLS ได้ วิธีการนี้จะสร้างตัวแบบของสมการถดถอยใหม่ โดยการ Transform ตัวแปรทุกตัวที่ใช้ในการประมาณค่า [ในตัวอย่างก็คือ การคูณตัวแปรทุกตัวด้วย "**1/RS**"] แล้วจึงประมาณค่าสัมประสิทธิ์ของสมการถดถอยด้วยวิธีการ OLS ส่วนวิธีการที่ 2 เป็นการประมาณค่าสัมประสิทธิ์ด้วยวิธีการ WLS โดยถ่วงน้ำหนักด้วย "**1/RS**"

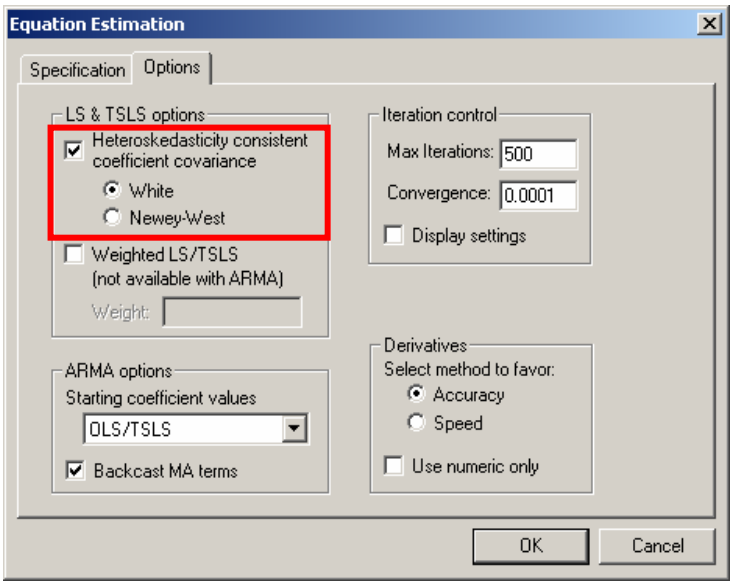
➢ ***Heteroskedasticity-Corrected Standard Errors***

วิธีการนี้เป็นการบรรเทาปัญหา Heteroskedasticity โดยโปรแกรม EViews มี Options ให้เลือก 2 วิธี คือ วิธีการของ White [บรรเทาปัญหา Heteroskedasticity] และ Newey-West [บรรเทาปัญหา Heteroskedasticity และ Autocorrelation] วิธีการทั้งสองแตกต่างกันตรงวิธีประมาณค่า The Heteroskedasticity Consistent Covariance Matrix ดังนั้นผลลัพธ์ที่ได้จึงแตกต่างกันตรงค่าความคลาดเคลื่อนมาตรฐานของตัวประมาณค่าสัมประสิทธิ์ส่วนค่าสัมประสิทธิ์ที่ได้จะมีค่าเหมือนกัน [เหมือนกับกรณีที่ยังไม่บรรเทาปัญหา Heteroskedasticity] ความแตกต่างดังกล่าวทำให้ค่า t-statistic ที่ได้หลังจากการบรรเทาปัญหา Heteroskedasticity ด้วยวิธีการทั้งสองมีความแตกต่างกัน และมีความน่าเชื่อถือกว่ากรณีที่ยังไม่มีการบรรเทาปัญหา Heteroskedasticity

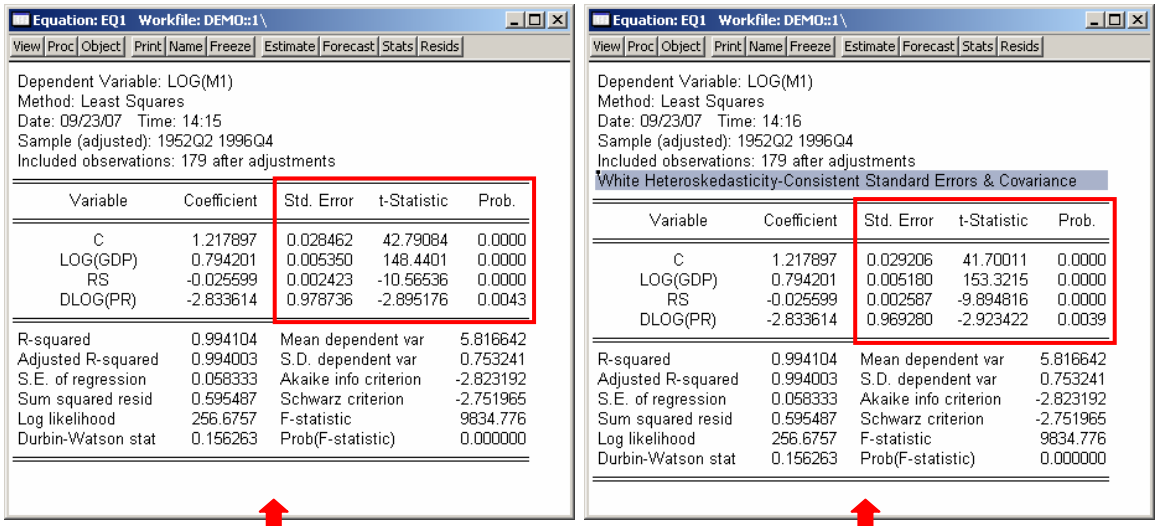
ขั้นตอนที่ 1 เปิด Objects Equation ขึ้นมา [Double click ที่ Objects Equation eq1] แล้วกดปุ่ม Estimate หรือ เลือก **Proc/Specify/Estimate** ที่แถบเครื่องมือของหน้าต่าง **Object Equation** โปรแกรม EViews จะเปิดหน้าต่าง **Equation Estimation** ขึ้นมาดังรูป [ถ้าช่อง Equation specification ไม่มีคำสั่งที่ใช้ในการประมาณค่าสัมประสิทธิ์ ให้ผู้ใช้พิมพ์คำสั่งที่ต้องการลงไปในช่องดังกล่าว แต่ถ้าเป็นการปรับปรุง/แก้ไขการประมาณค่าสัมประสิทธิ์ ในช่อง Equation specification จะมีคำสั่งเก่า ผู้ใช้สามารถแก้ไขได้ตามความต้องการ]



ขั้นตอนที่ 2 เลือก Options ที่หน้าต่าง **Equation Estimation** [ดังรูปข้างบน] จะปรากฏหน้าต่าง **Equation Options** ขึ้นมา ดังรูป แล้วให้คลิกที่ช่อง Heteroskedasticity consistent coefficient covariance เสร็จแล้ว ผู้ใช้จะสามารถเลือก Option ของวิธีการดังกล่าวได้ระหว่างวิธีการของ White กับ Newey-West ดังรูป



ขั้นตอนที่ 3 กดปุ่ม OK โปรแกรม EViews ก็จะประมาณค่าสัมประสิทธิ์ของสมการถดถอย พร้อมทั้งบรรเทาปัญหา Heteroskedasticity ด้วยวิธีการ Heteroskedasticity-Corrected Standard Errors [ในตัวอย่างเลือก White] เสร็จแล้วก็จะนำเสนอผลลัพธ์ดังรูป



Equation: EQ1 Workfile: DEMO::1\

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

Dependent Variable: LOG(M1)
Method: Least Squares
Date: 09/23/07 Time: 14:15
Sample (adjusted): 1952Q2 1996Q4
Included observations: 179 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.217897	0.028462	42.79084	0.0000
LOG(GDP)	0.794201	0.005350	148.4401	0.0000
RS	-0.025599	0.002423	-10.56536	0.0000
DLOG(PR)	-2.833614	0.978736	-2.895176	0.0043

R-squared	0.994104	Mean dependent var	5.816642
Adjusted R-squared	0.994003	S.D. dependent var	0.753241
S.E. of regression	0.058333	Akaike info criterion	-2.823192
Sum squared resid	0.595487	Schwarz criterion	-2.751965
Log likelihood	256.6757	F-statistic	9834.776
Durbin-Watson stat	0.156263	Prob(F-statistic)	0.000000

Equation: EQ1 Workfile: DEMO::1\

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

Dependent Variable: LOG(M1)
Method: Least Squares
Date: 09/23/07 Time: 14:16
Sample (adjusted): 1952Q2 1996Q4
Included observations: 179 after adjustments
White Heteroskedasticity-Consistent Standard Errors & Covariance

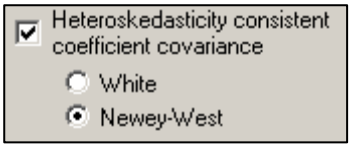
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.217897	0.029206	41.70011	0.0000
LOG(GDP)	0.794201	0.005180	153.3215	0.0000
RS	-0.025599	0.002587	-9.894816	0.0000
DLOG(PR)	-2.833614	0.969280	-2.923422	0.0039

R-squared	0.994104	Mean dependent var	5.816642
Adjusted R-squared	0.994003	S.D. dependent var	0.753241
S.E. of regression	0.058333	Akaike info criterion	-2.823192
Sum squared resid	0.595487	Schwarz criterion	-2.751965
Log likelihood	256.6757	F-statistic	9834.776
Durbin-Watson stat	0.156263	Prob(F-statistic)	0.000000

ผลลัพธ์ก่อนการบรรเทาปัญหา Heteroskedasticity **ผลลัพธ์หลังการบรรเทาปัญหา Heteroskedasticity**

จากผลลัพธ์ข้างต้น จะเห็นได้ว่า การบรรเทาปัญหา Heetroskedasticity ด้วยวิธีการ Heteroskedasticity-Corrected Standard Errors ค่าสัมประสิทธิ์จากการประมาณค่าไม่มีการเปลี่ยนแปลง แต่ค่าความคลาดเคลื่อนมาตรฐานของตัวประมาณค่าสัมประสิทธิ์ [ในช่อง Std. Error] มีความแตกต่างกัน ทำให้ค่า t-statistic มีความแตกต่างกัน [เนื่องจากค่า t-statistic ที่คำนวณมาจากการเอาค่า Coefficient หาร ด้วย Std. Error] การเปลี่ยนแปลงดังกล่าวย่อมมีผลต่อการตัดสินใจเลือกตัวแปรอิสระมาใช้ในการอธิบายความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตาม กล่าวคือ สมมติว่าสมการถดถอยมีปัญหา Heteroskedasticity ทำให้ตัวแปรอิสระบางตัวมีนัยสำคัญทางสถิติทั้งที่ไม่น่าจะมีนัยสำคัญทางสถิติ เมื่อผู้ใช้ได้บรรเทาปัญหาดังกล่าว ปรากฏว่า ตัวแปรอิสระดังกล่าวกลับไม่มีนัยสำคัญทางสถิติจริง ดังนั้นการปัญหา Heteroskedasticity อาจทำให้ผู้ใช้ตัดสินใจเลือกตัวแปรอิสระผิดก็ได้ รวมทั้งการนำมาซึ่งบทสรุปที่ผิดพลาดได้อีกด้วย

สำหรับการบรรเทาปัญหา Heteroskedasticity ด้วยวิธีการ Heteroskedasticity-Corrected Standard Errors ในกรณีของ Newey-West ผู้ใช้ก็สามารถทำได้ โดยการเปลี่ยน Option จาก White เป็น Newey-West ในช่องของ Heteroskedasticity consistent coefficient covariance ดังรูป



7. Autocorrelation

7.1 ปัญหา Autocorrelation

เป็นปัญหาหนึ่งที่เกี่ยวข้องกับตัวคลาดเคลื่อน (Error /Residuals: ε) เช่นเดียวกับปัญหา Hetroskedasticity โดยปัญหา Autocorrelation เกิดขึ้นเนื่องจากการที่ตัวคลาดเคลื่อนมีสหสัมพันธ์ระหว่างกัน หรือตัวคลาดเคลื่อนมีการกระจายที่ไม่เป็นอิสระแก่กัน [$\text{Cov}(\varepsilon_i, \varepsilon_j) = \text{E}(\varepsilon_i, \varepsilon_j) \neq 0$ สำหรับทุกค่าที่ $i \neq j$] ซึ่งผิดข้อสมมติพื้นฐานของวิธีการกำลังสองน้อยที่สุด (OLS) ที่มีข้อสมมติพื้นฐานว่า ตัวคลาดเคลื่อนต้องไม่มีสหสัมพันธ์ระหว่างกัน [$\text{Cov}(\varepsilon_i, \varepsilon_j) = \text{E}(\varepsilon_i, \varepsilon_j) = 0$ สำหรับทุกค่าที่ $i \neq j$] ความสัมพันธ์ระหว่างตัวคลาดเคลื่อนอาจเป็นความสัมพันธ์ในทิศทางบวก [เรียกว่า Positive Autocorrelation] หรือทิศทางลบ [เรียกว่า Negative Autocorrelation] ก็ได้ และตัวคลาดเคลื่อนอาจมีความสัมพันธ์ในช่วงเวลาที่แตกต่างกันได้อีกด้วย โดยทั่วไปการเกิดสหสัมพันธ์ของตัวคลาดเคลื่อนมักจะเกิดขึ้นกับข้อมูลอนุกรมเวลา [เรียกว่า Serial Correlation] แต่ในบางกรณีก็สามารถเกิดกับข้อมูลภาคตัดขวางได้ [เรียกว่า Spatial Correlation] การที่ตัวคลาดเคลื่อนมีสหสัมพันธ์ระหว่างกันมีอยู่ด้วยกัน 2 แบบ คือ ก. ***Impure Autocorrelation*** เป็นกรณีที่ความสัมพันธ์ของตัวคลาดเคลื่อนเป็นผลมาความผิดพลาดของรูปแบบหรือโครงสร้างของตัวแบบ (Specification Error) เช่น การละเลยตัวแปรที่สำคัญ, การใช้แบบจำลอง Cobweb, การใช้ Lagged variables เป็นต้น และ ข. ***Pure Autocorrelation*** เกิดจากตัวคลาดเคลื่อนมีความสัมพันธ์กันเองตามธรรมชาติ

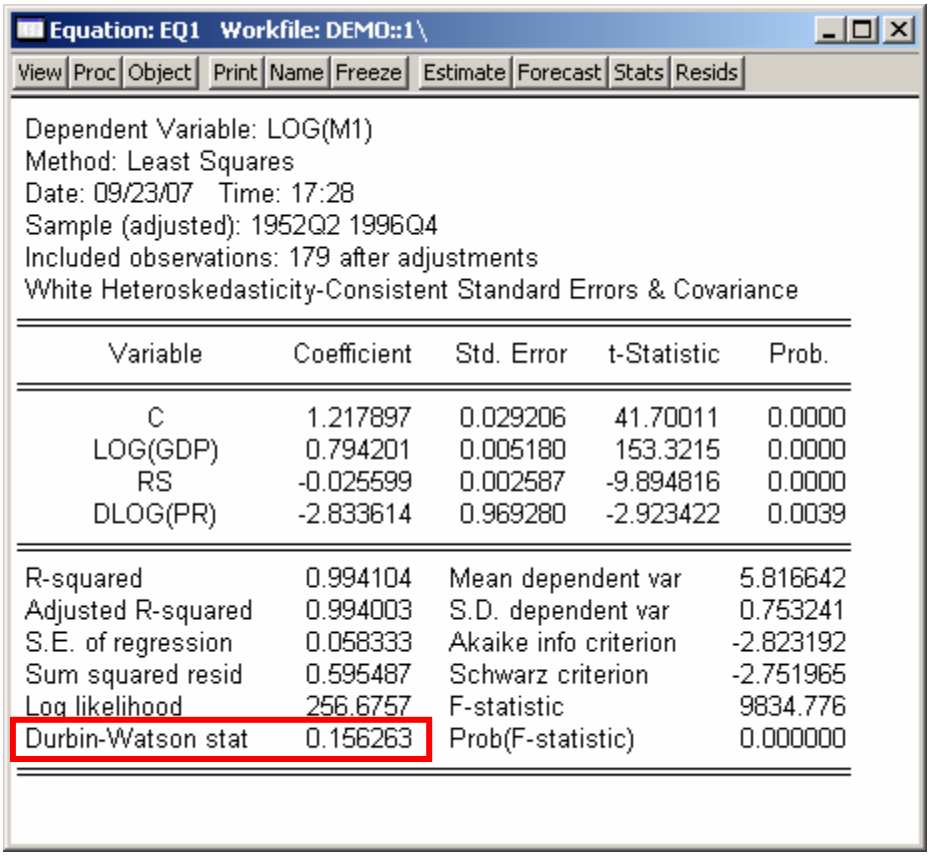
การที่ตัวคลาดเคลื่อนมีความสหสัมพันธ์ระหว่างกัน หรือเกิดปัญหา Autocorrelation จะทำให้ตัวประมาณค่าสัมประสิทธิ์ของสมการถดถอยยังคงมีคุณสมบัติ Unbiased แต่จะสูญเสียคุณสมบัติ Efficiency (ค่าความแปรปรวนของสัมประสิทธิ์จะไม่มีค่าต่ำที่สุด) ทำให้การใช้ OLS ในการประมาณค่าสัมประสิทธิ์ของสมการถดถอยขาดคุณสมบัติ BLUE ย่อมส่งผลทำให้ค่าพยากรณ์ที่เกิดจากสมการถดถอยที่มีปัญหา Autocorrelation มีค่าคลาดเคลื่อนในการพยากรณ์สูงกว่ากรณีที่ไม่มีปัญหา Autocorrelation นอกจากนี้ค่าความแปรปรวนของตัวคลาดเคลื่อนของตัวประมาณค่าสัมประสิทธิ์อาจมีค่าต่ำกว่า (Underestimate) ที่ควรจะเป็น จึงทำให้ค่า t-statistic ที่คำนวณได้สูงกว่าความเป็นจริง และนำมาสู่ข้อสรุปที่ผิดพลาดได้

7.2 การตรวจสอบปัญหา Autocorrelation

การตรวจสอบปัญหา Autocorrelation สามารถทำได้หลายวิธี เช่น การดูจากค่าสถิติ Durbin-Watson (D.W.), การพิจารณาจากกราฟของค่าคลาดเคลื่อน เป็นต้น ในโปรแกรม EViews 5.1 ก็สามารถตรวจสอบปัญหา Autocorrelation ได้หลายวิธีเช่นเดียวกัน แต่วิธีการที่ง่ายและสะดวก ได้แก่ การพิจารณาค่า Durbin-Watson statistic (D.W.), การพิจารณาจากกราฟของค่าคลาดเคลื่อน, การพิจารณาจาก Correlogram และ Q-statistics ของค่าคลาดเคลื่อน และค่า Breusch-Godfrey Test ดังนั้นในที่นี้จึงขอสาธิตวิธีการตรวจสอบปัญหา Autocorrelation ทั้ง 4 วิธี ดังมีรายละเอียดดังนี้

- ***การตรวจสอบจากค่า Durbin-Watson statistic (D.W.)***

ในการประมาณค่าสัมประสิทธิ์ของสมการถดถอย โปรแกรม EViews จะรายงานผลลัพธ์ออกมา 3 ส่วน [กล่าวไว้แล้วในคู่มือฉบับนี้ หน้าที่ 26 - 28] ซึ่งในส่วนที่ 3 ที่เป็นเรื่องของค่าสถิติการตัดสินใจนั้น โปรแกรม EViews จะรายงานผลการคำนวณค่า Durbin-Watson statistic มาให้ดังรูป



Equation: EQ1 Workfile: DEMO::1\

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

Dependent Variable: LOG(M1)
Method: Least Squares
Date: 09/23/07 Time: 17:28
Sample (adjusted): 1952Q2 1996Q4
Included observations: 179 after adjustments
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.217897	0.029206	41.70011	0.0000
LOG(GDP)	0.794201	0.005180	153.3215	0.0000
RS	-0.025599	0.002587	-9.894816	0.0000
DLOG(PR)	-2.833614	0.969280	-2.923422	0.0039

R-squared	0.994104	Mean dependent var	5.816642
Adjusted R-squared	0.994003	S.D. dependent var	0.753241
S.E. of regression	0.058333	Akaike info criterion	-2.823192
Sum squared resid	0.595487	Schwarz criterion	-2.751965
Log likelihood	256.6757	F-statistic	9834.776
Durbin-Watson stat	0.156263	Prob(F-statistic)	0.000000

วิธีการ Durbin-Watson Test เป็นวิธีที่ง่าย เหมาะสมกับกรณีที่ข้อมูลมีขนาดตัวอย่างเล็ก และวิธีการนี้เหมาะที่จะใช้ในการตรวจสอบว่า สมการถดถอยที่กำลังพิจารณามีสหสัมพันธ์ของตัวคลาดเคลื่อนในอันดับที่หนึ่ง (First Order Autocorrelation) หรือไม่ สมมติฐานที่ใช้ในการทดสอบ

H_0: $\rho = 0$ (Non – Autocorrelation)

H_1: $\rho \neq 0$ (Autocorrelation)

จากสมมติฐานดังกล่าว สถิติที่ใช้ในการทดสอบ คือ Durbin Watson statistic สามารถคำนวณได้จากสูตร

$$D.W. = \frac{\sum_{t=2}^{T}(\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{T}\varepsilon_t^2}$$

โดยที่ ε คือ ค่าคลาดเคลื่อนของสมการถดถอยที่ประมาณด้วยวิธี OLS

นอกจากสูตรนี้แล้ว ยังสามารถคำนวณได้จากสูตร

$$D.W. = 2(1-\hat{\rho})$$

โดยที่ ρ คือ ค่าสัมประสิทธิ์ของตัวคลาดเคลื่อนในสมการถดถอยที่แสดงความสัมพันธ์ระหว่างตัวคลาดเคลื่อนข้ามเวลาอันดับหนึ่ง [$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$] หรือ $\rho = \frac{\sum_{t=2}^{T} \varepsilon_t \varepsilon_{t-1}}{\sum_{t=1}^{T} \varepsilon_t^2}$

จากสูตรข้างต้น ค่า D.W. (Durbin Watson statistic) จะมีค่าระหว่าง 0 – 4 [0 เป็นกรณี Positive Autocorrelation สำหรับ 4 เป็นกรณี Negative Autocorrelation] โดยที่

ถ้าหากค่า ρ = -1 ค่า D.W. = 4 แสดงว่า มีปัญหา Perfect Negative Autocorrelation

ถ้าหากค่า ρ = 0 ค่า D.W. = 2 แสดงว่า ไม่มีปัญหา Autocorrelation

ถ้าหากค่า ρ = 1 ค่า D.W. = 0 แสดงว่า มีปัญหา Perfect Positive Autocorrelation

ดังนั้นโดยทั่วไปมักจะพิจารณาว่า ถ้าหากค่า D.W. มีค่าใกล้เคียง 2 ก็แสดงว่า สมการถดถอยที่กำลังพิจารณาไม่มีปัญหา Autocorrelation แต่อย่างไรก็ตาม ในบางครั้งก็ไม่สามารถสรุปได้อย่างนั้น ผู้ใช้จะต้องนำค่า D.W. ที่คำนวณได้ไปเทียบกับค่าวิกฤติในตารางสถิติ Durbin-Watson Test โดยมีเงื่อนไขในการพิจารณาค่า Durbin Watson Statistic มีดังนี้

ถ้าค่า $d_L > D.W. > 4 - d_L$ จะปฏิเสธสมมติฐานหลัก (H_0) แสดงว่ามีปัญหา Autocorrelation

ถ้าค่า $4 - d_U > D.W. > d_U$ จะไม่สามารถปฏิเสธสมมติฐานหลัก (H_0) แสดงว่าไม่มีปัญหา Autocorrelation

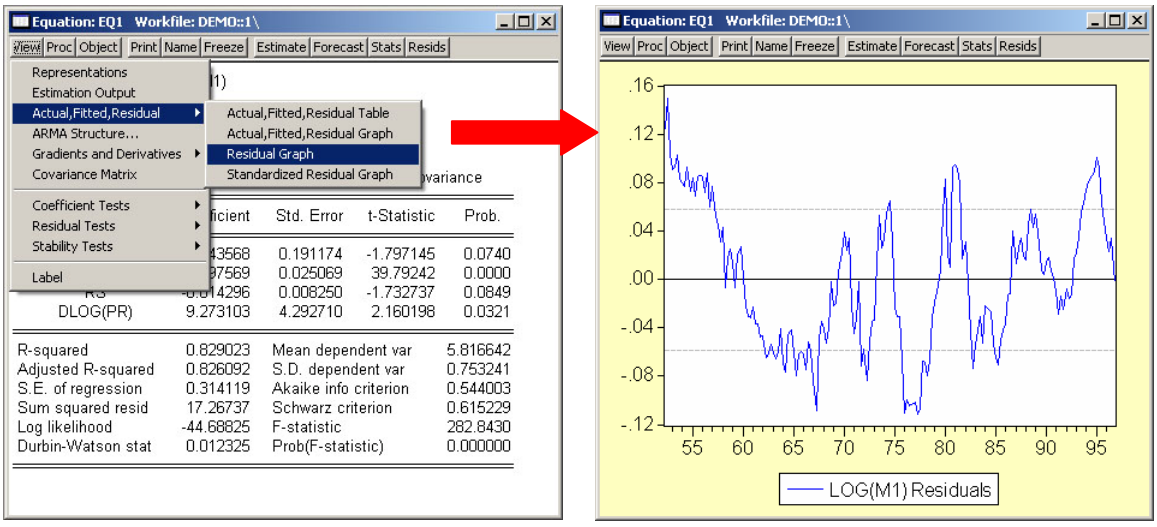
สำหรับกรณีอื่นๆ ไม่สามารถสรุปได้ว่ามีปัญหา Autocorrelation หรือไม่

ที่มา: ดัดแปลงมาจาก ไพฑูรย์ ไกรพรศักดิ์ (2546).

จากกรณีตัวอย่างข้างต้น ค่า D.W. ที่คำนวณได้มีค่า 0.156 ซึ่งเมื่อเปิดตารางสถิติ Durbin-Watson ที่ n=200 [n คือ จำนวนตัวอย่าง ในกรณีตัวอย่างมี n = 180 แต่ในตาราง Durbin-Watson ไม่มี n = 180 จึงเลือกที่ n มากกว่าแทน] และ k = 3 [k คือ จำนวนตัวแปรอิสระ ซึ่งในกรณีตัวอย่างมี k = 3 ตัว] พบว่า d_L = 1.738 ส่วน d_U = 1.799 ในขณะที่ค่า D.W. ที่คำนวณได้มีค่า 0.156 < d_L [1.738] แสดงว่า ปฏิเสธสมมติฐานหลักที่ว่า "H_0: ไม่มี Positive Autocorrelation" หมายความว่า สมการถดถอยนี้มีปัญหา Autocorrelation

➢ *การตรวจจากกราฟของค่าความคลาดเคลื่อน*

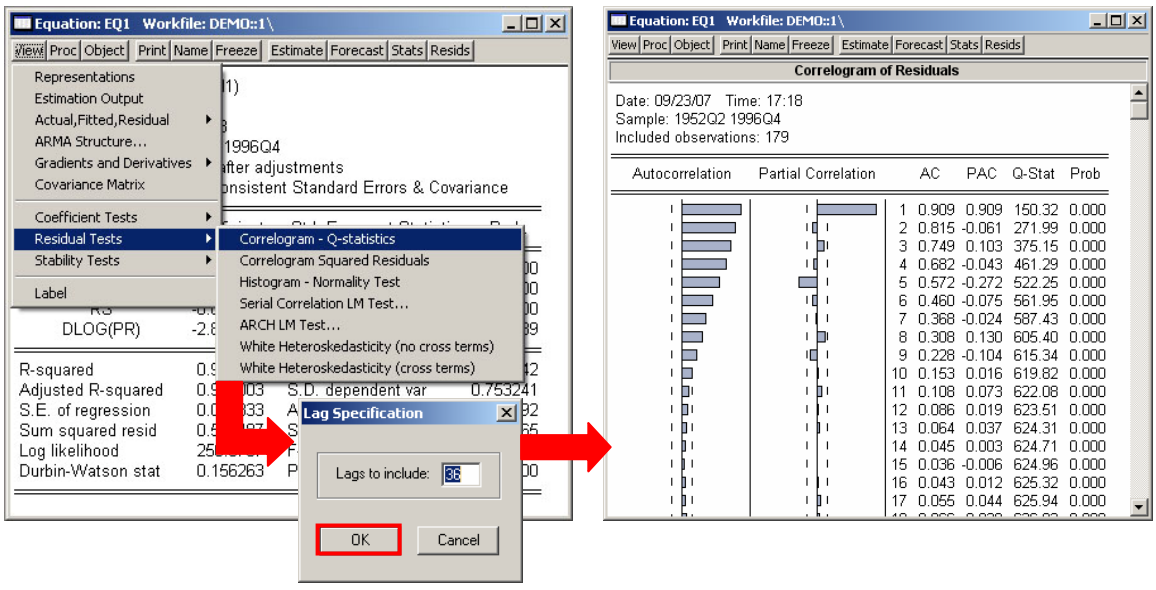
หลังจากผู้ใช้ประมาณค่าสัมประสิทธิ์ของสมการถดถอยเสร็จแล้ว [หน้าต่าง **Object Equation** จะถูกเปิดอยู่] ให้ผู้ใช้เลือก **View/Actual,Fitted,Residual/Residual Graph** โปรแกรม EViews จะนำเสนอกราฟของค่าคลาดเคลื่อนที่ได้จากสมการถดถอย ดังรูป



จากกราฟแสดงให้เห็นว่า ค่าคลาดเคลื่อนมีโอกาสที่จะมีสหสัมพันธ์เชิงบวก

➢ *การตรวจสอบจาก Correlogram – Q-statistics*

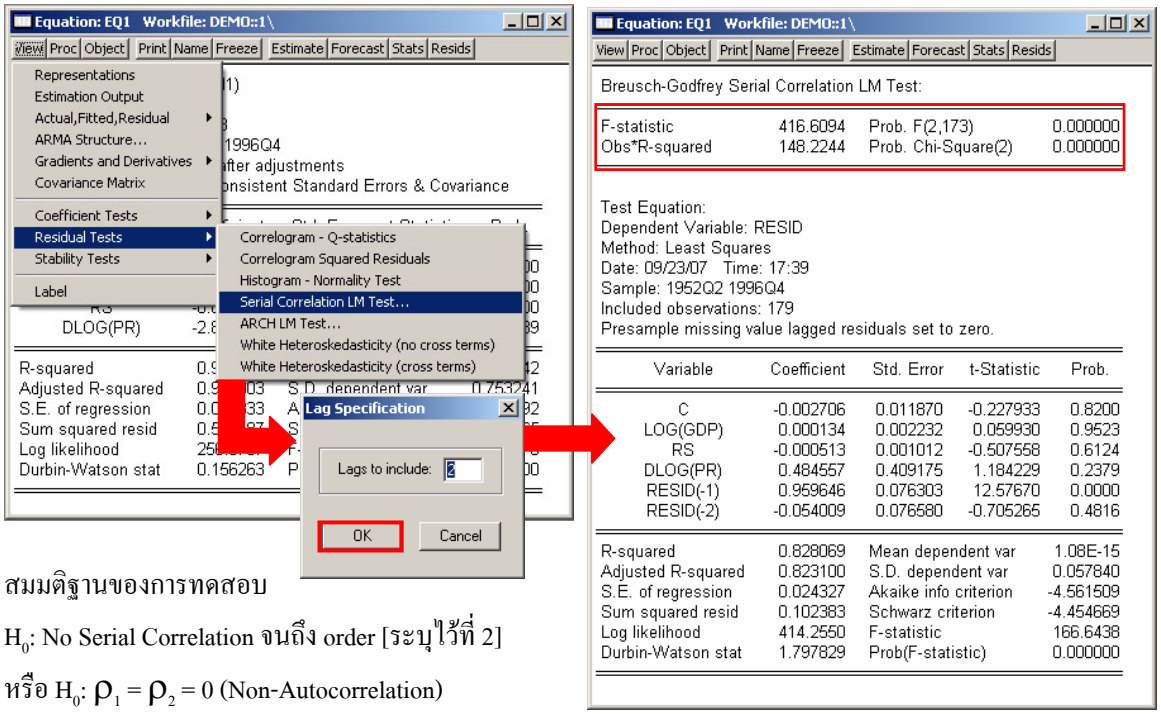
หลังจากผู้ใช้ประมาณค่าสัมประสิทธิ์ของสมการถดถอยเสร็จแล้ว [หน้าต่าง **Object Equation** จะถูกเปิดอยู่] ให้ผู้ใช้เลือก **View/Residual Tests/Correlogram – Q-statistics** ก็จะปรากฏหน้าต่าง **Lag Specification** แล้วผู้ใช้ต้องกำหนดค่า Lag ที่ต้องการ [ปกติถ้ามีจำนวนตัวอย่างมาก โปรแกรม EViews จะกำหนดค่า **Default** ไว้ที่ 36] เมื่อกำหนดค่า Lag เสร็จแล้วให้กดปุ่ม OK โปรแกรม EViews จะนำเสนอ Correlogram of Residuals พร้อมค่า Q-statistics ดังรูป



จากกรณีตัวอย่างจะเห็นได้ว่า Correlogram of Residuals ของ Autocorrelation (ACF) มีลักษณะลดลงแบบ Exponential ในขณะเดียวกันค่า Q-statistics ที่คำนวณได้มีค่าสูงกว่าค่าวิกฤตของ Chi-square ณ ระดับนัยสำคัญ 0.01 [Prob. < 0.01] แสดงว่า ค่าความคลาดเคลื่อนมีสหสัมพันธ์ในตัวเอง ดังนั้นจึงสรุปได้ว่า สมการถดถอยนี้น่าจะมีปัญหา Autocorrelation

➢ ***Breusch-Godfrey Test***

หลังจากผู้ใช้ประมาณค่าสัมประสิทธิ์ของสมการถดถอยเสร็จแล้ว [หน้าต่าง **Object Equation** จะถูกเปิดอยู่] ให้ผู้ใช้เลือก **View/Residual/Serial Correlation LM Test** ก็จะปรากฏหน้าต่าง **Lag Specification** แล้วผู้ใช้ต้องกำหนดค่า Lag ที่ต้องการ [ปกติโปรแกรม EViews จะกำหนดค่า **Default** ไว้ที่ 2] เมื่อกำหนดค่า Lag เสร็จแล้วให้กดปุ่ม OK โปรแกรม EViews จะนำเสนอผลลัพธ์ของวิธีการ Breusch-Godfrey Serial Correlation LM Test ดังรูป



สมมติฐานของการทดสอบ

H_0: No Serial Correlation จนถึง order [ระบุไว้ที่ 2]

หรือ H_0: $\rho_1 = \rho_2 = 0$ (Non-Autocorrelation)

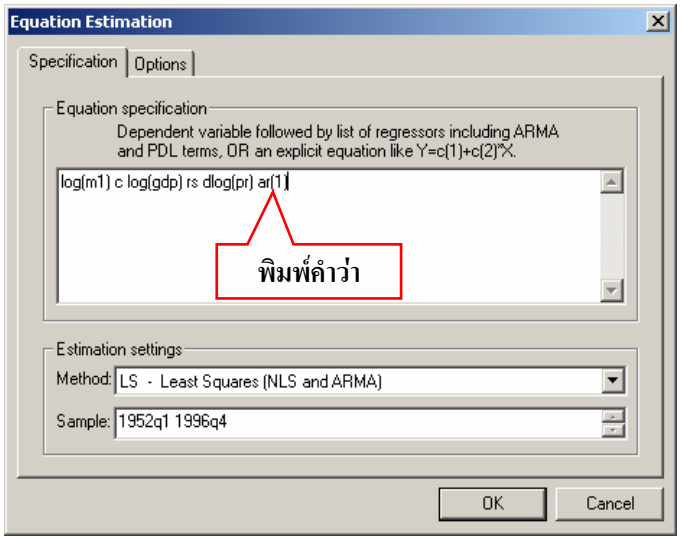
ผลการทดสอบ พบว่า ค่าสถิติ nR^2 ที่คำนวณได้มีค่าสูงกว่าค่าวิกฤต ณ ระดับความเชื่อมั่นที่ 99% [Prob. < 0.01] จึงปฏิเสธสมมติฐานหลัก แสดงว่า สมการถดถอยนี้มีปัญหา Autocorrelation

7.3 การแก้ไขกรณีเกิดปัญหา Autocorrelation

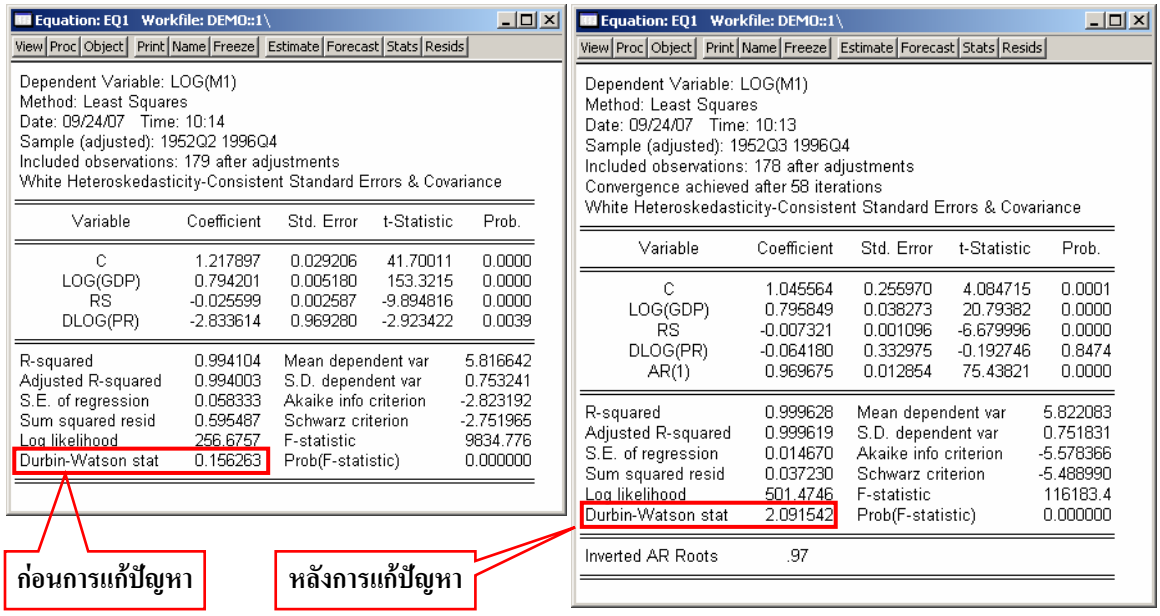
การแก้ไขปัญหา Autocorrelation สามารถทำได้หลายวิธี สำหรับในโปรแกรม EViews มีคำสั่งสำเร็จรูปในการแก้ไขปัญหา Autocorrelation โดยคำสั่งดังกล่าวจะแก้ไขปัญหา Autocorrelation ด้วยวิธี The Cochrane-Orcutt Iterative Method วิธีการนี้จะพยายามหาค่า ρ ที่แท้จริง เพื่อนำมาปรับตัวแปรในสมการถดถอย โดยทำการประมาณค่าหลายๆ รอบจนกว่าการเปลี่ยนแปลงของค่า ρ จะมีค่าน้อยภายใต้เงื่อนไขทางสถิติที่ยอมรับได้ ผู้ใช้สามารถใช้โปรแกรม EViews ในการแก้ไขปัญหา Autocorrelation ด้วยวิธีการดังกล่าว ได้ดังนี้

หลังจากผู้ใช้ประมาณค่าสัมประสิทธิ์ของสมการถดถอยเสร็จแล้ว [หน้าต่าง **Object Equation** จะถูกเปิดอยู่] และเมื่อผู้ใช้ทำการตรวจสอบปัญหา Autocorrelation [ด้วยวิธีการใดก็ได้] แล้วพบว่า สมการถดถอยของตัวเองมีปัญหา Autocorrelation [กรณีตัวอย่างในเอกสารฉบับนี้สมการถดถอยที่ประมาณค่าได้มีปัญหา Autocorrelation] ให้ผู้ใช้ดำเนินการตามขั้นตอนต่อไปนี้

ขั้นตอนที่ 1 จากหน้าต่าง **Object Equation** ให้ผู้ใช้กดปุ่ม Estimate หรือเลือก **Proc/Specify/Estimate…** โปรแกรม EViews จะเปิดหน้าต่าง **Equation Estimation** ขึ้นมาให้ แล้วให้ผู้ใช้พิมพ์คำสั่งว่า "**AR(1)**" [อักษรตัวใหญ่หรือเล็กก็ได้] ต่อท้ายคำสั่งเดิมที่อยู่ในช่อง Equation specification ดังรูป



ขั้นตอนที่ 2 เสร็จแล้วให้กดปุ่ม OK โปรแกรม EViews จะทำการประมวลผลใหม่ พร้อมทั้งแก้ไขปัญหา Autocorrelation ด้วยวิธีการ The Cochrane-Orcutt Iterative Method แล้วจะนำเสนอผลลัพธ์ใหม่ ดังรูป



จากกรณีตัวอย่าง ภายหลังจากที่โปรแกรม EViews ได้แก้ไขปัญหา Autocorrelation ด้วยวิธีการ The Cochrane-Orcutt Iterative Method แล้ว พบว่า ค่า D.W. มีค่าใกล้เคียง 2 และเมื่อนำค่าสถิติ D.W. ที่คำนวณได้ (2.092) ไปเปรียบเทียบกับค่าวิกฤติในตารางสถิติ Durbin-Watson ที่ n=200 และ k = 3 [d_L = 1.738 และ d_U = 1.799] พบว่า ค่า D.W. ที่คำนวณได้มีค่า 2.092 $>$ d_U [1.799] แสดงว่า ไม่สามารถปฏิเสธสมมติฐานหลักที่ว่า "H_0: ไม่มี Positive Autocorrelation" หมายความว่า สมการถดถอยนี้ไม่มีปัญหา Autocorrelation แล้ว

8. สมการถดถอยเมื่อตัวแปรตามเป็นตัวแปรหุ่น (Regression on Dummy Dependent Variable)

การวิเคราะห์สมการถดถอยโดยทั่วไปได้กำหนดให้ตัวแปรตามเป็นตัวแปรต่อเนื่อง หรือเป็นข้อมูลเชิงปริมาณ แต่ในบางกรณีตัวแปรตามก็เป็นตัวแปรเชิงคุณภาพ เช่น เลือกซื้อหรือไม่, จะจ่ายหรือไม่ เป็นต้น ตัวแปรตามลักษณะนี้จะมีค่าเพียง 2 ค่า หรือที่เรียกว่า Binary choice โดยจะมีค่า 0 และ 1 เท่านั้น ดังนั้นการวิเคราะห์แบบจำลองประเภทนี้จะเป็นการวิเคราะห์ เพื่อหาความน่าจะเป็นที่จะเกิดเหตุการณ์ใดเหตุการณ์หนึ่ง [ระหว่าง 0 กับ 1] ซึ่งความน่าจะเป็นดังกล่าวจะมีค่าระหว่าง 0 – 1 เสมอ โดยทั่วไปแบบจำลองที่ใช้ในการวิเคราะห์กรณีนี้มีอยู่ 3 แบบจำลอง ดังนี้

ก. แบบจำลองความน่าจะเป็นเชิงเส้น (Linear Probability Model)

ข. แบบจำลองโพรบิต (Probit Model)

ค. แบบจำลองโลจิต (Logit Model)

สำหรับตัวอย่างที่จะใช้สาธิตในการประมาณค่าของแบบจำลองทั้ง 3 เป็นตัวอย่างที่อยู่ในโปรแกรม EViews ซึ่งได้อ้างอิงข้อมูลมาจาก Greene (2003, p. 675) ซึ่งสามารถเปิดดูได้ที่ c:\programfile\EViews5\Example\data\binary.wf1 หรือ binary.wf1 ในตัวอย่างนี้เป็นการพิจารณาประสิทธิภาพของวิธีการสอน โดยมีข้อมูลจำนวน 32 ตัวอย่าง [ปกติการวิเคราะห์ด้วย Probit /Logit Model จะต้องมีจำนวนตัวอย่างไม่น้อยกว่า 30 ตัวอย่าง ต่อ 1 ค่าสัมประสิทธิ์ที่ต้องการประมาณค่า] และมีรูปแบบฟังก์ชันดังนี้

$$GRADE = \alpha + \beta_1 GPA + \beta_2 TUCE + \beta_3 PSI + \varepsilon$$

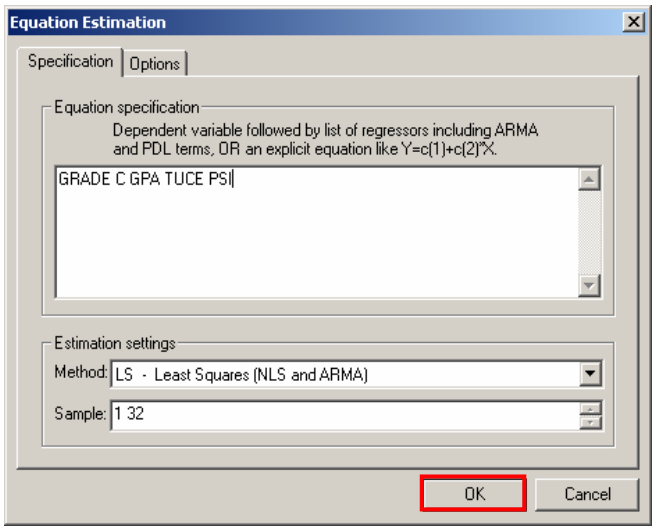
โดยที่	GRADE	คือ	เป็นตัวแปรหุ่นที่แสดงถึงการพัฒนาของ grades	
			GRADE = 1 เพิ่มขึ้น	GRADE = 0 อื่นๆ
	GPA	คือ	ระดับ GPA ของนักเรียน	
	TUCE	คือ	คะแนนทดสอบ	
	PSI	คือ	เป็นตัวแปรหุ่นที่แสดงถึงวิธีการสอน	
			PSI = 1 วิธีการสอนใหม่	PSI = 0 อื่นๆ

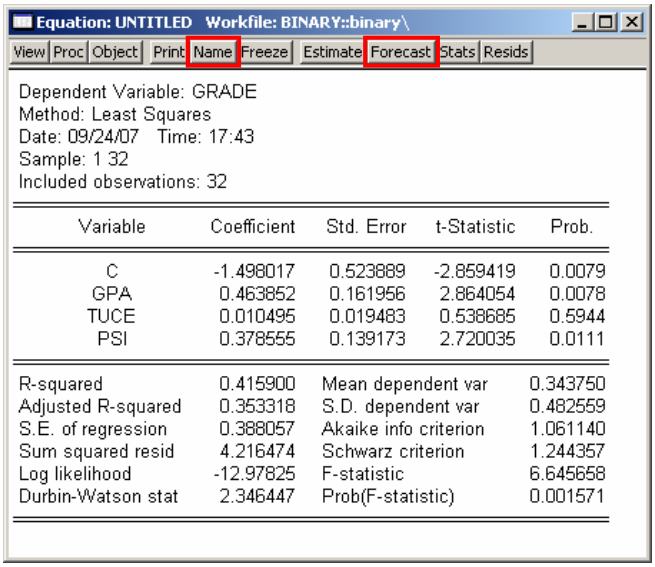
จากกรณีตัวอย่างดังกล่าวสามารถวิเคราะห์แบบจำลองทั้ง 3 แบบจำลอง ด้วยโปรแกรม EVeiws ได้ดังต่อไปนี้

8.1 แบบจำลองความน่าจะเป็นเชิงเส้น (Linear Probability Model)

ขั้นตอนที่ 1 นำข้อมูลเข้าโปรแกรม EViews [แล้วแต่ว่าผู้ใช้จะใช้วิธีใด]

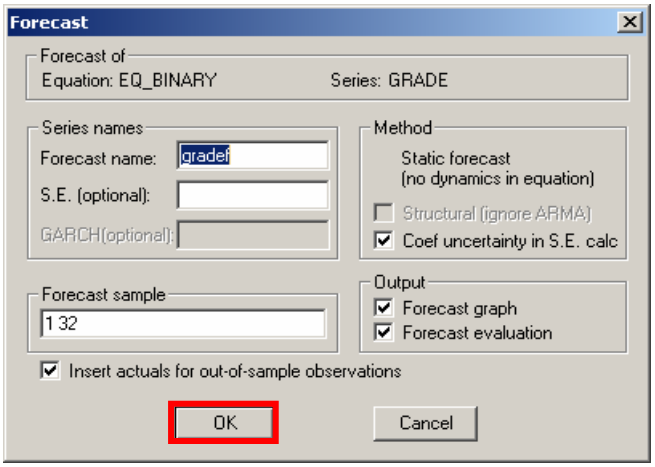
ขั้นตอนที่ 2 เลือก **Objects/New Object/ Equation** จากเมนู **Workfile** แล้วตั้งชื่อ Objects Equation หรือเลือก **Quick/Estimate Equation** จากเมนูหลัก (**Main menu**) จะปรากฏหน้าต่าง **Equation Specification** ให้ผู้ใช้พิมพ์คำสั่งว่า **"GRADE C GPA TUCE PSI"** ลงในช่อง Equation specification และให้เลือกวิธีการ **"LS-Least Squares (NLS and ARMA)"** ดังรูป เสร็จแล้วกดปุ่ม OK โปรแกรม EViews จะนำเสนอผลลัพธ์ ดังรูป



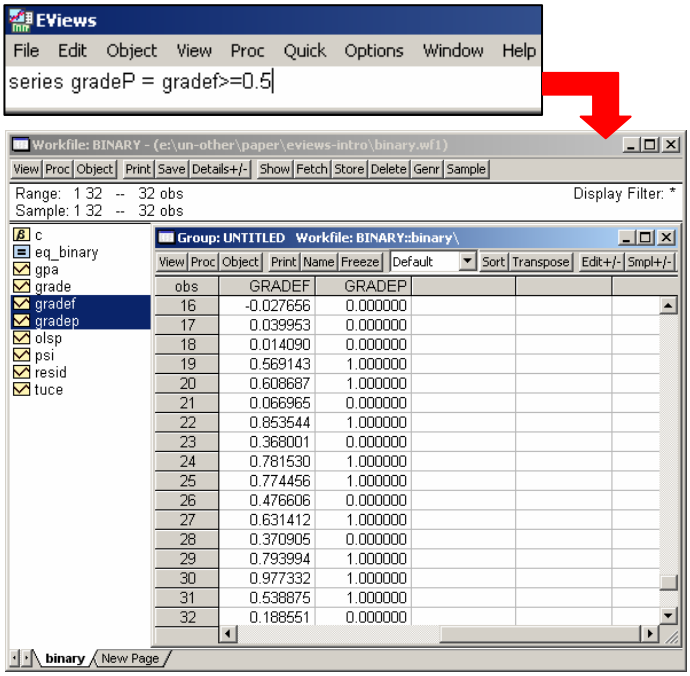


ขั้นตอนที่ 3 บันทึก Objects Equation โดยกดปุ่ม Name ที่แถบเครื่องมือของหน้าต่าง **Objects Equation** แล้วตั้งชื่อให้ Objects นี้ว่า"EQ_Binary" ซึ่งที่หน้าต่าง Workfile จะปรากฏ Objects eq_binary [eq_binary] หลังจากนั้นให้ Save Workfile นี้

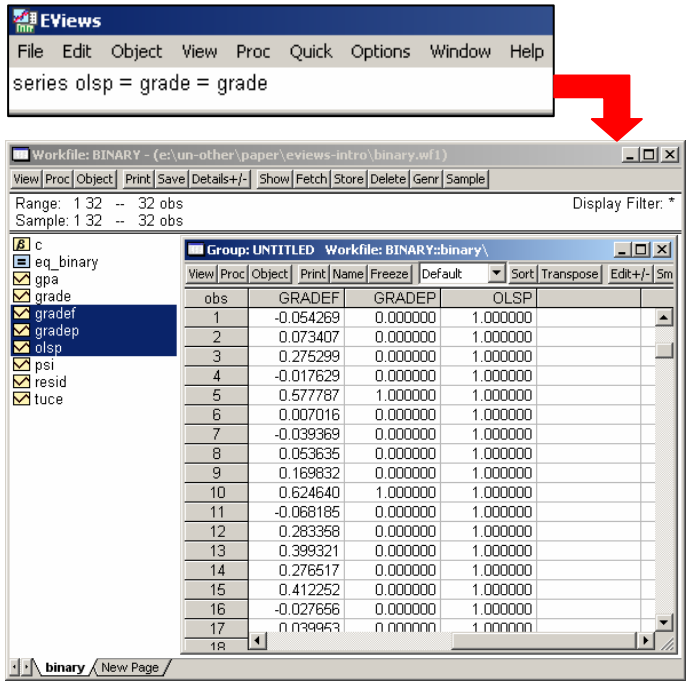
ขั้นตอนที่ 4 คำนวณหาค่าความน่าจะเป็นในการทำนายเหตุการณ์ โดยกดปุ่ม Forecast ที่แถบเครื่องมือของหน้าต่าง **Objects Equation** ก็จะปรากฏหน้าต่าง **Forecast** ให้เติมชื่อตัวแปรที่จะ Forecast ในช่อง Forecast name ว่า "gradef" เสร็จแล้วให้กดปุ่ม OK ดังรูป โปรแกรม EViews จะสร้าง Series gradef ใน Workfile



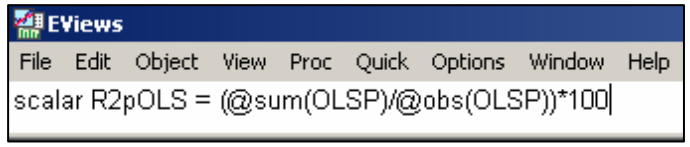
ขั้นตอนที่ 5 ให้พิมพ์คำสั่งว่า "**series gradep= gradef >= 0.5**" [***หมายความว่า ให้สร้าง Series gradep โดยมีเงื่อนไขว่า ถ้าค่าใน Series gradef ≥ 0.5 ให้ค่าใน Series gradep มีค่าเป็น 1 ในกรณีอื่นๆ มีค่าเป็น 0***] ลงในหน้าต่าง **Command** เสร็จแล้วกด **Enter** โปรแกรมก็จะทำตามคำสั่งดังกล่าว โดยจะสร้าง Series gradep ตามเงื่อนไขที่กำหนด ดังรูป



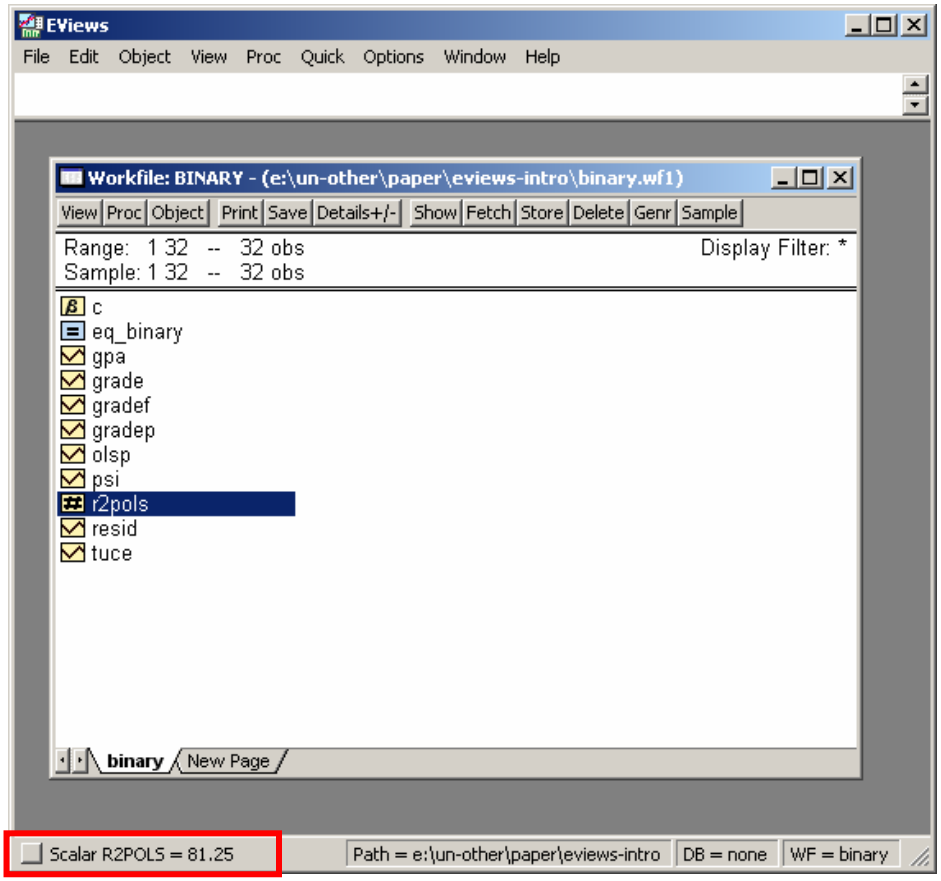
ขั้นตอนที่6 ให้พิมพ์คำสั่งว่า "**series olsp = gradep = grade**" [***หมายความว่า ให้สร้าง Series olsp โดยมีเงื่อนไขว่า ถ้าแบบจำลองที่ประมาณค่าถูก (ทำนายถูก) ให้ค่าใน Series olsp มีค่าเป็น 1 ในกรณีอื่นๆ มีค่าเป็น 0***] ลงในหน้าต่าง **Command** เสร็จแล้วกด **Enter** โปรแกรมก็จะทำตามคำสั่งดังกล่าว โดยจะสร้าง Series olsp ตามเงื่อนไขที่กำหนด ดังรูป



ขั้นตอนที่ 7 คำนวณเปอร์เซ็นต์การทำนายถูกของแบบจำลอง โดยการพิมพ์คำสั่งว่า "**scalar R2pOLS = (@sum(OLSP)/@obs(OLSP)) *100**" [***หมายความว่า ให้คำนวณค่าเปอร์เซ็นต์การทำนายถูกโดยการนำจำนวนค่าการทำนายถูกหารด้วยจำนวนตัวอย่างทั้งหมดแล้วคูณด้วย 100***] ลงในหน้าต่าง **Command** เสร็จแล้วกด **Enter** โปรแกรมก็จะสร้าง objects Scalar [r2pols]



ขั้นตอนที่ 8 ถ้าหากต้องการดูค่าเปอร์เซ็นต์การทำนายถูก ให้กด Double click ที่ Objects Scalar [r2pols] โปรแกรม EViews จะแสดงค่าเปอร์เซ็นต์การทำนายถูกที่ Status Line [อยู่ด้านล่างของหน้าต่างหลัก] ดังรูป



ผลการวิเคราะห์ สรุปได้ว่า การเปลี่ยนแปลงของ GPA และวิธีการสอนใหม่ ได้ทำให้ความน่าจะเป็นในการพัฒนา grades ของนักเรียนดีขึ้น อย่างมีนัยสำคัญทางสถิติที่ ณ ระดับความเชื่อมั่น 99% [พิจารณาจากเครื่องหมายของค่าสัมประสิทธิ์ที่ประมาณได้ และพิจารณาจากค่า Prob. ใน Column สุดท้าย ซึ่งค่า Prob. < 0.01] โดยแบบจำลองนี้สามารถทำนายเหตุการณ์ได้ถูกต้อง 81.25%

การประมาณแบบจำลองความน่าจะเป็นเชิงเส้นข้างต้น มีปัญหาบางประการ ได้แก่

ก. ตัวคลาดเคลื่อนมีการกระจายแบบไม่ปกติ (Nonnormality of Distribution)

ข. ตัวคลาดเคลื่อนมีความแปรปรวนไม่คงที่ (Heteroskedasticity)

ค. ค่าพยากรณ์ของ Y [$\hat{Y}$] มีค่าไม่อยู่ในช่วง 0 กับ 1 ($0 \leq E(Y|X) \leq 1$)

ง. ค่า R^2 ไม่สามารถวัดประสิทธิภาพของแบบจำลองได้

จากปัญหาข้างต้นจึงทำให้การประมาณค่าแบบจำลองความน่าจะเป็นเชิงเส้นด้วยวิธีการ OLS ขาดความน่าเชื่อถือ จึงได้มีการเสนอแบบจำลองโพรบิตและโลจิต (Probit and Logit Model) มาใช้ในกรณีที่ตัวแปรตามมีลักษณะเป็น Binary และใช้วิธีการ MLE (Maximum Likelihood Estimation) ในการประมาณค่าสัมประสิทธิ์ของแบบจำลองดังกล่าว

8.2 แบบจำลองโพรบิตและโลจิต (Probit and Logit Model)

ในการวิเคราะห์ด้วยแบบจำลองโลจิตและโพรบิตมีเงื่อนไขพื้นฐานบางประการดังนี้

ก. ตัวแปรตามต้องเป็น Binary Response ส่วนตัวแปรอิสระอาจเป็น Dummy Variable /Interval /Ratio Scale ก็ได้

ข. ค่าคาดหวัง (ค่าเฉลี่ย) ของตัวคลาดเคลื่อนมีค่าเป็น 0 [$E(\varepsilon_i) = 0$]

ค. ค่าคลาดเคลื่อนไม่มีความสัมพันธ์กันเอง [$Cov(\varepsilon_i,\varepsilon_j) = 0$]

ง. ตัวแปรอิสระกับค่าคลาดเคลื่อนจะต้องเป็นอิสระแก่กัน

จ. ตัวแปรอิสระจะต้องไม่มีความสัมพันธ์กันเอง

ฉ. จำนวนตัวอย่างต้องมีอย่างน้อยหรือเท่ากับ 30*P [$n \geq 30*P$] [P คือ จำนวน Parameter]

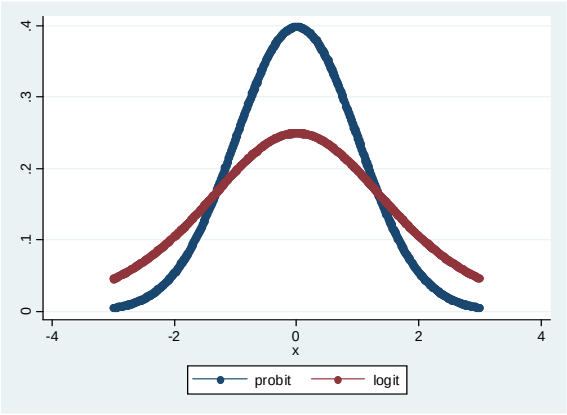
รูปแบบทั่วไปของแบบจำลอง $$\text{prob}\left(y_i = 1 \mid X\right) = F\left(x_i'\beta\right)$$

ฟังก์ชัน Probit ของ Probit Model $$\text{prob}\left(y_i = 1\right) = \Phi\left(\frac{x_i'\beta}{\sigma}\right) = \int_{-\alpha}^{\frac{x_i'\beta}{\sigma}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz$$

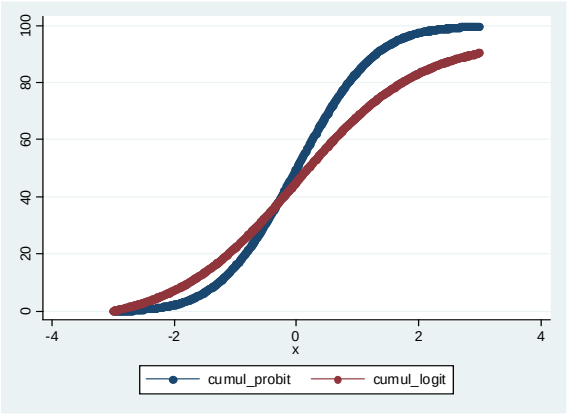
$$\text{prob}\left(y_i = 1\right) = \Phi\left(\frac{x_i'\beta}{\sigma}\right)$$

ฟังก์ชัน Logistic ของ Logit Model $$\text{prob}\left(y_i = 1\right) = \frac{e^{\beta' x}}{1 + e^{\beta' x}}$$

ความแตกต่างระหว่างแบบจำลอง Probit และ Logit อยู่ที่การกำหนดการแจกแจงของตัวคลาดเคลื่อน โดยแบบจำลอง Probit ได้กำหนดให้ตัวคลาดเคลื่อนมีการแจกแจงแบบปกติ (Normal Distribution) ในขณะที่แบบจำลอง Logit ได้กำหนดให้ตัวคลาดเคลื่อนมีการแจกแจงแบบโลจิต (Logistic Distribution) ซึ่งสามารถดูได้จากกราฟแสดงความแตกต่างระหว่างแบบจำลองโพรบิตและโลจิตดังนี้



Probability Density Function ของ Probit และ Logit

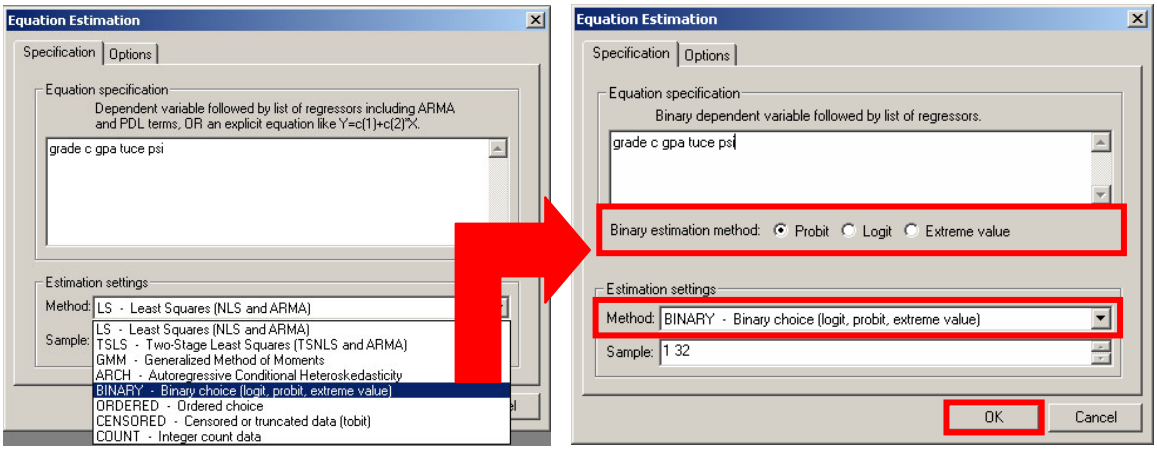


Cumulative Distribution function ของ Probit และ Logit

จากกรณีตัวอย่างสามารถวิเคราะห์แบบจำลอง Probit และ Logit ด้วยโปรแกรม EVeiws ได้ดังนี้

ขั้นตอนที่ 1 นำข้อมูลเข้าโปรแกรม EViews [แล้วแต่ว่าผู้ใช้จะใช้วิธีใด]

ขั้นตอนที่ 2 เลือก **Objects/New Object/Equation** จากเมนู **Workfile** แล้วตั้งชื่อ Objects Equation หรือเลือก **Quick/Estimate Equation** จากเมนูหลัก (**Main menu**) จะปรากฏหน้าต่าง **Equation Estimation** ต่อมาให้ผู้ใช้พิมพ์คำสั่งว่า "**GRADE C GPA TUCE PSI**" ลงในช่อง Equation specification และในช่อง Method ให้เลือกวิธีการ "**BINARY – Binary Choice (logit, probit, extreme value)**" จะมี Options ขึ้นมาให้เลือกวิธีการประมาณค่า 3 วิธีการ [ในตัวอย่างเลือก Probit] เมื่อเลือกเสร็จแล้วก็กดปุ่ม OK ดังรูป

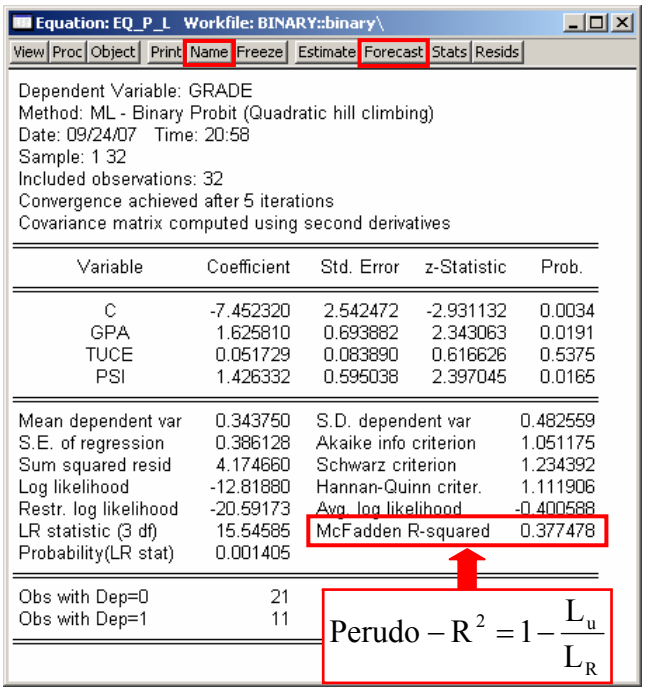


ขั้นตอนที่ 3 โปรแกรม EViews จะนำเสนอผลการวิเคราะห์ ผู้ใช้จะต้องบันทึก Objects Equation ด้วยการกดปุ่ม Name ที่แถบเครื่องมือของหน้าต่าง **Objects Equation** แล้วตั้งชื่อให้ Objects นี้ เช่น "EQ_P_L" ซึ่งที่หน้าต่าง Workfile จะปรากฏ Object eq_p_l [eq_p_l] หลังจากนั้นให้ Save Workfile นี้

ค่าสัมประสิทธิ์ที่ได้ของแบบจำลองโพรบิตและโลจิตจะมีค่าสูงกว่าของแบบจำลองความน่าจะเป็นเชิงเส้น และค่าสัมประสิทธิ์ที่ได้จากโปรแกรม EViews ไม่ใช่ค่า Marginal effect ของตัวแปรอิสระที่มีเงื่อนไขความน่าจะเป็นว่า

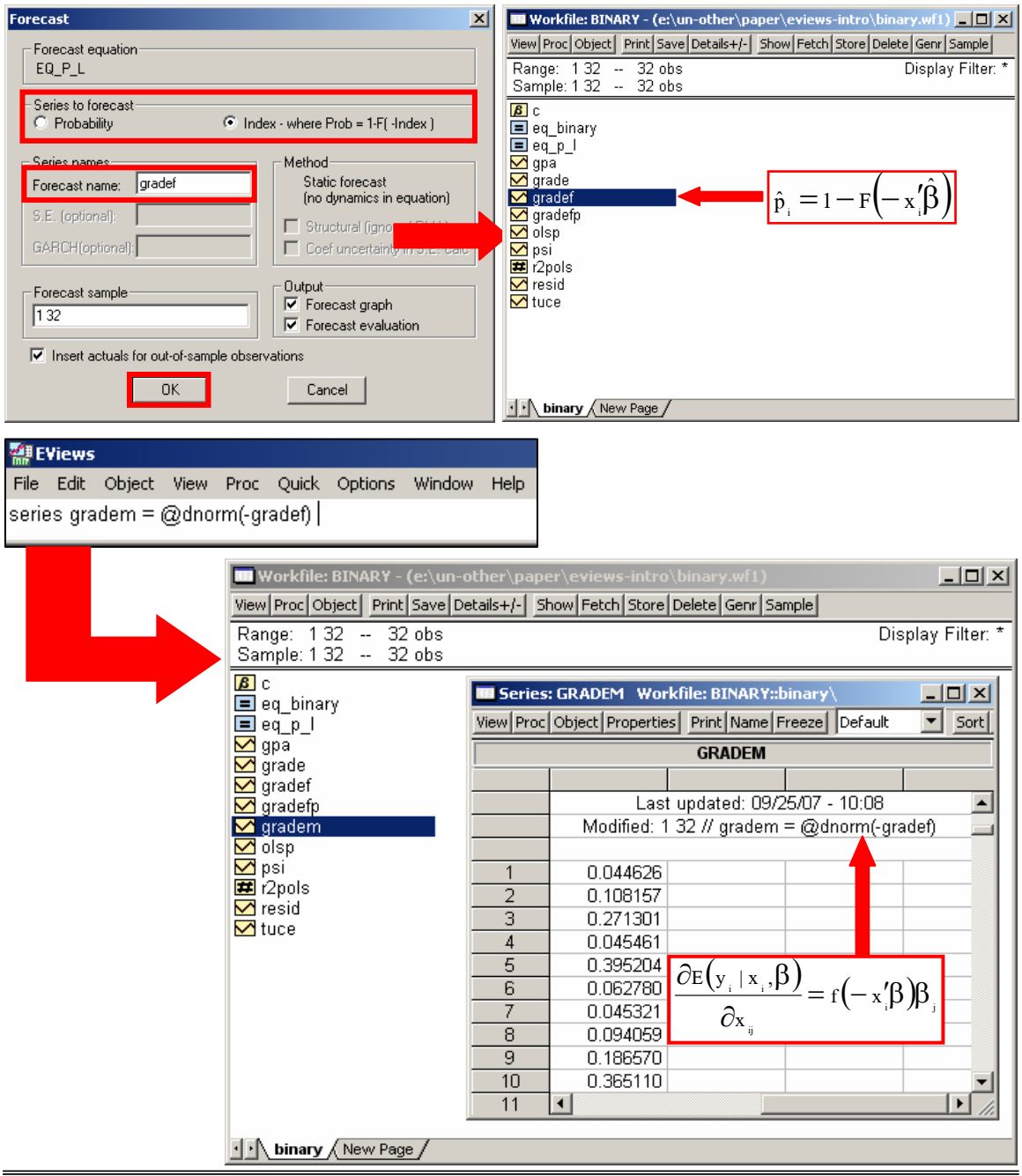
$$\frac{\partial E(y_i \mid x_i, \beta)}{\partial x_{ij}} = f(-x_i'\beta)\beta_j$$

ดังนั้นถ้าผู้ใช้ต้องการคำนวณค่า Marginal effect ผู้ใช้จะต้องคำนวณหาค่า Fitted probability [$\hat{p}_i = 1 - F(-x_i'\hat{\beta})$] ก่อน โดยเลือก **Proc/ Forecast (Fitted Probability/Index)…** หรือกดปุ่ม Forecast ที่หน้าต่าง **Objects Equation** ดังรูป

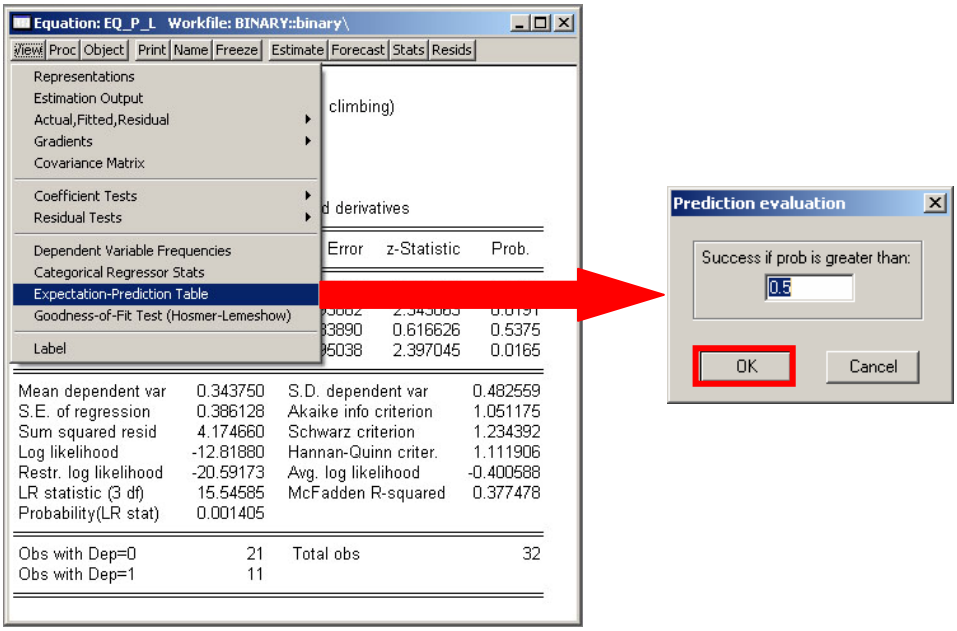


เสร็จแล้วให้สร้าง Series ของค่า Fitted probability [สมมติว่า gradef] ต่อมาให้พิมพ์คำสั่งว่า **"series gradem = @dnorm (-gradef)"** [สำหรับ Probit Model] หรือ **"series gradem = @dlogistic(-gradef)"** [สำหรับ Logit Model] ลงในหน้าต่าง **Command** แล้วกด **Enter** โปรแกรมจะคำนวณค่าคาดหวังของ y_i ที่เป็น Marginal effect ของตัวแปรอิสระภายใต้เงื่อนไขความน่าจะเป็นที่ว่า $\frac{\partial E(y_i \mid x_i, \beta)}{\partial x_{ij}} = f(-x_i'\beta)\beta_j$ [ดูขั้นตอนตามรูปข้างล่าง]

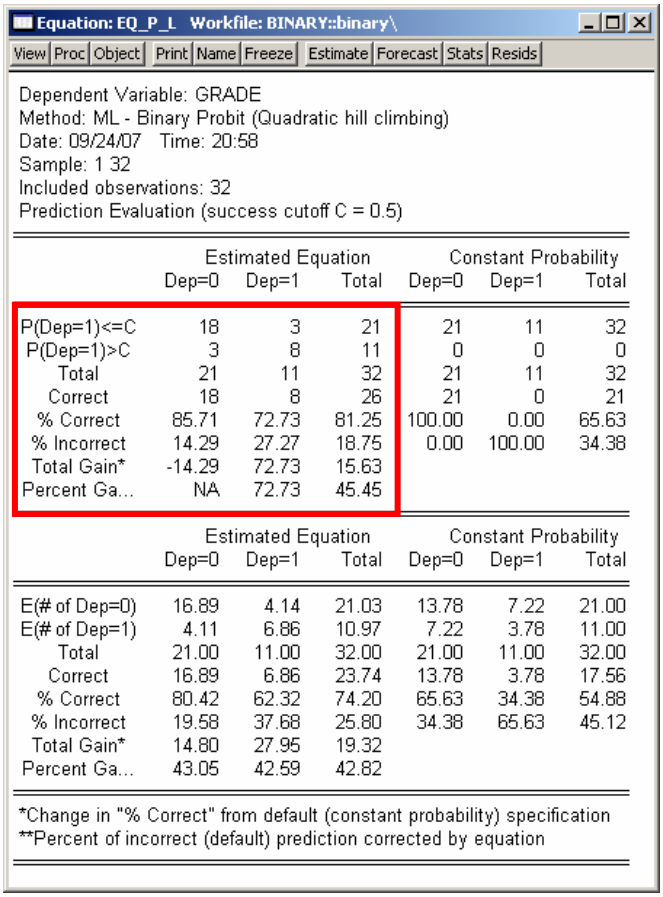
สำหรับค่าสถิติในการตัดสินใจของแบบจำลอง Probit และ Logit ส่วนใหญ่นิยมใช้ค่า McFadden R-squared ที่คำนวณมาจาก $\text{Perudo} - R^2 = 1 - \frac{L_u}{L_R}$ ซึ่งโดยปกติจะมีค่าไม่เกิน 0.40



ขั้นตอนที่ 4 ถ้าหากผู้ใช้ต้องการดูค่าเปอร์เซ็นต์การทำนาย ให้เลือก **View/Expectation-Prediction Table** ที่หน้าต่าง **Object Equation** แล้วจะปรากฏหน้าต่าง **Prediction evaluation** ให้เติมค่า Success if prob is greater than ซึ่งโปรแกรม EViews จะ **Default** ที่ 0.5 ถ้าหากผู้ใช้ไม่มีการเปลี่ยนแปลงก็ให้กดปุ่ม OK ดังรูป



ภายหลังจากกดปุ่ม OK โปรแกรม EViews จะนำเสนอผลลัพธ์ ดังรูป



ถ้าหากผู้ใช้ต้องการประมาณค่าแบบจำลอง Logit ก็ให้ผู้ใช้เลือก Option ที่หน้าต่าง **Equation specification** เป็น Logit ดังรูป

แบบจำลองโพรบิตและโลจิตใช้วิธีการ MLE ในการประมาณค่าสัมประสิทธิ์ และได้แก้ไขปัญหาในเรื่อง ค่าประมาณของ Y ไม่อยู่ในช่วง 0 กับ 1 ($0 \leq E(Y|X) \leq 1$) ของแบบจำลองความน่าจะเป็นเชิงเส้น ทำให้แบบจำลองโพรบิตและโลจิตมีความน่าเชื่อถือมากกว่า ในทางปฏิบัติมักนิยมใช้แบบจำลองโลจิตมากกว่าโพรบิต เนื่องจากเปรียบเทียบในทางคณิตศาสตร์ได้ง่ายกว่า และค่าสัมประสิทธิ์ของแบบจำลองโพรบิตจะน้อยกว่าแบบจำลองโลจิต [Amemiya (1981) ได้เสนอว่า $\beta_{probit} = 0.625 \times \beta_{logit}$]

9. การวิเคราะห์ข้อมูลอนุกรมเวลา (Time Series Analysis)

อนุกรมเวลา คือ เซตของข้อมูลเชิงปริมาณที่มีการจัดเก็บในช่วงเวลาที่ติดต่อกัน ส่วนข้อมูลอนุกรมเวลา (Time Series Data) ก็คือ ชุดของข้อมูลที่มีการเก็บรวบรวมตามระยะเวลาที่ติดต่อกันอย่างเป็นระบบโดยทั่วไปข้อมูลอนุกรมเวลาจะประกอบด้วยองค์ประกอบ 4 ส่วน คือ แนวโน้ม (Trend: T), ฤดูกาล (Seasonal: S), วัฏจักร (Cycle: C) และเหตุการณ์ที่ผิดปกติ/เหตุการณ์ความไม่แน่นอน (Irregular: I) สำหรับรูปแบบของอนุกรมเวลาโดยทั่วไปนั้นมีอยู่ 2 รูปแบบคือ

ก. รูปแบบบวก $Y = T+S+C+I$

ข. รูปแบบคูณ $Y = T\times S\times C\times I$

โดยที่ Y คือ อนุกรมเวลา
T คือ อิทธิพลของแนวโน้ม
S คือ อิทธิพลของฤดูกาล
C คือ อิทธิพลของวัฏจักร
I คือ อิทธิพลของความไม่แน่นอน

สำหรับในคู่มือฉบับนี้จะแบ่งเรื่องของการวิเคราะห์ข้อมูลอนุกรมเวลาออกเป็น 2 ส่วน คือ การทดสอบข้อมูลอนุกรมเวลาและวิธีการพยากรณ์ โดยจะสาธิตการใช้โปรแกรม EViews ในการวิเคราะห์ทั้ง 2 ส่วน ดังนี้

9.1 การทดสอบข้อมูลอนุกรมเวลา

การทดสอบข้อมูลอนุกรมเวลา เป็นสิ่งที่ควรกระทำก่อนที่จะนำข้อมูลอนุกรมเวลามาใช้ในการวิเคราะห์ โดยเฉพาะเงื่อนไขความคงที่ของอนุกรมเวลา (Stationary) ซึ่งเป็นเงื่อนไขที่สำคัญในการนำข้อมูลอนุกรมเวลามาใช้ ดังนั้นเพื่อความเข้าใจในที่นี้ขออธิบายในประเด็นดังกล่าวพอสังเขป ดังนี้

<u>การคงที่ของอนุกรมเวลา (Stationary)</u>

วิธีการวิเคราะห์/การพยากรณ์หลายๆ วิธี ได้มีเงื่อนไขที่ว่า ข้อมูลอนุกรมเวลาที่นำมาใช้จะต้องคงที่ ดังนั้นถ้าหากอนุกรมเวลาที่ใช้ไม่คงที่จะต้องทำให้อนุกรมเวลาดังกล่าวคงที่ก่อน โดยการหาผลต่างของอนุกรมเวลา การคงที่ของอนุกรมเวลา หมายถึง อนุกรมเวลาที่อยู่ในสภาวะสมดุลเชิงสถิติ (Statistical equilibrium) ซึ่งก็คือ การที่คุณสมบัติทางสถิติของอนุกรมเวลาไม่เปลี่ยนแปลงไปตามกาลเวลา

เมื่อสมมติให้ตัวแปร X_t เป็นอนุกรมเวลาที่คงที่ (Stationary) ดังนั้น ตัวแปร X_t จะมีคุณสมบัติดังนี้

Mean: $E(X_t) = \mu$

Variance: $Var(X_t) = E(X_t-\mu)^2 = \sigma^2$

Covariance: $E\left[(X_t-\mu)(X_{t+k}-\mu)\right] = \gamma_k$

ถ้าตัวแปร X_t เป็นอนุกรมเวลาที่ไม่คงที่ (Nonstationary) ดังนั้น ตัวแปร X_t จะมีคุณสมบัติดังนี้

Mean: $E(X_t) = t\mu$

Variance: $Var(X_t) = E(X_t - \mu)^2 = t\sigma^2$

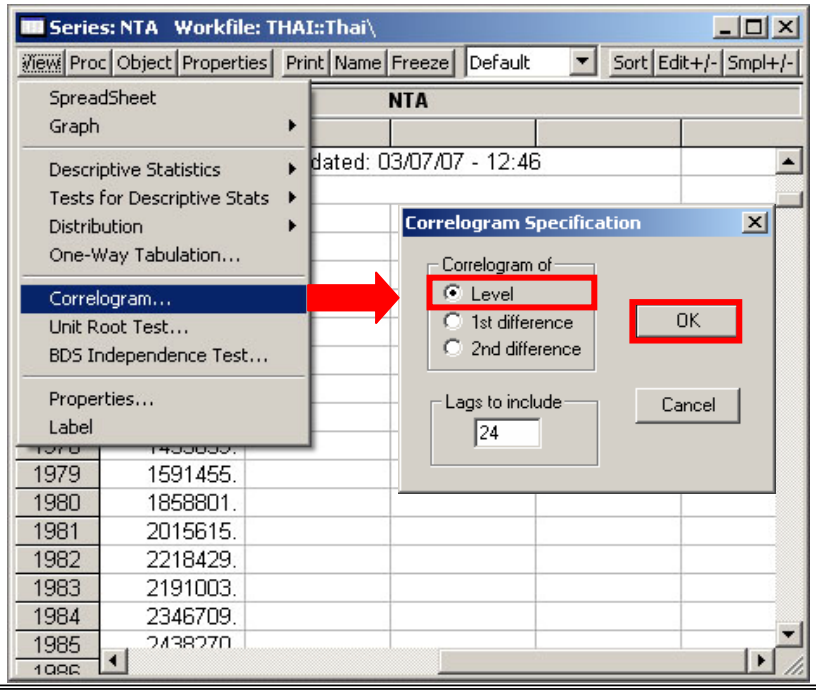
Covariance: $E[(X_t - \mu)(X_{t+k} - \mu)] = t\gamma_k$

สำหรับโปรแกรม EViews มีคำสั่งสำเร็จรูปในการตรวจสอบความคงที่ของข้อมูลอนุกรมเวลา ซึ่งในที่นี้จะนำเสนอวิธีการที่มักนิยมเอามาใช้ในการตรวจสอบความคงที่ของข้อมูล 2 วิธีการ คือ การพิจารณาจาก Correlogram and Q-statistics และวิธีการ Unit Root Test โดยข้อมูลที่นำมาใช้ในการสาธิตในครั้งนี้ เป็นข้อมูลสถิติจำนวนนักท่องเที่ยวชาวต่างชาติที่เดินทางเข้ามาท่องเที่ยวในประเทศไทย ระหว่างปี พ.ศ. 2523 – 2546 จำนวน 33 ตัวอย่าง

➢ ***Correlogram and Q-statistics***

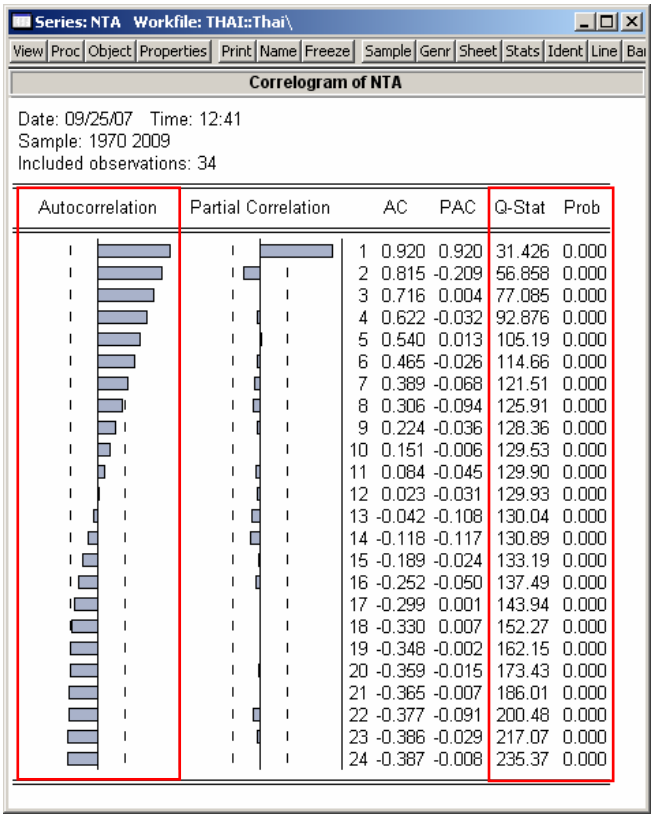
การพิจารณา Correlogram and Q-statistics มีเงื่อนไขในการพิจารณาว่ากราฟ Correlogram ของ Autocorrelation ของตัวแปรที่กำลังพิจารณาจะต้องไม่มีลักษณะการลดลงแบบ Exponential และค่า Q-statistics ที่คำนวณได้มีค่ามากกว่าค่าวิกฤตของ Chi-square ณ ระดับนัยสำคัญ 0.10 [Prob. < 0.10] สามารถใช้โปรแกรม EViews ทดสอบความคงที่ของข้อมูลด้วยการพิจารณา Correlogram and Q-statistics ได้ดังนี้

ขั้นตอนที่1 เปิดหน้าต่าง Series ของข้อมูลที่ต้องการทดสอบขึ้นมา ด้วยการ Double click ที่ Series ที่ต้องการ [Double click nta] เมื่อหน้าต่าง series เปิดแล้ว ให้เลือก **View/Correlogram…** โปรแกรมจะเปิดหน้าต่าง **Correlogram Specification** ขึ้นมา ให้กำหนดค่าต่างๆ ตามความต้องการ ในกรณีตัวอย่างกำหนด Correlation of Level และกำหนด Lags = 24 เสร็จแล้วให้กดปุ่ม OK ดังรูป

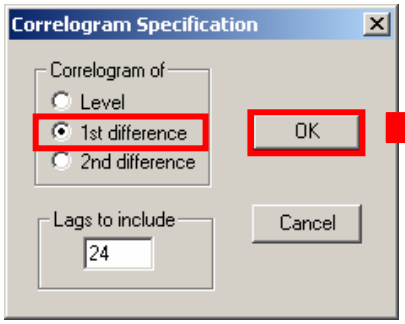


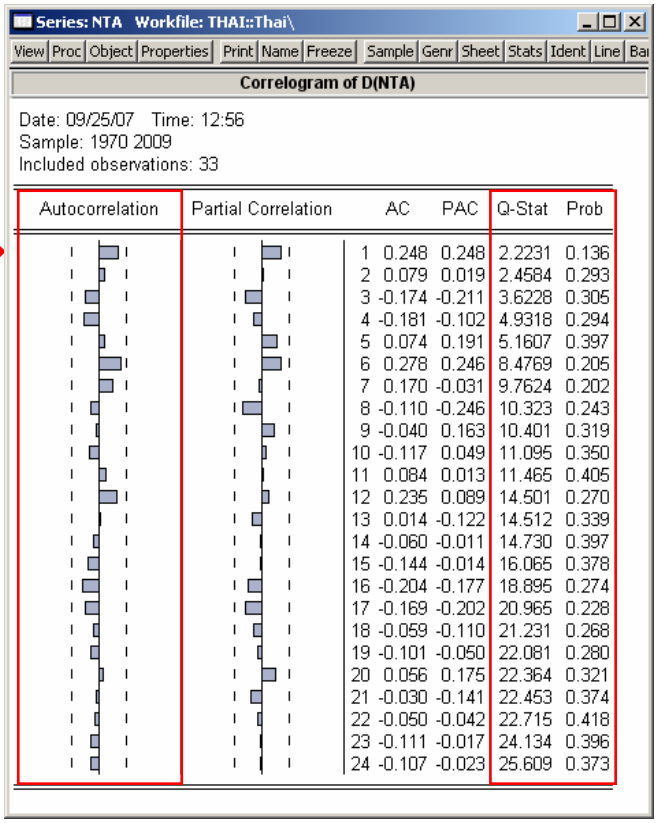
ขั้นตอนที่ 2 เมื่อกดปุ่ม OK โปรแกรม EViews จะนำเสนอผลลัพธ์ ดังรูป

ผลลัพธ์ที่โปรแกรม EViews วิเคราะห์มาให้นั้น จะเห็นได้อย่างชัดเจนว่า Correlogram ของ Autocorrelation ของตัวแปรที่ทดสอบ มีลักษณะลดลงแบบ Exponential ในขณะเดียวกันค่าสถิติ Q-statistics ก็มีค่ามากกว่าค่าวิกฤติ ณ ระดับความเชื่อมั่นทางสถิติที่ 90% [Prob. $<$ 0.10] แสดงว่า ณ ที่ Level [ยังไม่มีการแปลงข้อมูลใดๆ] ตัวแปรอนุกรมเวลาตัวนี้ มีลักษณะไม่คงที่อย่างแน่นอน ดังนั้นจึงต้องทดสอบต่อไปว่า ณ ที่ 1st Difference ตัวแปรตัวนี้จะมีลักษณะคงที่หรือไม่



ขั้นตอนที่ 3 ให้ทำตามขั้นตอนที่ 1 โดยที่หน้าต่าง Correlogram Specification ให้กำหนด Correlation of 1st Difference และกำหนด Lags = 24 เช่นเดิม เสร็จแล้วให้กดปุ่ม OK โปรแกรม EViews จะนำเสนอผลลัพธ์ใหม่ของตัวแปรที่มีการแปลงค่าโดยการ 1st Difference ดังรูป





ผลลัพธ์ใหม่แสดงให้เห็นว่า ตัวแปรอนุกรมเวลาตัวนี้ มีลักษณะคงที่เมื่อมีการแปลงข้อมูลด้วยการ 1st Difference ดังนั้นจึงสรุปได้ว่า ตัวแปรตัวนี้มีลักษณะไม่คงที่ที่ Level แต่จะคงที่ที่ 1st Difference

➢ ***Unit Root Test***

ปัจจุบันวิธีการทดสอบ Unit Root มีอยู่หลายวิธี และในโปรแกรม EViews มีวิธีการทดสอบ Unit Root อยู่หลายวิธีเช่นเดียวกัน เช่น Dickey-Fuller, Augmented Dickey-Fuller, Phillips-Perron เป็นต้น สำหรับในคู่มือฉบับนี้จะสาธิตเฉพาะวิธีการ Dickey-Fuller, Augmented Dickey-Fuller ซึ่งเป็นวิธีการที่นิยมใช้อย่างกว้างขวางในปัจจุบัน วิธีการของ Dickey-Fuller มีสมการที่ต้องการทดสอบอยู่ 3 สมการ (At Level) คือ

$$\Delta x_t = \gamma X_{t-1} + \varepsilon_t \qquad \text{(random walk process)}$$

$$\Delta x_t = \alpha + \gamma X_{t-1} + \varepsilon_t \qquad \text{(random walk with drift)}$$

$$\Delta x_t = \alpha + \beta t + \gamma X_{t-1} + \varepsilon_t \qquad \text{(random walk with drift และมี linear time trend)}$$

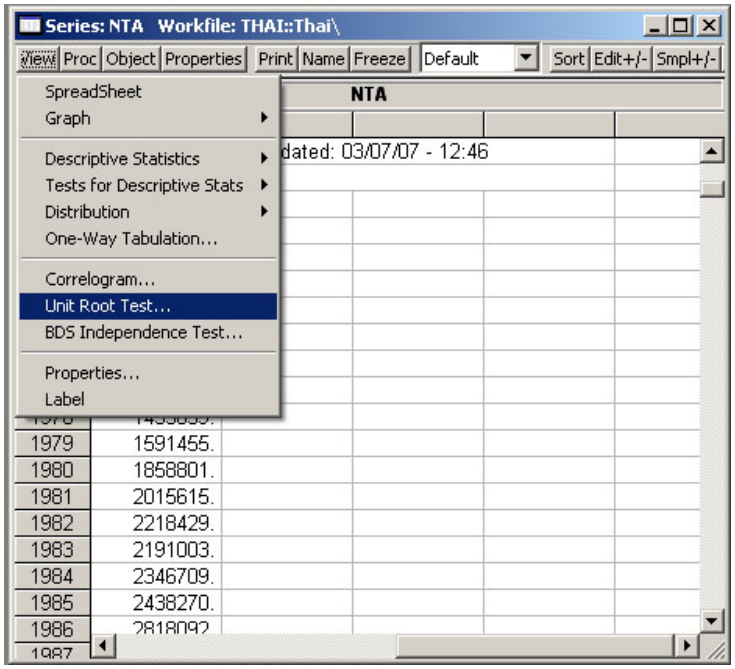
สมมติฐานที่ทดสอบ

$H_0 : \gamma = 0$

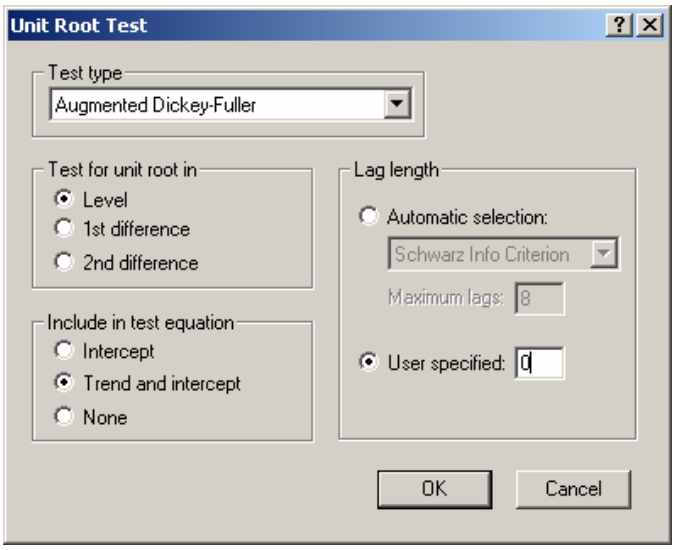
$H_a : \gamma \neq 0$

ถ้ายอมรับ H_0 แสดงว่า X_t มีลักษณะไม่นิ่ง (Nonstationary) เนื่องจาก $\gamma = (1-\rho)$ ในสมการ $X_t = \rho X_{t-1} + \varepsilon_t$ แสดงว่า ค่า X_t มีการเปลี่ยนแปลงเมื่อเวลาเปลี่ยนแปลง การทดสอบนี้สามารถทำได้ด้วยโปรแกรม EViews ดังนี้

ขั้นตอนที่ 1 เปิดหน้าต่าง Series ของข้อมูลที่ต้องการทดสอบขึ้นมา ด้วยการ Double click ที่ Series ที่ต้องการ [Double click nta] เสร็จแล้วให้เลือก **View/Unit Root Test…** ที่แถบเครื่องมือของหน้าต่าง **Series** ดังรูป



ขั้นตอนที่ 2 หลังจากนั้นจะปรากฏหน้าต่าง **Unit Root Test** ดังรูป



ขั้นตอนที่ 3 เลือก **Test Type: Augmented Dickey-Fuller**, เลือก **Test for unit root in: Level**, เลือก **Include in test equation: Trend and intercept**, เลือก **Lag length: User specified: 0** (ดังรูปข้างต้น) แล้วกดปุ่ม OK จะได้ผลดังนี้

Series: NTA Workfile: THAI::Thai\

View | Proc | Object | Properties | Print | Name | Freeze | Sample | Genr | Sheet | Stats | Ident | Line

Augmented Dickey-Fuller Unit Root Test on NTA

Null Hypothesis: NTA has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Fixed)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.772271	0.6952
Test critical values:	1% level	-4.262735	
	5% level	-3.552973	
	10% level	-3.209642	

*MacKinnon (1996) one-sided p-values.

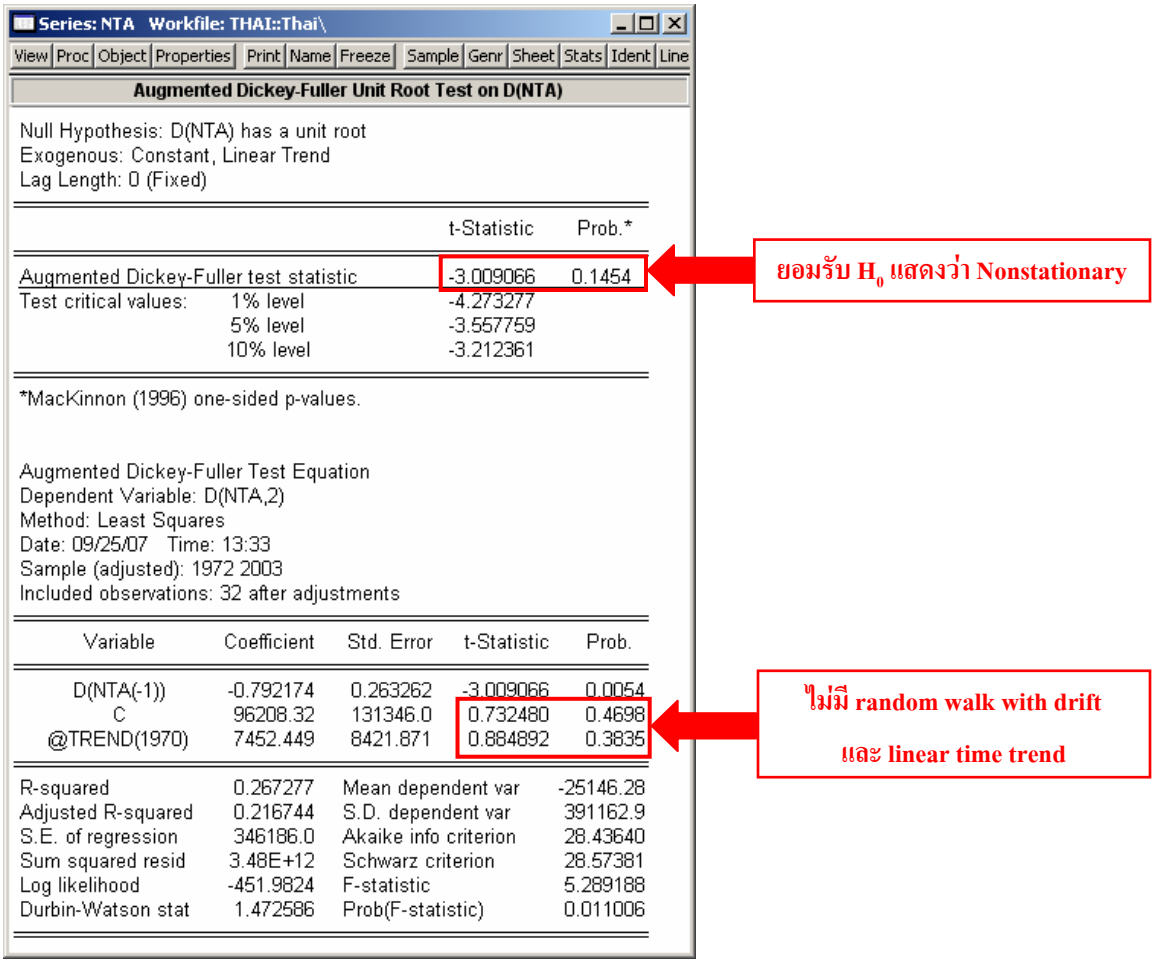
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(NTA)
Method: Least Squares
Date: 09/25/07 Time: 13:26
Sample (adjusted): 1971 2003
Included observations: 33 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
NTA(-1)	-0.124330	0.070153	-1.772271	0.0865
C	-55609.93	139353.8	-0.399056	0.6927
@TREND(1970)	49840.04	22150.47	2.250066	0.0319

R-squared	0.193229	Mean dependent var	284114.6
Adjusted R-squared	0.139444	S.D. dependent var	353191.7
S.E. of regression	327642.3	Akaike info criterion	28.32374
Sum squared resid	3.22E+12	Schwarz criterion	28.45979
Log likelihood	-464.3417	F-statistic	3.592634
Durbin-Watson stat	1.358963	Prob(F-statistic)	0.039926

ยอมรับ H_0
แสดงว่า Nonstationary

จากผลลัพธ์ข้างต้น แสดงให้เห็นว่า ที่ Level ข้อมูลชุดนี้ไม่คงที่ ดังนั้นจึงต้องทำการทดสอบที่ 1st Difference โดยทำตามขั้นตอนที่ 1 – 3 ใหม่อีกครั้งหนึ่ง แต่ให้เลือก **Test for unit root in: 1st difference** แทนการเลือก **Level** ก็จะได้ผลลัพธ์ใหม่ดังรูป



Series: NTA Workfile: THAI::Thai\

View | Proc | Object | Properties | Print | Name | Freeze | Sample | Genr | Sheet | Stats | Ident | Line

Augmented Dickey-Fuller Unit Root Test on D(NTA)

Null Hypothesis: D(NTA) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Fixed)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.009066	0.1454
Test critical values:	1% level	-4.273277	
	5% level	-3.557759	
	10% level	-3.212361	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(NTA,2)
Method: Least Squares
Date: 09/25/07 Time: 13:33
Sample (adjusted): 1972 2003
Included observations: 32 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(NTA(-1))	-0.792174	0.263262	-3.009066	0.0054
C	96208.32	131346.0	0.732480	0.4698
@TREND(1970)	7452.449	8421.871	0.884892	0.3835

R-squared	0.267277	Mean dependent var	-25146.28
Adjusted R-squared	0.216744	S.D. dependent var	391162.9
S.E. of regression	346186.0	Akaike info criterion	28.43640
Sum squared resid	3.48E+12	Schwarz criterion	28.57381
Log likelihood	-451.9824	F-statistic	5.289188
Durbin-Watson stat	1.472586	Prob(F-statistic)	0.011006

ผลลัพธ์ใหม่แสดงให้เห็นว่า ตัวแปรอนุกรมเวลาตัวนี้ มีลักษณะคงที่เมื่อมีการแปลงข้อมูลด้วยการ 1st Difference แต่เมื่อพิจารณาในระดับรายละเอียด พบว่า ข้อมูลชุดนี้อาจไม่มี random walk with drift และ linear time trend ดังนั้นผู้ใช้อาจทำการทดสอบ Unit Root ใหม่อีกครั้ง โดยการเลือก **Include in equation**: **Intercept** ถ้าไม่มี random walk with drift ให้เลือก **Include in equation: None** ตามลำดับ

ที่ผ่านมาเป็นการทดสอบ Unit Root โดยวิธี Dickey-Fuller Test ซึ่งถ้าหากแบบจำลองที่ใช้ในการทดสอบมีปัญหา Autocorrelation ค่าสถิติที่ได้จะขาดความถูกต้อง ดังนั้นจึงได้มีการเสนอให้ปรับสมการของวิธีการ Dickey-Fuller ใหม่ โดยใส่ตัวแปรล่า (Lag) ของ X ในลำดับที่สูงขึ้น แล้วเรียกวิธีการนี้ว่า Augmented Dickey-Fuller Test การทดสอบ Unit Root โดยวิธีการ Augmented Dickey-Fuller Test มีสมการที่ต้องการทดสอบอยู่ 3 สมการ (At level) คือ

$$\Delta X_t = \gamma X_{t-1} + \sum_{i=1}^{p} \phi \Delta X_{t-i} + \varepsilon_t \qquad \text{(random walk process)}$$

$$\Delta X_t = \alpha + \gamma X_{t-1} + \sum_{i=1}^{p} \phi \Delta X_{t-i} + \varepsilon_t \qquad \text{(random walk with drift)}$$

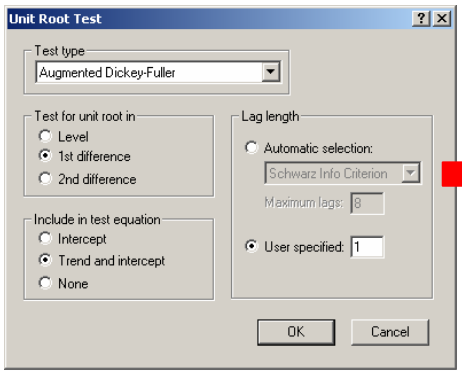
$$\Delta X_t = \alpha + \beta t + \gamma X_{t-1} + \sum_{i=1}^{p} \phi \Delta X_{t-i} + \varepsilon_t \qquad \text{(random walk with drift และมี linear time trend)}$$

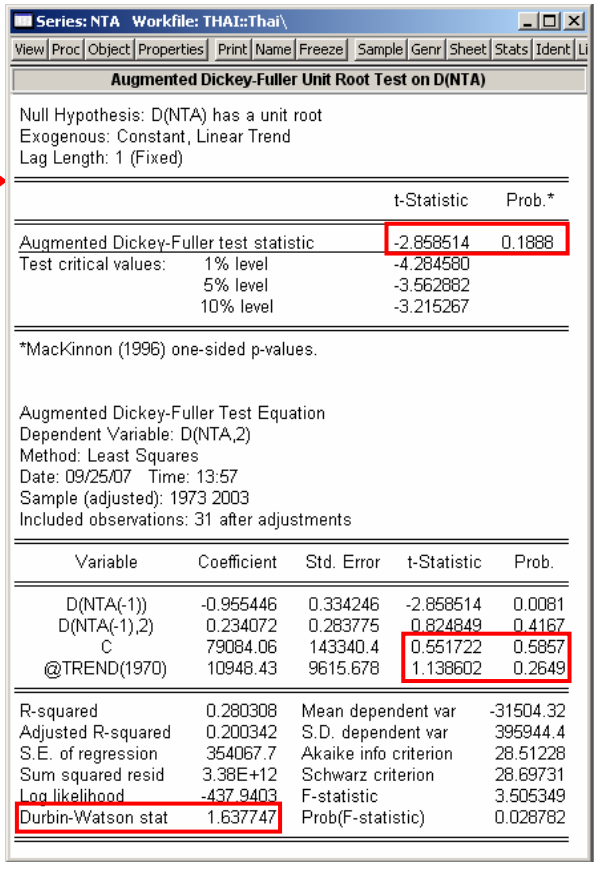
สมมติฐานที่ทดสอบ

$H_0 : \gamma = 0$

$H_a : \gamma \neq 0$

ถ้ายอมรับ H_0 แสดงว่า X_t มีลักษณะไม่นิ่ง (Nonstationary) วิธีการนี้สามารถทำได้ด้วยโปรแกรม EViews โดยการทำตามขั้นตอนที่ 1 – 3 เหมือนกับการทดสอบด้วยวิธี Dickey-Fuller แต่ในขั้นตอนของการกำหนดค่าต่างๆ ที่หน้าต่าง **Unit Root Test** ให้ผู้ใช้กำหนด Lag length อย่างน้อย 1 Lag [โปรแกรม EViews มี Option ให้เลือกกำหนด Lag length ระหว่าง Automatic selection กับ User specified ผู้ใช้สามารถเลือกได้ตามความต้องการ] ถ้าหากค่า D.W. ในผลลัพธ์ของ Unit Root Test แสดงให้เห็นว่ามีปัญหา Autocorrelation ให้ผู้ใช้เพิ่ม Lag ในการทดสอบ Unit Root





จะเห็นได้ว่า ไม่ว่าจะทดสอบด้วยวิธี Correlogram - Q-statistics หรือ Unit Root Test ก็จะให้ผลเหมือนกันว่า ตัวแปรอนุกรมเวลาตัวนี้ มีความไม่คงที่ ซึ่งถ้าหากแปลงข้อมูลด้วยการ 1st Difference แล้ว ข้อมูลที่แปลงใหม่จะคงที่ สรุปได้ว่า ข้อมูลอนุกรมเวลาตัวนี้มีลักษณะไม่คงที่ที่ Level แต่จะคงที่ที่ 1st Difference

ความไม่คงที่ของอนุกรมเวลาอาจเกิดจากการที่อนุกรมเวลานั้นมีแนวโน้มและ/ฤดูกาล หรืออนุกรมเวลามีค่าความแปรปรวนไม่คงที่ ก่อนการกำหนดรูปแบบให้กับอนุกรมเวลาที่ไม่คงที่จะต้องแปลงอนุกรมเวลาดังกล่าวให้คงที่เสียก่อน โดยการหาผลต่างหรือผลต่างของฤดูกาลของอนุกรมเวลา หรือแปลงโดยการหา Natural logarithm ของอนุกรมเวลาที่ไม่คงที่ ดังนี้

ก. การหาผลต่าง

กรณีที่อนุกรมเวลาที่กำลังพิจารณามีแนวโน้มและ/ฤดูกาล การแปลงอนุกรมเวลาดังกล่าว สามารถทำได้โดยการหาผลต่างหรือผลต่างของฤดูกาลของอนุกรมเวลาก่อน ดังนี้

เมื่อกำหนดให้ Δ เป็นผลต่างครั้งที่ 1 Δ^d เป็นผลต่างครั้งที่ d จะได้ว่า

$$\Delta X_t = X_t - X_{t-1}$$

$$\Delta^d X_t = \Delta^{d-1} X_t - \Delta^{d-1} X_{t-1}$$

เนื่องจากในหัวข้อต่อไปจะกล่าวถึงวิธีการพยากรณ์ Box and Jenkins [ARIMA] ซึ่งรูปแบบของ Box and Jenkins มักจะเขียนอยู่ในรูปของอนุกรมเวลาย้อนหลัง (Backward shift operator) ดังนั้นเพื่อความสอดคล้องของเนื้อหาจึงกำหนดสัญลักษณ์ย้อนหลังเป็น B เมื่อกำหนดให้ $BX_t = X_{t-1}$ และ $B^d X_t = X_{t-d}$

จาก $\Delta X_t = X_t - X_{t-1}$

$$\Delta X_t = X_t - BX_t$$

$$\Delta X_t = (1-B)X_t$$

$$\therefore \Delta = 1-B$$

สำหรับการหาผลต่างอนุกรมเวลาที่มีฤดูกาล จะคล้ายกับการหาผลต่างของข้อมูลที่ไม่มีฤดูกาล โดยทั่วไปผลต่างของอนุกรมเวลา ปกติจะคำนวณระยะเวลาหนึ่งไปอีกระยะเวลาหนึ่ง ($X_t - X_{t-1}$) ส่วนผลต่างของฤดูกาลจะหาผลต่างของข้อมูลที่ห่างกัน S หน่วยเวลา ($X_t - X_{t-S}$) จากคำจำกัดความนี้สามารถเขียนให้อยู่ในรูปแบบของอนุกรมเวลาย้อนหลัง (B) และสัญลักษณ์ความแตกต่าง (Δ) ได้ดังนี้

เมื่อกำหนด $B^K X_t = X_{t-k}$ ให้ K = S และ D เป็นอันดับของผลต่างของอนุกรมเวลา X_t ดังนั้นสามารถเขียนผลต่างของอนุกรมเวลา X_t ในอันดับ D ได้ดังนี้

$$\Delta_S^D X_t = (1-B^S)^D X_t$$

เมื่อ D = 1 ดังนั้นผลต่างอันดับที่ 1 ของ X_t ก็คือ

$$\Delta_S X_t = (1-B^S) X_t$$

$$= X_t - B^S X_t$$

$$= X_t - X_{t-S}$$

ส่วนผลต่างของอนุกรมเวลาที่มีฤดูกาลและไม่มีฤดูกาล ส่วนใหญ่ที่นิยมใช้รูปแบบการคูณของอนุกรมเวลาที่มีฤดูกาล – ไม่มีฤดูกาล (seasonal – nonseasonal multiplicative models) ดังนี้

$$\Delta^d \Delta_S^D X_t = (1-B)^d (1-B^s)^D X_t$$

ข. การแปลงอนุกรมเวลา

กรณีที่อนุกรมเวลาที่กำลังพิจารณามีความแปรปรวนไม่คงที่ ต้องแปลงอนุกรมเวลาให้มีความแปรปรวนคงที่ก่อน วิธีการหนึ่งที่ได้รับความนิยม ก็คือ การหาค่า Natural logarithm ของค่าสังเกตในอนุกรมเวลา นั่นคือ แปลงอนุกรมเวลาเดิม (X_t) ให้เป็นอนุกรมเวลาใหม่ (X_t') ซึ่ง $X_t' = \ln(X_t)$

สำหรับคำสั่งในโปรแกรม EViews ที่ใช้ในการแปลงข้อมูลตามวิธีการข้างต้น มีดังนี้

วิธีการแปลงข้อมูล	คำสั่งในโปรแกรม Eviews
$\Delta^d X_t$	d(X,d)
$\Delta^d \Delta_S^D X_t$	d(X,d,s)
$\ln(X_t)$	log(X)
$\Delta^d \ln(X_t)$	dlog(X,d)
$\Delta^d \Delta_S^D \ln(X_t)$	dlog(X,d,s)

➢ ***Granger Causality Tests***

ถ้ามีตัวแปร 2 ตัวแปร มักจะมีคำถามในการวิเคราะห์อยู่เสมอว่าตัวแปรใดเป็นสาเหตุของการเปลี่ยนแปลงของอีกตัวแปรหนึ่ง หรือตัวแปรทั้งสองกำหนดซึ่งกันและกัน หรือต่างก็เป็นตัวแปร Endogenous ในปี ค.ศ. 1969 Prof. Granger ได้นำเสนอตัวทดสอบที่เรียกว่า "***Granger Causality Test***" สำหรับทดสอบในประเด็นดังกล่าว

สมมติว่าเรามีตัวแปรอนุกรมเวลาอยู่ 2 ตัวแปร คือ X และ Y แนวคิดของ Granger ต้องการทดสอบดูว่าการเปลี่ยนแปลงของตัวแปร X เป็นสาเหตุของการเปลี่ยนแปลงของตัวแปร Y หรือว่าการเปลี่ยนแปลงของตัวแปร Y จะเป็นสาเหตุของการเปลี่ยนแปลงของตัวแปร X โดยมีสมมติฐานหลักของการทดสอบทั้งสองกรณี คือ

H_0: X ไม่ได้เป็นสาเหตุของ Y (X does not Granger Cause Y)

H_0: Y ไม่ได้เป็นสาเหตุของ X (Y does not Granger Cause X)

โดยสมการที่ใช้ในการทดสอบสมมติฐาน ก็คือ

$$y_t = \alpha_0 + \alpha_1 y_{t-p} + ... + \alpha_1 y_{t-1} + \beta_1 x_{t-1} + ... + \beta_1 x_{t-p} \quad \text{(Unrestricted regression)}$$

$$y_t = \alpha_0 + \alpha_1 y_{t-p} + ... + \alpha_1 y_{t-1 t-p} \quad \text{(Restricted regression)}$$

หรือ

$$x_t = \alpha_0 + \alpha_1 x_{t-p} + ... + \alpha_1 x_{t-1} + \beta_1 y_{t-1} + ... + \beta_1 y_{t-p} \quad \text{(Unrestricted regression)}$$

$$x_t = \alpha_0 + \alpha_1 x_{t-p} + ... + \alpha_1 x_{t-1} \quad \text{(Restricted regression)}$$

สมมติฐานหลักในเชิงสถิติของการทดสอบสมการแต่ละคู่ระหว่าง Unrestricted regression กับ Restricted regression [การทดสอบมี 2 ชุด คือ X ไม่ได้เป็นสาเหตุของ Y และ Y ไม่ได้เป็นสาเหตุของ X] ก็คือ

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_1 = 0$$

$$H_A : \beta_1 \neq \beta_2 \neq \ldots \neq \beta_1 \neq 0$$

สำหรับสถิติทดสอบ (Test statistic) ได้แก่ สถิติ F (F-statistic) โดยมีสูตรการคำนวณ ดังนี้

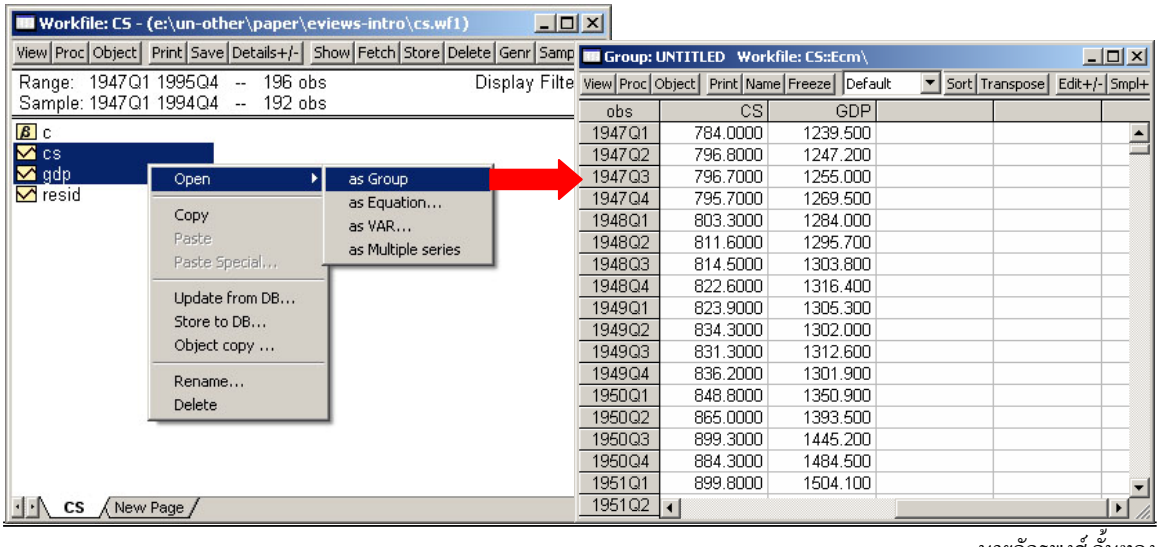
$$F_{p,(n-k)} = \frac{(RSS_r - RSS_{ur})p}{RSS_{ur}/(n-k)}$$

จากสมมติฐานหลักที่ว่า "***H_0: X ไม่ได้เป็นสาเหตุของ Y (X does not Granger Cause Y)***" ถ้าค่า F-statistic ที่คำนวณได้สูงกว่าค่าวิกฤติ [Prob. $< \alpha$] แสดงว่า ปฏิเสธสมมติฐานหลัก (H_0) หมายความว่า X เป็นสาเหตุของการเปลี่ยนแปลงของ Y ในทำนองเดียวกันจากสมมติฐานหลักที่ว่า "***H_0: Y ไม่ได้เป็นสาเหตุของ X (Y does not Granger Cause X)***" ถ้าค่า F-statistic ที่คำนวณได้สูงกว่าค่าวิกฤติ [Prob. $< \alpha$] แสดงว่า ปฏิเสธสมมติฐานหลัก (H_0) หมายความว่า X เป็นสาเหตุของการเปลี่ยนแปลงของ Y

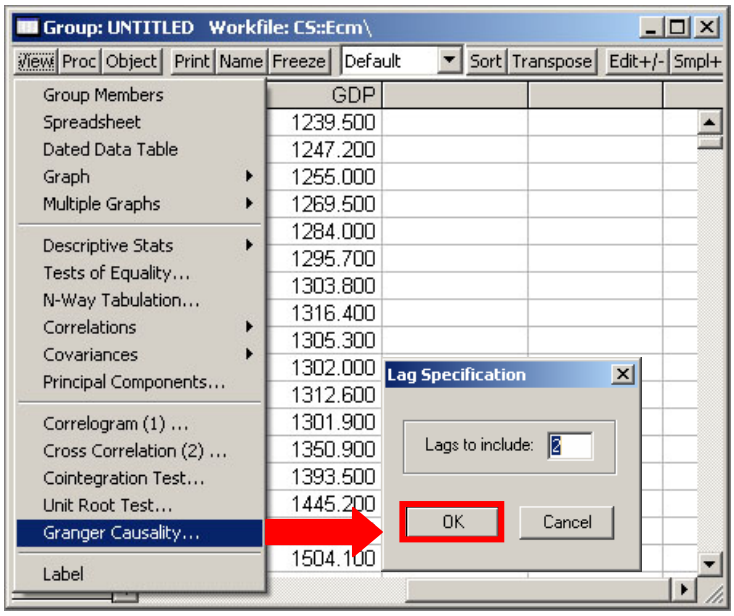
จากแบบจำลองข้างต้น มีประเด็นคำถามว่า "เราควรจะใส่จำนวน Lag เท่าไหร่" เนื่องจากว่าในแบบจำลองข้างต้นผู้ใช้สามารถใส่จำนวน Lag ได้ตามความต้องการ โดยทั่วไปจะทำการทดสอบโดยใช้ค่า Lag ที่แตกต่างกัน 2 – 3 ค่า เพื่อให้แน่ใจว่าผลลัพธ์ที่ได้จากการทดสอบไม่อ่อนไหวต่อการเปลี่ยนแปลงของค่า Lag

โปรแกรม EViews มีคำสั่งสำเร็จรูปในการทดสอบ Granger Causality ดังนั้นเพื่อความเข้าใจในการใช้โปรแกรม EViews ในการทดสอบ Granger Causality จึงขอให้ตัวอย่างที่โปรแกรม EViews ให้มาเป็นตัวสาธิต ซึ่งสามารถเปิดดูได้ที่ c:\programfile\EViews5\Example\data\cs.wf1 หรือ cs.wf1 ในตัวอย่างนี้ ประกอบด้วยตัวแปร CS_t [PERSONAL CONSUMPTION EXPENDITURES (BIL. 1987$)] และตัวแปร GDP_t [GROSS DOMESTIC PRODUCT (BIL. 1987$)] เป็นข้อมูลรายไตรมาสตั้งแต่ปี 1947: 01 – 1994: 04 รวม 192 Obs. โดยในที่นี้จะทดสอบ Granger Causality ว่า การบริโภคเป็นสาเหตุของรายได้ หรือรายได้เป็นสาเหตุของการบริโภค โดยมีขั้นตอนในการใช้โปรแกรม EViews ในการทดสอบดังกล่าว ดังนี้

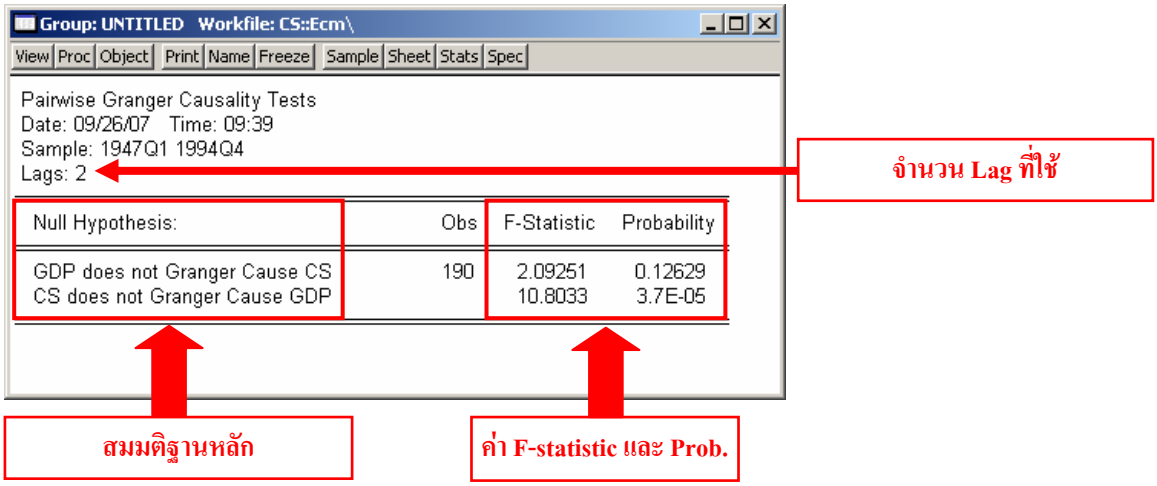
ขั้นตอนที่ 1 เปิด Objects Group ของตัวแปร CS และ GDP ขึ้นมา โดยการเลือกตัวแปรทั้งสอง แล้วคลิก Mouse ด้านขวาเลือก **Open/as Group** โปรแกรม EViews ก็จะเปิดหน้าต่าง **Objects Group** ให้ดังรูป



ขั้นตอนที่ 2 เลือก **View/Granger Causality…** ที่หน้าต่าง **Objects Group** ก็จะปรากฏหน้าต่าง **Lag specification** ให้กำหนดค่า Lag โดยปกติ มีค่า **Default** อยู่ที่ 2 แต่ผู้ใช้สามารถเลือกค่า Lag ได้ตามต้องการ เสร็จแล้วให้กดปุ่ม OK ดังรูป



ขั้นตอนที่ 3 หลังจากกดปุ่ม OK โปรแกรม EViews จะแสดงผลลัพธ์ ดังรูป



จากผลลัพธ์ที่ได้ ไม่สามารถปฏิเสธสมมติฐานหลักที่ว่า "***GDP does not Granger Cause CS***" [เนื่องจากค่า F-statistic ที่คำนวณได้มีค่าต่ำกว่าค่าวิกฤติ ณ ระดับความเชื่อมั่นที่ 99% [Prob. $>$ 0.01]] แต่สามารถปฏิเสธสมมติฐานหลักที่ว่า "***CS does not Granger Cause GDP***" [เนื่องจากค่า F-statistic ที่คำนวณได้มีค่ามากกว่าค่าวิกฤติ ณ ระดับความเชื่อมั่นที่ 99% [Prob. $<$ 0.01]] แสดงว่า ตัวแปร CS เป็นสาเหตุของการเปลี่ยนแปลงของตัวแปร GDP เพียงทิศทางเดียว

ขั้นตอนที่ 4 ดำเนินการตามขั้นตอนที่ 1 – 3 โดยกำหนดค่า Lag ที่แตกต่างกันออกไปอีก 2 – 3 ค่า เพื่อให้แน่ใจว่าผลลัพธ์ที่ได้ไม่อ่อนไหวต่อการเปลี่ยนแปลงของค่า Lag

9.2 วิธีการพยากรณ์

ประโยชน์ของข้อมูลอนุกรมเวลาที่สำคัญ ก็คือ การนำข้อมูลดังกล่าวมาใช้วิเคราะห์เพื่อพยากรณ์ค่าในอนาคต การพยากรณ์ค่าอนุกรมเวลาในอนาคตสามารถทำได้หลายวิธี เช่น การวิเคราะห์แนวโน้ม (Time trend) วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential smoothing), วิธีแยกส่วนประกอบ (Decomposition), วิธีถดถอยเชิงพหุ (Multiple regression) และวิธีบอกซ์ – เจนกินส์ (Box – Jenkins) เป็นต้น สำหรับในคู่มือฉบับนี้จะอธิบายวิธีการวิเคราะห์ 3 วิธี ได้แก่ การวิเคราะห์แนวโน้ม (Time trend), วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential smoothing) และวิธีบอกซ์ – เจนกินส์ (Box – Jenkins) เนื่องจากเป็นวิธีการที่นิยมในปัจจุบัน

- ***วิธีการวิเคราะห์แนวโน้ม (Time trend)***

แบบจำลองที่ใช้ในการวิเคราะห์แนวโน้มโดยทั่วไปมักจะใช้สมการแนวโน้มที่ประมาณค่าสัมประสิทธิ์ด้วยวิธีการกำลังสองน้อยที่สุด (Ordinary Least Square: OLS) ซึ่งเป็นวิธีการที่ทำให้ผลรวมกำลังสองของผลต่างระหว่างค่าแนวโน้มกับข้อมูลที่ใช้มีค่าน้อยที่สุด (Least-square Error) วิธีการนี้เป็นวิธีการหนึ่งที่นิยมนำมาใช้ในการพยากรณ์แนวโน้ม เนื่องจากเป็นวิธีการที่ง่ายในการคำนวณ และสามารถสร้างสมการพยากรณ์ได้หลายรูปแบบทั้งที่เป็นเส้นตรงหรือไม่ใช่เส้นตรงก็ได้ หรือสามารถประยุกต์ใช้วัดอิทธิพลของฤดูกาลได้ด้วย สำหรับวิธีการประมาณค่าพารามิเตอร์ด้วยวิธีการกำลังสองน้อยที่สุด (OLS) สามารถทำได้ดังนี้

$$\hat{y} = \alpha + \beta t$$

โดยที่ $\hat{y}$ คือ ค่าพยากรณ์ของ y

α, β คือ ค่าพารามิเตอร์

t คือ เวลา โดยที่ t = 1, 2, …, n

จากสมการข้างต้นมีค่าพารามิเตอร์ที่ต้องประมาณค่า 2 ตัว คือ ค่า α และ β ในที่นี้จะใช้วิธีการกำลังสองน้อยที่สุดในการประมาณค่าพารามิเตอร์ทั้งสอง โดยมีสูตรคำนวณค่าพารามิเตอร์ทั้งสองดังนี้

$$\beta = \frac{n\sum_{t=0}^{n-1} y_t t - \sum_{t=0}^{n-1} y_t \sum_{t=0}^{n-1} t}{n\sum_{t=0}^{n-1} t^2 - \left(\sum_{t=0}^{n-1} t\right)^2}$$

$$\alpha = \frac{\sum_{t=0}^{n-1} y_t}{n} - \frac{\beta\sum_{t=0}^{n-1} t}{n} = \bar{y}_t - \beta\bar{t}$$

ถ้าจำนวนข้อมูลของอนุกรมเวลามีจำนวนมาก การหาค่า α และ β ด้วยสูตรข้างต้นจะต้องใช้เวลาในการคำนวณมาก ดังนั้นสามารถปรับสูตรข้างต้นให้คำนวณได้ง่ายและรวดเร็วขึ้น โดยการย้ายจุดเริ่มต้น ณ จุดที่ t = 0 เป็นจุดใดจุดหนึ่งที่ทำให้ $\sum t = 0$ แล้วจะได้ว่า

$$\beta = \frac{\sum y_t t}{\sum t^2}$$

$$\alpha = \bar{y}_t$$

วิธีการที่ทำให้ $\sum t = 0$ สามารถทำได้ 2 กรณี ดังนี้

ก. ถ้าจำนวนอนุกรมเวลาเป็นเลขคี่ ให้เวลาที่อยู่กึ่งกลางของอนุกรมเวลามีค่า t = 0 ส่วนเวลาที่อยู่ก่อนเวลากึ่งกลางให้มีค่า t เท่ากับ -1, -2, -3, … สำหรับเวลาที่อยู่หลังเวลากึ่งกลางจะให้มีค่า t เท่ากับ 1, 2, 3, …

ข. ถ้าจำนวนอนุกรมเวลาเป็นเลขคู่ จะให้ t ที่อยู่ระหว่างเวลากึ่งกลางมีค่าเท่ากับ 0 เวลาที่อยู่ก่อนเวลากึ่งกลางให้มีค่า t เท่ากับ -1, -3, -5, …ส่วนเวลาที่อยู่หลังเวลากึ่งกลางให้ค่า t เท่ากับ 1, 3, 5,…

สมการข้างต้นเป็นสมการแนวโน้มเส้นตรง แต่ในบางกรณีเส้นสมการแนวโน้มอาจเป็นเส้นโค้งก็ได้ ดังนั้นการใช้สมการแนวโน้มเส้นตรงในการพยากรณ์ก็อาจทำให้เกิดความคลาดเคลื่อน รูปแบบสมการแนวโน้มที่ไม่เป็นเส้นตรงที่มักจะพบอยู่เสมอ ได้แก่

1. แนวโน้มโพลิโนเมียล (Polynomial trend) มีรูปแบบสมการดังนี้

$$\hat{y} = \alpha_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + ... + \beta_n t^n$$

แนวโน้มโพลิโนเมียลที่มักพบอยู่เสมอ ก็คือ แนวโน้มพาราโบลา (Parabola trend) ดังนี้

$$\hat{y} = \alpha_0 + \beta_1 t + \beta_2 t^2$$

2. แนวโน้มเอ๊กซ์โปเนนเชียล (Exponential trend) มีรูปแบบดังนี้

$$\hat{y} = \alpha_0 + \beta^t$$

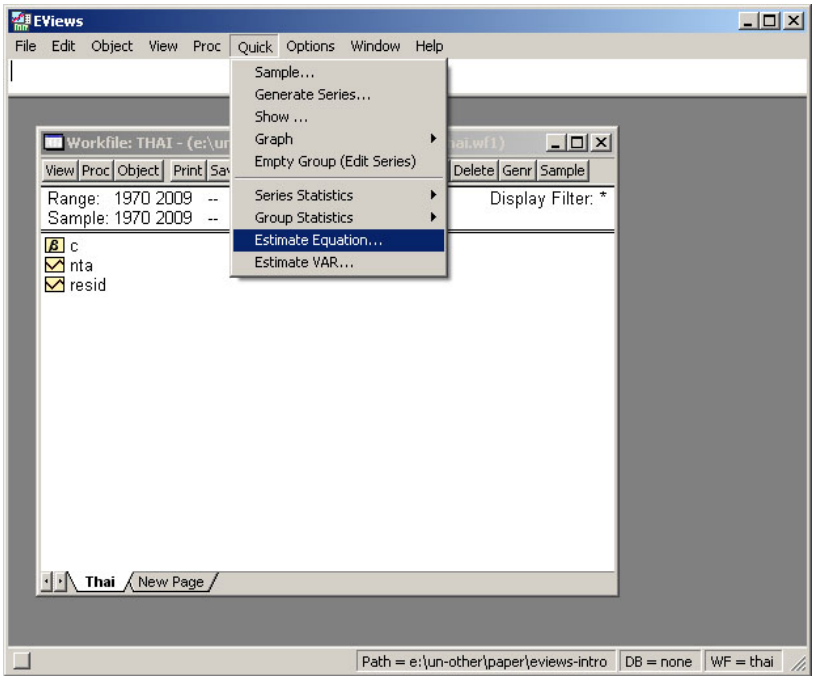
3. แนวโน้มเอ๊กซ์โปเนนเชียลลำดับสอง (Second exponential trend) มีรูปแบบดังนี้

$$\hat{y} = \alpha_0 + \beta_1^t \beta_2^{t^2}$$

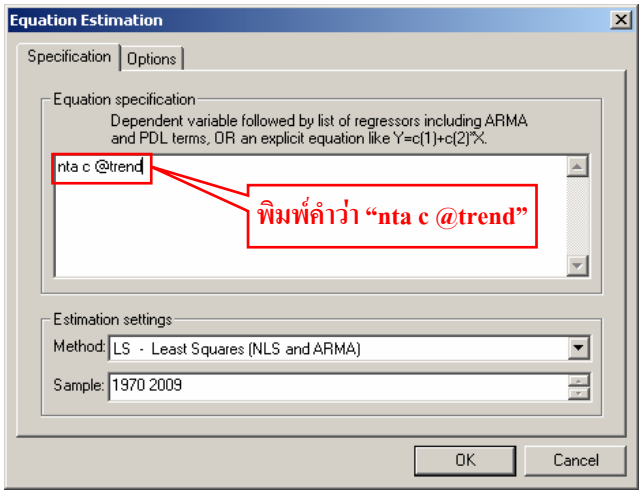
การเลือกรูปแบบของสมการแนวโน้มแบบใดนั้น ขึ้นอยู่กับลักษณะการกระจายของข้อมูลอนุกรมเวลา ดังนั้นก่อนการสร้างสมการพยากรณ์จะต้องทำการทดสอบหรือตรวจสอบว่าตัวแปร/ข้อมูลที่ต้องการพยากรณ์นั้นมีลักษณะของการกระจาย และมีเส้นกราฟแนวโน้มประเภทใด แล้วเลือกรูปแบบของสมการพยากรณ์ให้สอดคล้องกับรูปแบบการกระจายและเส้นกราฟแนวโน้ม เพื่อให้สมการและค่าพยากรณ์ที่ได้จากการคำนวณมีความแม่นยำและคลาดเคลื่อนน้อยที่สุด

โปรแกรม EViews มีคำสั่งสำเร็จรูปในการกำหนดค่าแนวโน้มของเวลา (Time trend) คือ "**@trend**" ส่วนการประมาณค่าสัมประสิทธิ์ของแบบจำลองก็จะใช้วิธีการ OLS สำหรับข้อมูลที่นำมาใช้ในการสาธิตในครั้งนี้เป็นข้อมูลสถิติจำนวนนักท่องเที่ยวชาวต่างชาติที่เดินทางเข้ามาท่องเที่ยวในประเทศไทย ระหว่างปี พ.ศ. 2523 – 2546 จำนวน 33 ตัวอย่าง ส่วนสมการแนวโน้มที่จะทำการทดสอบในครั้งนี้ได้แก่ สมการแนวโน้มแบบเส้นตรง (Linear time trend) และแบบพาราโบลา (Parabola time trend) จากกรณีตัวอย่างสามารถวิเคราะห์โดยใช้โปรแกรม EViews ได้ดังนี้

ขั้นตอนที่1 เลือก **Quick/Estimate Equation** ที่แถบเครื่องมือหลัก ดังรูป



ขั้นตอนที่2 เสร็จแล้ว โปรแกรม EViews จะเปิดหน้าต่าง **Equation Estimation** ให้ผู้ใช้พิมพ์คำสั่งว่า "**nta c @trend**" [nta = ตัวแปรที่ต้องการประมาณค่า, c = ค่าคงที่, @trend = ค่าแนวโน้มเวลา และเป็นสมการแนวโน้มแบบเส้นตรง (Linear time trend)] ลงในช่อง Equation specification ที่หน้าต่าง **Equation Estimation** ดังรูป



ขั้นตอนที่3 กดปุ่ม OK โปรแกรม EViews ก็จะนำเสนอผลลัพธ์ ดังรูป

Equation: UNTITLED Workfile: THAI::Thai\

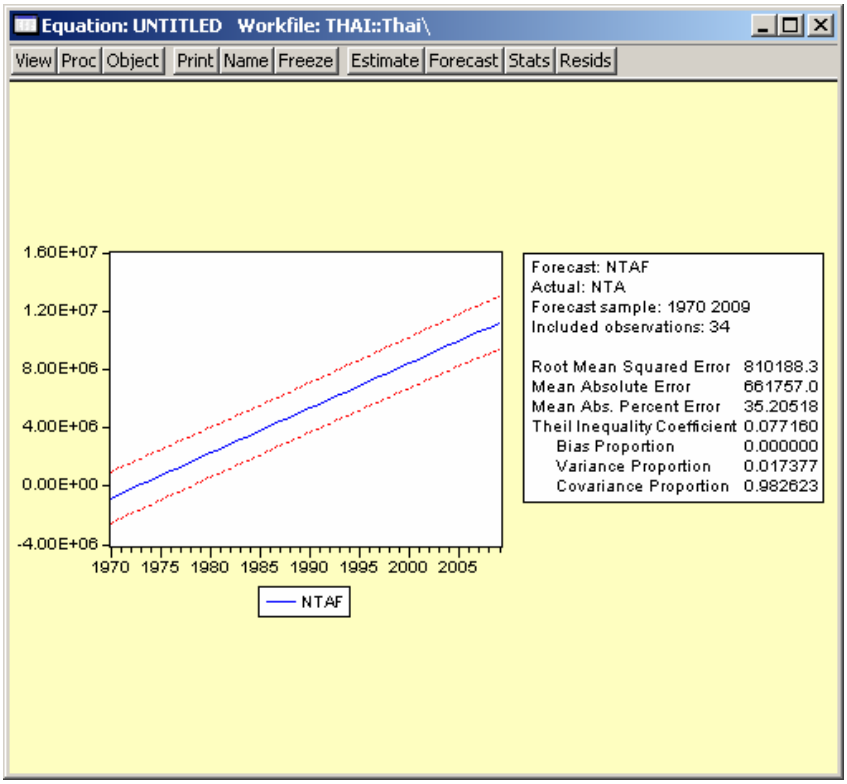
View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: NTA
Method: Least Squares
Date: 09/26/07 Time: 11:00
Sample (adjusted): 1970 2003
Included observations: 34 after adjustments

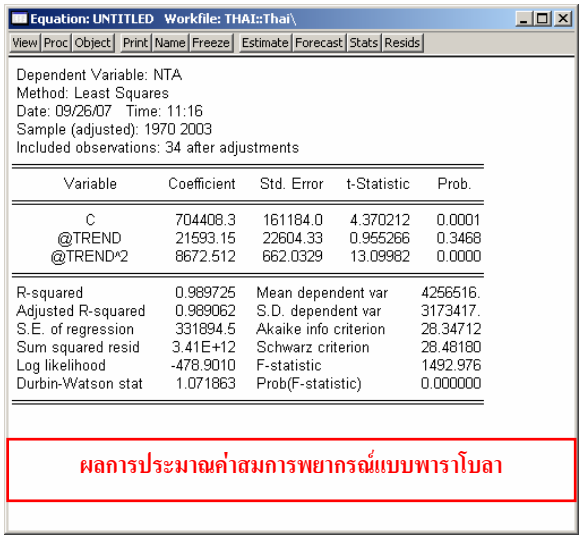
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-821953.9	280239.5	-2.933041	0.0062
@TREND	307786.1	14598.58	21.08329	0.0000

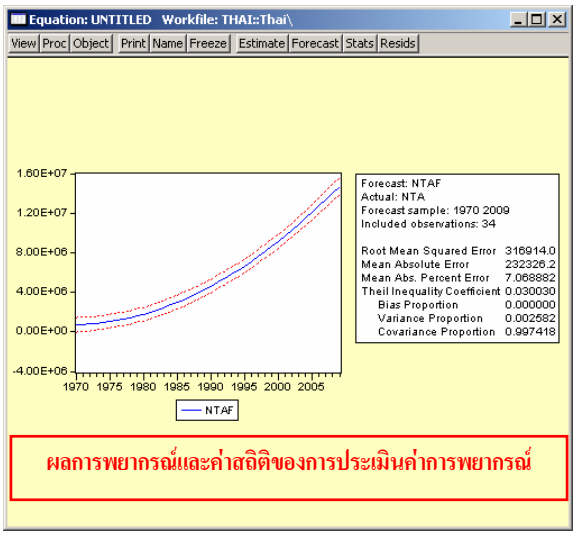
R-squared	0.932844	Mean dependent var	4256516.
Adjusted R-squared	0.930746	S.D. dependent var	3173417.
S.E. of regression	835122.9	Akaike info criterion	30.16557
Sum squared resid	2.23E+13	Schwarz criterion	30.25535
Log likelihood	-510.8147	F-statistic	444.5050
Durbin-Watson stat	0.179691	Prob(F-statistic)	0.000000

ขั้นตอนที่4 ถ้าหากผู้ใช้ต้องการให้โปรแกรม EViews พยากรณ์ ให้ทำตามขั้นตอนในหัวข้อ “การสร้างค่าพยากรณ์ หน้าที่ 30 ของคู่มือฉบับนี้” โปรแกรม EViews จะทำการพยากรณ์ และเสนอค่าสถิติต่างๆ ดังรูป



ขั้นตอนที่5 ถ้าต้องการสร้างสมการพยากรณ์แบบพาราโบลา (Parabola trend) ให้พิมพ์คำสั่งว่า "**nta c @trend @trend^2**" [@trend^2 = ค่าแนวโน้มเวลายกกำลังสอง] ลงในช่อง Equation specification ที่หน้าต่าง **Equation Estimation** เสร็จแล้วก็ทำตามขั้นตอนที่ 3 – 4 ก็จะได้สมการพยากรณ์และค่าพยากรณ์ของสมการพยากรณ์แบบพาราโบลา ดังรูป





➢ *วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential smoothing)*

วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential smoothing) เป็นวิธีการพยากรณ์ที่เหมาะกับการพยากรณ์ในระยะสั้นและปานกลาง วิธีการนี้ได้ให้ความสำคัญกับข้อมูลล่าสุด และความสำคัญของข้อมูลที่ห่างออกไปจะลดลง วิธีการ Exponential smoothing ที่นิยมในปัจจุบันมีอยู่ 3 วิธีการ ดังนี้

ก. Single Exponential Smoothing (SES)

เป็นวิธีการทำให้เรียบอย่างง่าย โดยใช้การหาค่าเฉลี่ยแบบถ่วงน้ำหนัก และสมมติให้ค่าน้ำหนักหรือค่าความสำคัญของข้อมูล คือ α (Alpha) วิธีการนี้มีเงื่อนไขว่า ข้อมูลอนุกรมเวลาที่ใช้ในการวิเคราะห์จะต้องไม่มีแนวโน้ม (Trend) และอิทธิพลของฤดูกาล (Seasonality) หมายความว่า ข้อมูลอนุกรมเวลาที่นำมาใช้ในการวิเคราะห์ต้องมีลักษณะคงที่ สำหรับสมการที่ใช้ในการพยากรณ์สามารถแสดงได้ดังนี้

$$\hat{y}_t = \alpha y_t + (1+\alpha)\hat{y}_{t-1} \qquad ; t = 1, 2, \ldots, n$$

หรือ

$$\hat{y}_t = \alpha \sum_{n=0}^{t-1} (1+\alpha)^n y_{t-n}$$

โดยที่ y_t = ข้อมูล ณ เวลาที่ t; t = 1, 2, …, n

α = ค่าน้ำหนักความสำคัญที่ให้แก่ข้อมูล ณ เวลาที่ t ($0 \le \alpha \le 1$)

$\hat{y}_t$ = ค่าประมาณหรือค่าพยากรณ์ของข้อมูล ณ เวลาที่ t

$\hat{y}_{t-1}$ = ค่าประมาณหรือค่าพยากรณ์ของข้อมูล ณ เวลาที่ t-1

ข. Holt's Two-Parameter Method

จากเงื่อนไขของวิธีการ Single Exponential Smoothing (SES) ที่ว่า ข้อมูลอนุกรมเวลาที่ใช้ในการวิเคราะห์ต้องมีลักษณะคงที่ ซึ่งในความเป็นจริงนั้นข้อมูลอนุกรมเวลาโดยส่วนใหญ่จะมีแนวโน้มของเวลา ดังนั้น Holt (1957) จึงได้ปรับปรุงวิธีการ Single Exponential Smoothing ใหม่ โดยให้สามารถใช้กับข้อมูลอนุกรมเวลาที่มีแนวโน้ม (Trend) ของเวลา เรียกวิธีการนี้ว่า "***Holt's Two-Parameter Method***" วิธีการนี้ให้ความสำคัญกับข้อมูลล่าสุดและแนวโน้มเวลา ดังนั้นจึงมีค่าคงที่ในการทำให้เรียบ 2 ค่า คือ α (Alpha) และ β (Beta) โดยมีสมการที่ใช้ในการพยากรณ์ดังนี้

$$\hat{y}_{t+k} = a + bk \quad ; \hat{y}_{t+k} = \text{ค่าพยากรณ์ ณ เวลาที่ } t+k$$

ค่า a และ b คำนวณจาก

$$a_t = \alpha y_t + (1+\alpha)(a_{t-1} + b_{t-1})$$

$$b_t = \beta(a_t - a_{t-1}) + 1 - \beta b_{t-1}$$

โดยที่ α = ค่าคงที่ที่ทำให้เรียบระหว่างข้อมูลกับค่าพยากรณ์ $(0 \le \alpha \le 1)$

β = ค่าคงที่ที่ทำให้เรียบระหว่างแนวโน้มจริงกับค่าประมาณของแนวโน้ม $(0 \le \beta \le 1)$

ค. Holt-Winters-Trend and Seasonal (Three-Parameter)

วิธีนี้ถูกพัฒนาเพิ่มขึ้นจากวิธีการของ Holt โดย Winters (1960) ได้พัฒนาให้วิธีการนี้สามารถวิเคราะห์ข้อมูลอนุกรมเวลาที่มีทั้งแนวโน้ม (Trend) และฤดูกาล (Seasonality) วิธีการนี้ให้ความสำคัญกับข้อมูลล่าสุด แนวโน้มเวลาและฤดูกาล ดังนั้นจึงมีค่าคงที่ในการทำให้เรียบ 3 ค่า คือ α (Alpha), β (Beta) และ γ (Gamma) โดยมีสมการที่ใช้ในการพยากรณ์ดังนี้

- ***Holt-Wintes-Mulitplicative***

$$\hat{y}_{t+k} = (a + bk)c_{t+k} \quad ; \hat{y}_{t+k} = \text{ค่าพยากรณ์ ณ เวลาที่ } t+k$$

ค่า a, b และ c จะคำนวณจาก

$$a_t = \alpha \frac{y_t}{c_{t-s}} + (1+\alpha)(a_{t-1} + b_{t-1})$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}$$

$$c_t = \gamma \frac{y_t}{a_t} + (1-\gamma)c_{t-s}$$

โดยที่ α = ค่าคงที่ที่ทำให้เรียบระหว่างข้อมูลกับค่าพยากรณ์ $(0 \le \alpha \le 1)$

β = ค่าคงที่ที่ทำให้เรียบระหว่างแนวโน้มจริงกับค่าประมาณของแนวโน้ม $(0 \le \beta \le 1)$

γ = ค่าคงที่ที่ทำให้เรียบระหว่างฤดูกาลจริงกับค่าประมาณของฤดูกาล $(0 \le \gamma \le 1)$

- ***Holt-Wintes-Additive***

$$\hat{y}_{t+k} = (a + bk) + c_{t+k} \quad ; \ \hat{y}_{t+k} = \text{ค่าพยากรณ์ ณ เวลาที่ } t+k$$

ค่า a, b และ c จะคำนวณจาก

$$a_t = \alpha(y_t - c_{t-s}) + (1-\alpha)(a_{t-1} + b_{t-1})$$

$$b_t = \beta(a_t - a_{t-1}) + 1 - \beta b_{t-1}$$

$$c_t = \gamma(y_t - a_{t+1}) - \gamma c_{t-s}$$

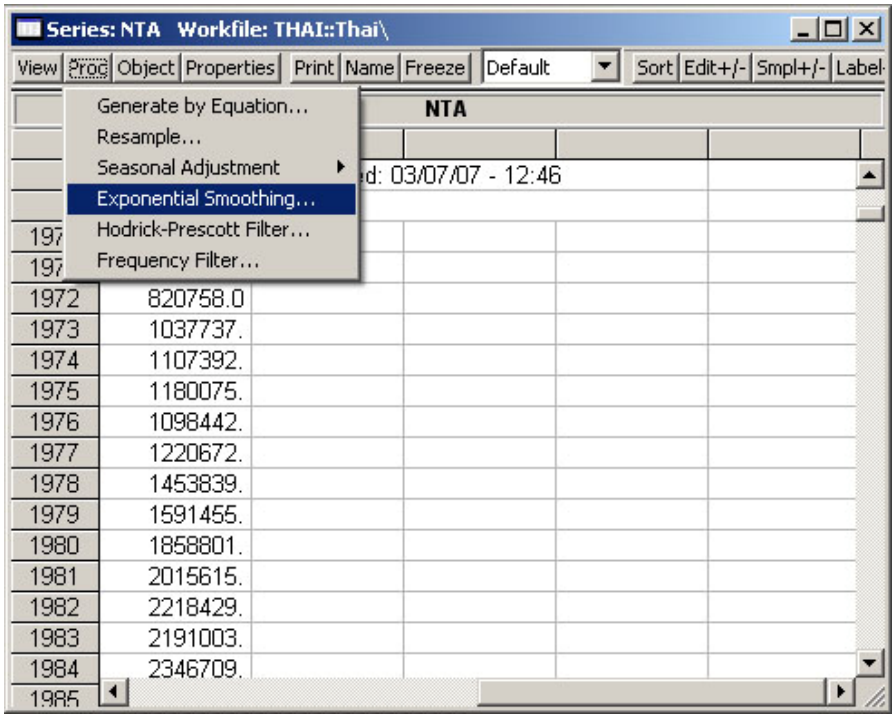
โดยที่ α = ค่าคงที่ที่ทำให้เรียบระหว่างข้อมูลกับค่าพยากรณ์ $(0 \le \alpha \le 1)$

β = ค่าคงที่ที่ทำให้เรียบระหว่างแนวโน้มจริงกับค่าประมาณของแนวโน้ม $(0 \le \beta \le 1)$

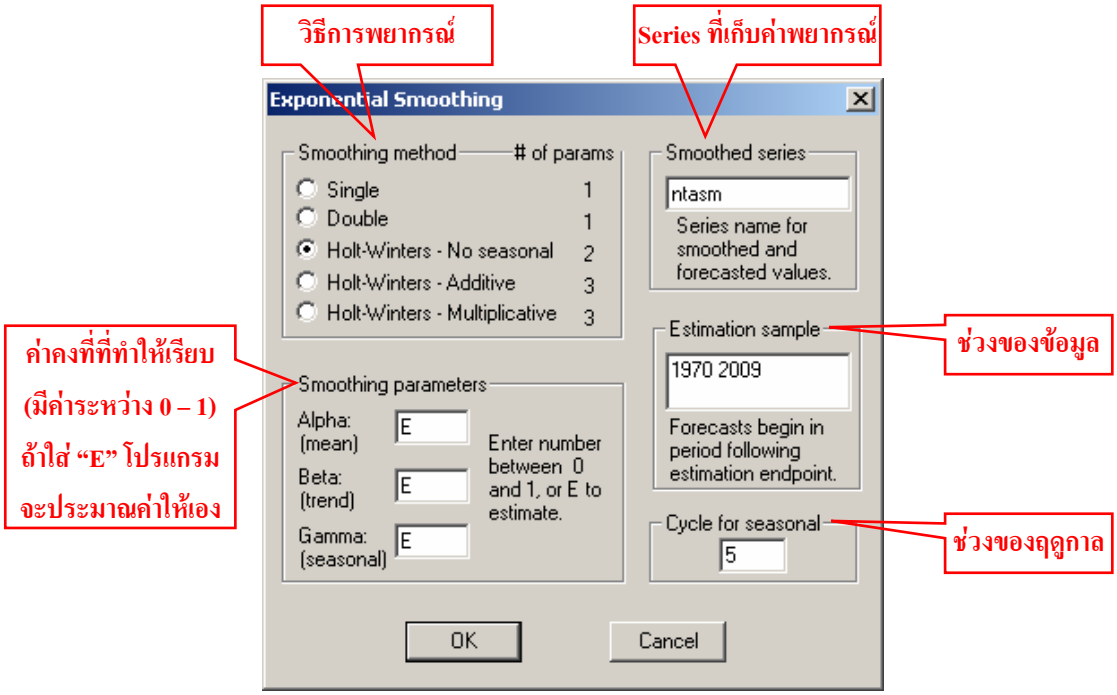
γ = ค่าคงที่ที่ทำให้เรียบระหว่างฤดูกาลจริงกับค่าประมาณของฤดูกาล $(0 \le \gamma \le 1)$

โปรแกรม EViews มีคำสั่งสำเร็จรูปในการพยากรณ์ด้วยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential smoothing) ทั้ง 3 วิธีการ ซึ่งข้อมูลที่นำมาใช้ในการสาธิตในครั้งนี้ เป็นข้อมูลสถิติจำนวนนักท่องเที่ยวชาวต่างชาติที่เดินทางเข้ามาท่องเที่ยวในประเทศไทย ระหว่างปี พ.ศ. 2523 – 2546 จำนวน 33 ตัวอย่าง จากกรณีตัวอย่างสามารถวิเคราะห์โดยใช้โปรแกรม EViews ได้ดังนี้

ขั้นตอนที่ 1 เปิดหน้าต่าง Series ของข้อมูลที่ต้องการทดสอบ ด้วยการ Double click ที่ Series ที่ต้องการ [Double click nta] เสร็จแล้ว ให้เลือก **Proc/Exponential Smoothing …** ที่แถบเครื่องมือของหน้าต่าง **Series** ดังรูป

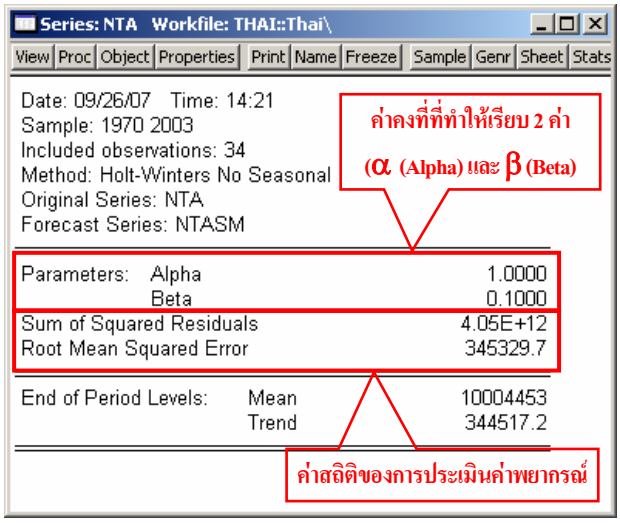


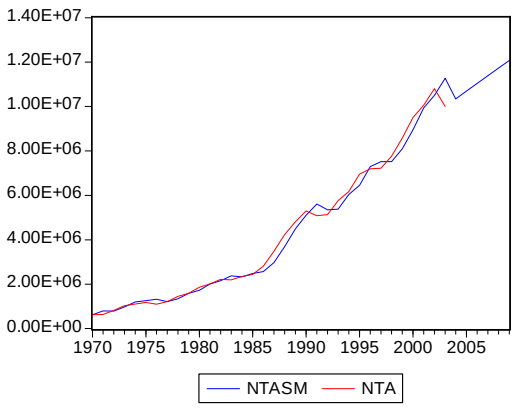
ขั้นตอนที่ 2 โปรแกรม EViews จะเปิดหน้าต่าง **Exponential Smoothing** ขึ้นมา ให้ผู้ใช้กำหนดค่าต่างๆ ตามที่ต้องการ ดังรูป



กรณีตัวอย่าง เป็นข้อมูลอนุกรมเวลาที่มีแนวโน้ม (Trend) แต่ไม่มีฤดูกาล (Seasonality) ดังนั้นจึงเลือกใช้วิธีการ Holt-Winters – No Seasonal [Holt-Winters - No seasonal 2] ซึ่งจะต้องกำหนดค่าคงที่ที่ทำให้เรียบจำนวน 2 ค่า คือ α (Alpha) และ β (Beta) [ในที่นี้ได้กำหนดเป็นค่า "**E**" หมายความว่า ให้โปรแกรมประมาณค่าที่เหมาะสมให้] เนื่องจากไม่มีฤดูกาลดังนั้นในช่อง Cycle for seasonal จึงไม่มีการเปลี่ยนแปลงใดๆ **[Default = 5]** ดังรูป

ขั้นตอนที่ 3 กดปุ่ม OK โปรแกรม EViews นำเสนอผลลัพธ์ของการพยากรณ์ พร้อมทั้งสร้าง Series ค่าพยากรณ์ไว้ที่ Series ชื่อ **ntasm** [ntasm] ผลลัพธ์ที่โปรแกรม EViews นำเสนอประกอบด้วยค่าคงที่ที่ทำให้เรียบ และค่าสถิติของการประเมินค่าพยากรณ์ ดังรูป





➢ ***วิธีการ Box and Jenkins (ARIMA)***

วิธีการ Box and Jenkins เป็นวิธีการพยากรณ์ค่าในอนาคตที่พัฒนาและเสนอโดยนักสถิติผู้มีชื่อเสียงสองท่านคือ George E.P. Box และ Gwilym M. Jenkins ในปี ค.ศ. 1970 วิธีนี้เป็นวิธีที่ให้ค่าพยากรณ์ในระยะสั้นที่ดี คือ มีค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (Mean Square Error: MSE) ของการพยากรณ์ต่ำกว่าวิธีอื่น เหมาะสำหรับการพยากรณ์ไปข้างหน้าในช่วงเวลาสั้นๆ และต้องมีอนุกรมเวลาที่ยาวพอสมควร

แบบจำลองที่ใช้ในการพยากรณ์ คือ ตัวแบบ ARIMA (p,d,q) ซึ่งมีส่วนประกอบที่สำคัญ 3 ส่วน ได้แก่ AutoRegressive AR : (p), Integrated (I) และ Moving Average MA : (q) สำหรับ AR (p) เป็นรูปแบบที่แสดงว่าค่าสังเกต y_t ขึ้นอยู่กับค่าของ $y_{t-1},\ldots,y_{t-p}$ หรือค่าสังเกตที่เกิดขึ้นก่อนหน้า p ค่า ส่วนรูปแบบ MA (q) เป็นรูปแบบที่แสดงว่าค่าสังเกต y_t ขึ้นอยู่กับค่าความคลาดเคลื่อน $\varepsilon_{t-1},\ldots,\varepsilon_{t-q}$ หรือความคลาดเคลื่อนที่อยู่ก่อนหน้า q ค่า ส่วน Integrated (I) เป็นการหาผลต่าง (Difference) ของอนุกรมเวลา เหตุผลสำคัญที่ต้องหาผลต่างของอนุกรมเวลา เนื่องจากแบบจำลอง ARIMA จะต้องใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีคุณสมบัติคงที่ (Stationary) เท่านั้น ในกรณีที่ข้อมูลอนุกรมเวลาที่ใช้ในการวิเคราะห์มีคุณสมบัติไม่คงที่ (Nonstationary) จะต้องทำการแปลงข้อมูลอนุกรมเวลาดังกล่าวให้มีคุณสมบัติคงที่ก่อน โดยการหาผลต่างของข้อมูลอนุกรมเวลา หรือการหาค่า Natural logarithm ของอนุกรมเวลาก่อนที่จะนำข้อมูลไปใช้สร้างแบบจำลอง ARIMA

รูปแบบทั่วไปของ ARIMA ที่ใช้ในการประมาณการ คือ

$$\phi(B)\nabla^{d}y_{t} = \delta + \theta(B)\varepsilon_{t}$$

โดยที่

$$\phi(B) = 1 - \phi_{1}B - \phi_{2}B^{2} - \ldots - \phi_{P}B^{p}$$

$$\theta(B) = 1 - \theta_{1}B - \theta_{2}B^{2} - \ldots - \theta_{P}B^{q}$$

y_t = ค่าสังเกตในอนุกรมเวลา ณ เวลา t

B = Backward shift operation โดยที่ $B_m = \nabla y_{t-m}$

d = จำนวนครั้งของการหาผลต่างเพื่อให้อนุกรมเวลามีคุณสมบัติคงที่ (stationary)

p = อันดับของออโตรีเกรสซีฟ (Autoregressive Order)

q = อันดับของค่าเฉลี่ยเคลื่อนที่ (Moving Average)

δ = ค่าคงที่ (Constant Term)

$\phi_1,\ldots,\phi_p$ = พารามิเตอร์ของ ออโตรีเกรสซีฟ (Autoregressive parameter)

$\theta_1,\ldots,\theta_q$ = พารามิเตอร์ของ ค่าเฉลี่ยเคลื่อนที่ (Moving–Average parameter)

ε_t = กระบวนการ white noise ซึ่งก็คือ ค่าความคลาดเคลื่อน ณ เวลา t ภายใต้ข้อสมมติว่าความคลาดเคลื่อนที่คนละเวลาเป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน โดยมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็นศูนย์และความแปรปรวนคงที่ [$\varepsilon_t \sim N(0,\sigma^2)$]

จากสมการข้างต้นอาจเขียนใหม่ได้เป็น

$$\nabla^d y_t = \delta + \phi\nabla^d y_{t-1} + \phi\nabla^d y_{t-2} + \ldots + \phi\nabla^d y_{t-p} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \ldots - \theta_q\varepsilon_{t-q}$$

จากรูปแบบทั่วไปตามสมการข้างต้นนำไปใช้ในการกำหนดรูปแบบที่เหมาะสมและประมาณค่าต่อไป ซึ่งอนุกรมเวลาที่จะนำมาวิเคราะห์ ด้วยวิธีของ Box and Jenkins นี้ ต้องมีเงื่อนไขบางประการเกี่ยวกับค่าพารามิเตอร์ในตัวแบบเพื่อให้อนุกรมเวลามีคุณสมบัติคงที่ (Stationary) และคุณสมบัติผกผัน (Invertibility) สำหรับคุณสมบัติคงที่ (Stationary) เป็นคุณสมบัติของรูปแบบ AR (p) ซึ่งเป็นคุณสมบัติที่ทำให้ $E(y_t)$ และ $V(y_t)$ คงที่ และ Cov $(y_t,..,y_{t-k})$ มีค่าคงที่ ขึ้นกับค่า Lag k อย่างเดียว ส่วนคุณสมบัติผกผัน (Invertible) เป็นคุณสมบัติของรูปแบบ MA (q) ซึ่งเป็นคุณสมบัติที่ทำให้ค่าคลาดเคลื่อนของการพยากรณ์ ε_t ในเทอมของ y_t , y_{t-1} มีค่าคงที่ (ทรงศิริ แต้สมบัติ, 2539) มีขั้นตอนของวิธีการ Box and Jenkins

ก. **กำหนดรูปแบบ (Identification)** เพื่อหารูปแบบที่คาดว่าเหมาะสมให้กับอนุกรมเวลา โดยใช้วิธีพิจารณาเปรียบเทียบจาก Correlogram ของค่า r_k และ r_{kk} ของอนุกรมเวลา

ข. **การประมาณค่าสัมประสิทธิ์ (Estimation)** ในรูปแบบ โดยทั่วไปใช้วิธีการประมาณค่าสัมประสิทธิ์ด้วยวิธีกำลังสองน้อยที่สุดแบบธรรมดา (Ordinary Least Square Method: OLS)

ค. **การตรวจสอบรูปแบบ (Diagnostic checking)** เมื่อกำหนดรูปแบบและประมาณค่าพารามิเตอร์ในรูปแบบแล้ว ต้องตรวจสอบอีกครั้งว่ารูปแบบที่กำหนดมีความเหมาะสมจริงหรือไม่ โดยการพิจารณาค่าสหสัมพันธ์ในตัวเองของค่าความคลาดเคลื่อน (ดูจากกราฟ Correlogram) การทดสอบค่าพารามิเตอร์ในรูปแบบ โดยการพิจารณาจากค่าสถิติ t (t – statistic) และการทดสอบความเหมาะสมของรูปแบบโดยการทดสอบของ Box and Pierce หรือการทดสอบของ Box and Ljung [Q-statistic]

ง. **การพยากรณ์ (Forecasting)** นำสมการพยากรณ์ที่สร้างจากรูปแบบการพยากรณ์ที่กำหนดและผ่านการตรวจสอบรูปแบบ มาพยากรณ์ค่าในอนาคต โดยสามารถทำได้ทั้งการพยากรณ์แบบจุด (Point forecast) และการพยากรณ์แบบช่วง (Interval forecast) การพยากรณ์โดยวิธีการของ Box and Jenkins จะให้ค่าพยากรณ์ไปข้างหน้าที่ดีในช่วงเวลาสั้นๆ

โปรแกรม EViews ไม่มีคำสั่งสำเร็จรูปในการพยากรณ์ด้วยวิธีการ Box and Jenkins (ARIMA) แต่สามารถใช้โปรแกรม EViews ในการพยากรณ์ตามขั้นตอนของวิธีการดังกล่าวได้ เนื่องจากโปรแกรม EViews มีคำสั่งสำเร็จรูปในการพยากรณ์ตามขั้นตอนต่างๆ ของวิธีการ Box and Jenkins (ARIMA) ทั้ง 4 ขั้นตอน สำหรับข้อมูลที่นำมาใช้ในการสาธิตเป็นข้อมูลสถิติจำนวนนักท่องเที่ยวชาวต่างชาติที่เดินทางเข้ามาท่องเที่ยวในประเทศไทย ระหว่างปี พ.ศ. 2523 – 2546 จำนวน 33 ตัวอย่าง จากกรณีตัวอย่างสามารถพยากรณ์แบบจำลอง ARIMA โดยใช้โปรแกรม EViews ได้ดังนี้

ก. การกำหนดรูปแบบ (Identification)

ขั้นตอนที่ 1 ทดสอบ Unit Root ของข้อมูลที่ใช้ในการวิเคราะห์ว่า มีลักษณะคงที่ที่เท่าไหร่ [ทำตามขั้นตอนการทดสอบ Unit Root ของคู่มือฉบับนี้ หน้าที่ 67 – 70] จากกรณีตัวอย่าง จะได้ผลการทดสอบดังนี้

Null Hypothesis: NTA has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Fixed)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.116828	0.5173
Test critical values:	1% level	-4.273277	
	5% level	-3.557759	
	10% level	-3.212361	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(NTA)
Method: Least Squares
Date: 03/06/07 Time: 15:36
Sample (adjusted): 1972 2003
Included observations: 32 after adjustments

	Coefficient	Std. Error	t-Statistic	Prob.
NTA(-1)	-0.163537	0.077256	-2.116828	0.0433
D(NTA(-1))	0.366902	0.259859	1.411928	0.1690
C	-108518.5	157342.0	-0.689698	0.4961
@TREND(1970)	55420.00	24016.81	2.307550	0.0286

R-squared	0.234607	Mean dependent var	292678.6
Adjusted R-squared	0.152601	S.D. dependent var	355344.9
S.E. of regression	327110.2	Akaike info criterion	28.35045
Sum squared resid	3.00E+12	Schwarz criterion	28.53367
Log likelihood	-449.6072	Hannan-Quinn criter.	28.41118
F-statistic	2.860841	Durbin-Watson stat	1.627586
Prob(F-statistic)	0.054676		

ผลการทดสอบ Unit Root ที่ Level (d = 0)

Null Hypothesis: D(NTA) has a unit root
Exogenous: Constant
Lag Length: 1 (Fixed)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.813946	0.0679
Test critical values:	1% level	-3.661661	
	5% level	-2.960411	
	10% level	-2.619160	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(NTA,2)
Method: Least Squares
Date: 03/06/07 Time: 15:44
Sample (adjusted): 1973 2003
Included observations: 31 after adjustments

	Coefficient	Std. Error	t-Statistic	Prob.
D(NTA(-1))	-0.699303	0.248513	-2.813946	0.0089
D(NTA(-1),2)	0.109674	0.263283	0.416562	0.6802
C	195121.7	101330.6	1.925595	0.0644

R-squared	0.245752	Mean dependent var	-31504.32
Adjusted R-squared	0.191877	S.D. dependent var	395944.4
S.E. of regression	355936.9	Akaike info criterion	28.49466
Sum squared resid	3.55E+12	Schwarz criterion	28.63343
Log likelihood	-438.6672	Hannan-Quinn criter.	28.53990
F-statistic	4.561527	Durbin-Watson stat	1.620342
Prob(F-statistic)	0.019284		

ผลการทดสอบ Unit Root ที่ 1st Difference (d = 1)

จากการทดสอบ Unit Root สรุปได้ว่า ข้อมูลที่นำมาใช้มีลักษณะไม่คงที่ และต้องทำ 1st Difference จึงจะทำให้ข้อมูลมีลักษณะคงที่ ดังนั้นแบบจำลอง ARIMA ของข้อมูลชุดนี้จะมี Integrated (I) = 1

ขั้นตอนที่ 2 ดู Correlogram ของสหสัมพันธ์ในตัวเอง [Autocorrelation (ACF)] และสหสัมพันธ์ในตัวเองบางส่วน [Partial Correlation (PACF)] เพื่อกำหนด AutoRegressive AR: (p) และ Moving Average MA: (q) [ทำตามขั้นตอนของคู่มือฉบับนี้ หน้าที่ 65 – 66] จากกรณีตัวอย่าง ได้ผลลัพธ์ดังนี้

Date: 03/06/07 Time: 15:30
Sample: 1970 2009
Included observations: 34

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.920	0.920	31.426	0.000
		2	0.815	-0.209	56.858	0.000
		3	0.716	0.004	77.085	0.000
		4	0.622	-0.032	92.876	0.000
		5	0.540	0.013	105.19	0.000
		6	0.465	-0.026	114.66	0.000
		7	0.389	-0.068	121.51	0.000
		8	0.306	-0.094	125.91	0.000
		9	0.224	-0.036	128.36	0.000
		10	0.151	-0.006	129.53	0.000
		11	0.084	-0.045	129.90	0.000
		12	0.023	-0.031	129.93	0.000

ACF และ PACF ที่ Level (d = 0)

Date: 03/06/07 Time: 15:31
Sample: 1970 2009
Included observations: 33

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.248	0.248	2.2231	0.136
		2	0.079	0.019	2.4584	0.293
		3	-0.174	-0.211	3.6228	0.305
		4	-0.181	-0.102	4.9318	0.294
		5	0.074	0.191	5.1607	0.397
		6	0.278	0.246	8.4769	0.205
		7	0.170	-0.031	9.7624	0.202
		8	-0.110	-0.246	10.323	0.243
		9	-0.040	0.163	10.401	0.319
		10	-0.117	0.049	11.095	0.350
		11	0.084	0.013	11.465	0.405
		12	0.235	0.089	14.501	0.270

ACF และ PACF ที่ 1st Difference (d = 1)

พิจารณา Correlogram ของ ACF และ PACF เพื่อกำหนดรูปแบบของ AR และ MA โดยมีเงื่อนไขของการพิจารณาดังตารางรูปแบบ ARMA(p,q) และเงื่อนไขข้างล่างนี้

ตารางรูปแบบ ARMA(p,q) เงื่อนไข

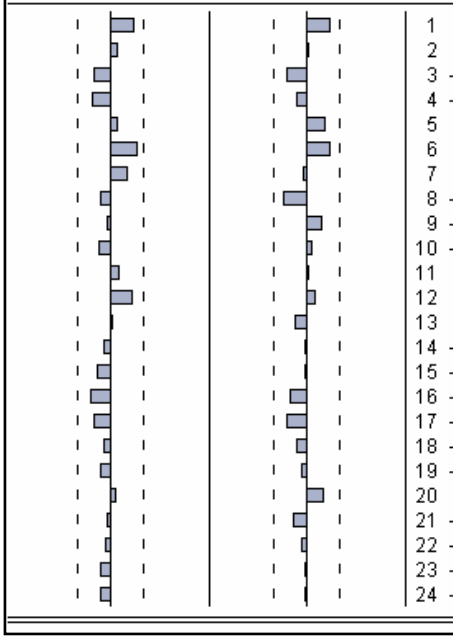
กระบวนการ	ACF	PACF	เงื่อนไข
AR(1)	ลดลงแบบ Exponential	เฉพาะค่า $\phi_{11} \neq 0$ ค่าอื่นๆ เป็น 0	$1 < \phi 1 < 1$
MA(1)	$\rho_1 \neq 0$ ค่าอื่นๆ เป็น 0 และ $-0.5 < \rho_1 < 0.5$	ลดลงในกรอบของการลดแบบ Exponential	$1 < \theta 1 < 1$
AR(2)	ลดลงแบบ Exponential หรือลดลงแบบคลื่นไซน์	เฉพาะ ϕ_{11} และ $\phi_{22} \neq 0$ ค่าอื่นๆ เป็น 0	$\phi_2 + \phi_1 < 1, \phi_2 - \phi_1 < 1,$ $\phi\|_2\| < 1$
MA(2)	เฉพาะ ρ_1 และ $\rho_2 \neq 0$ ค่าอื่นๆ เป็น 0	ลดลงแบบ Exponential หรือลดลงแบบคลื่นไซน์	$\theta_2 + \theta_1 < 1, \theta_2 - \theta_1 < 1,$ $\theta\|_2\| < 1$

ที่มา: ดัดแปลงมาจากดร.มุกดา แม้นมินทร์. 2549. อนุกรมเวลาและการพยากรณ์. หน้า 218.

จาก Correlogram ของ ACF และ PACF ที่ 1st Difference (d = 1) ที่ได้ และเมื่อพิจารณาตามเงื่อนไขในตารางข้างต้น สามารถสรุปได้ว่า รูปแบบ ARIMA ที่เป็นไปได้ของข้อมูลชุดนี้มี 2 รูปแบบ คือ ARIMA (2,1,1) และ ARIMA (2,1,2) ดังนั้นในขั้นตอนต่อไปจะทำการประมาณค่าสัมประสิทธิ์ของรูปแบบทั้งสอง

Date: 03/06/07 Time: 16:57
Sample: 1970 2003
Included observations: 33

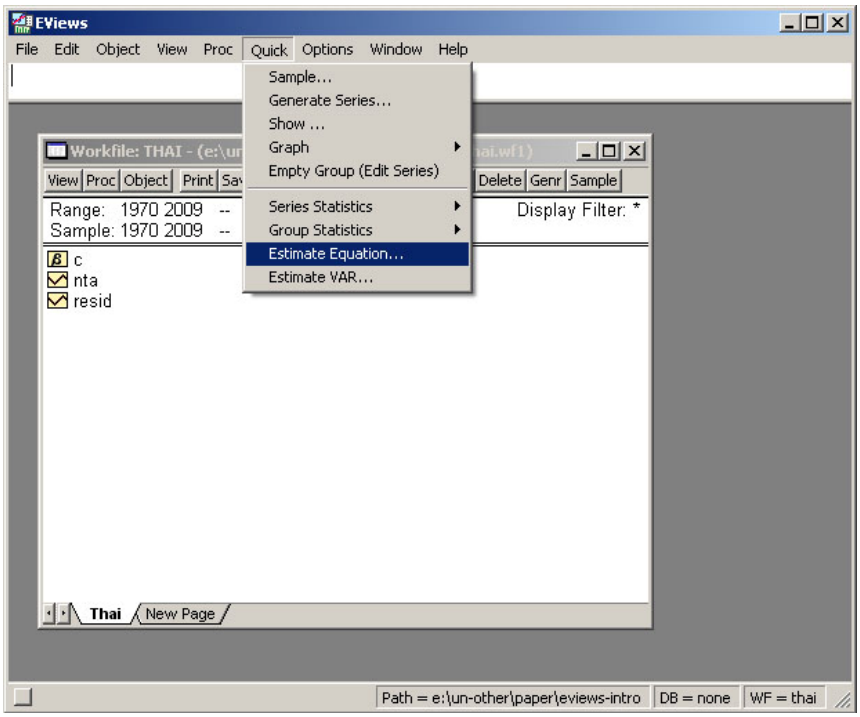
ACF และ PACF ที่ 1st Difference (d = 1)



	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1			0.248	0.248	2.2231	0.136
2			0.079	0.019	2.4584	0.293
3			-0.174	-0.211	3.6228	0.305
4			-0.181	-0.102	4.9318	0.294
5			0.074	0.191	5.1607	0.397
6			0.278	0.246	8.4769	0.205
7			0.170	-0.031	9.7624	0.202
8			-0.110	-0.246	10.323	0.243
9			-0.040	0.163	10.401	0.319
10			-0.117	0.049	11.095	0.350
11			0.084	0.013	11.465	0.405
12			0.235	0.089	14.501	0.270
13			0.014	-0.122	14.512	0.339
14			-0.060	-0.011	14.730	0.397
15			-0.144	-0.014	16.065	0.378
16			-0.204	-0.177	18.895	0.274
17			-0.169	-0.202	20.965	0.228
18			-0.059	-0.110	21.231	0.268
19			-0.101	-0.050	22.081	0.280
20			0.056	0.175	22.364	0.321
21			-0.030	-0.141	22.453	0.374
22			-0.050	-0.042	22.715	0.418
23			-0.111	-0.017	24.134	0.396
24			-0.107	-0.023	25.609	0.373

ข. การประมาณค่าสัมประสิทธิ์ (Estimation)

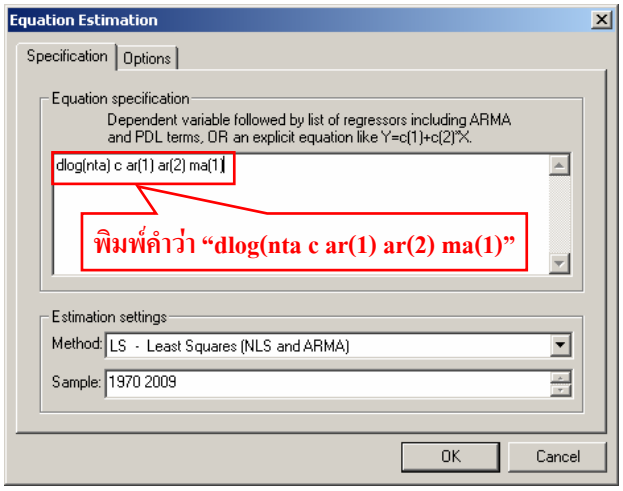
ขั้นตอนที่1 เลือก **Quick/Estimate Equation** ที่แถบเครื่องมือหลัก ดังรูป



ขั้นตอนที่2 เสร็จแล้ว โปรแกรม EViews จะเปิดหน้าต่าง **Equation Estimation** ให้ผู้ใช้พิมพ์คำว่า "**dlog(nta) c ar(1) ar(1) ma(1)**" ลงในช่อง Equation specification ในหน้าต่าง **Equation Estimation** สำหรับแบบจำลอง ARIMA (2,1,1) ดังรูป

คำสั่ง	ความหมาย
dlog(nta)	$\ln(nta_t) - \ln(nta_{t-1})$
ar(#)	AutoRegressive
ma(#)	Moving Average

หมายเหตุ: # จำนวนตัวเลขของ p, q ที่ต้องการ



ขั้นตอนที่ 3 เสร็จแล้วให้กดปุ่ม OK โปรแกรม EViews จะนำเสนอผลลัพธ์ดังรูป

ขั้นตอนที่ 4 ให้ทำขั้นตอนที่ 1 – 3 ซ้ำอีกครั้งหนึ่ง สำหรับแบบจำลอง ARIMA (2,1,2) โดยที่ในช่อง Equation specification ในหน้าต่าง **Equation Estimation** ให้พิมพ์คำสั่งเพิ่มว่า "**ma(2)**" ต่อท้ายคำสั่งเดิม ก็จะได้ผลลัพธ์ของแบบจำลอง ARIMA (2,1,2) ดังรูป

Dependent Variable: D(LOG(NTA))
Method: Least Squares
Date: 03/06/07 Time: 17:32
Sample (adjusted): 1973 2003
Included observations: 31 after adjustments
Convergence achieved after 14 iterations
MA Backcast: 1972

	Coefficient	Std. Error	t-Statistic	Prob.
C	0.078816	0.009604	8.206851	0.0000
AR(1)	0.797921	0.261387	3.052638	0.0050
AR(2)	-0.427187	0.167246	-2.554249	0.0166
MA(1)	-0.540559	0.304442	-1.775574	0.0871

R-squared	0.258839	Mean dependent var	0.080663
Adjusted R-squared	0.176488	S.D. dependent var	0.074651
S.E. of regression	0.067744	Akaike info criterion	-2.426248
Sum squared resid	0.123910	Schwarz criterion	-2.241217
Log likelihood	41.60684	Hannan-Quinn criter.	-2.365933
F-statistic	3.143119	Durbin-Watson stat	1.964683
Prob(F-statistic)	0.041466		

Inverted AR Roots	.40-.52i	.40+.52i
Inverted MA Roots	.54	

ผลลัพธ์ของแบบจำลอง ARIMA (2,1,1)

Dependent Variable: D(LOG(NTA))
Method: Least Squares
Date: 03/06/07 Time: 17:34
Sample (adjusted): 1973 2003
Included observations: 31 after adjustments
Convergence achieved after 24 iterations
MA Backcast: 1971 1972

	Coefficient	Std. Error	t-Statistic	Prob.
C	0.075034	0.017628	4.256646	0.0002
AR(1)	0.679535	0.172027	3.950165	0.0005
AR(2)	-0.572315	0.120602	-4.745497	0.0001
MA(1)	-0.377614	0.105000	-3.596323	0.0013
MA(2)	0.910978	0.067082	13.58001	0.0000

R-squared	0.493772	Mean dependent var	0.080663
Adjusted R-squared	0.415890	S.D. dependent var	0.074651
S.E. of regression	0.057054	Akaike info criterion	-2.742962
Sum squared resid	0.084633	Schwarz criterion	-2.511673
Log likelihood	47.51590	Hannan-Quinn criter.	-2.667567
F-statistic	6.340058	Durbin-Watson stat	2.127757
Prob(F-statistic)	0.001064		

Inverted AR Roots	.34+.68i	.34-.68i
Inverted MA Roots	.19+.94i	.19-.94i

ผลลัพธ์แบบจำลอง ARIMA (2,1,2)

จากผลการประมาณค่าสัมประสิทธิ์ของแบบจำลองทั้งสอง พบว่า แบบจำลอง ARIMA (2,1,2) จะมีค่าสัมประสิทธิ์ในการตัดสินใจดีกว่า [ดูรายละเอียดเรื่องค่าสัมประสิทธิ์ในการตัดสินใจได้ในคู่มือฉบับนี้ หน้า 26] ดังนั้นจึงตัดสินใจเลือกแบบจำลอง ARIMA (2,1,2) มาใช้ในการพยากรณ์

ค. การตรวจสอบรูปแบบ (Diagnostic Checking)

ขั้นตอนที่ 1 เลือก**View/Residual/Correlogram – Q-statistics** ที่แถบเครื่องมือของหน้าต่าง **Objects Equation** ให้ผู้ใช้ทำตามขั้นตอนการตรวจสอบปัญหา Autocorrelation ด้วยการตรวจสอบจาก Correlogram – Q-statistics [ดูรายละเอียดของขั้นตอนได้ในคู่มือฉบับนี้ หน้า 54] หลังจากทำตามขั้นตอนดังกล่าวแล้ว โปรแกรม EViews จะเสนอ Correlogram of Residuals พร้อมทั้งค่า Q-statistics ดังรูป

Date: 03/06/07 Time: 17:35
Sample: 1973 2003
Included observations: 31
Q-statistic probabilities adjusted for 4 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.108	-0.108	0.4001	
		2	-0.031	-0.043	0.4336	
		3	-0.071	-0.080	0.6167	
		4	-0.232	-0.255	2.6530	
		5	-0.021	-0.097	2.6704	0.102
		6	0.144	0.106	3.5219	0.172
		7	0.127	0.127	4.2063	0.240
		8	-0.073	-0.106	4.4436	0.349
		9	0.157	0.156	5.5913	0.348
		10	-0.205	-0.099	7.6475	0.265
		11	-0.007	0.026	7.6498	0.364
		12	0.226	0.227	10.389	0.239
		13	0.004	0.086	10.390	0.320
		14	-0.180	-0.265	12.333	0.263
		15	0.005	-0.045	12.335	0.339
		16	-0.205	-0.142	15.190	0.231
		17	-0.104	-0.140	15.979	0.250
		18	0.184	-0.052	18.638	0.179
		19	-0.154	-0.272	20.651	0.148
		20	0.154	0.034	22.845	0.118
		21	-0.015	-0.055	22.868	0.154
		22	-0.034	0.056	22.995	0.191
		23	-0.080	0.021	23.816	0.203
		24	0.066	0.015	24.457	0.223

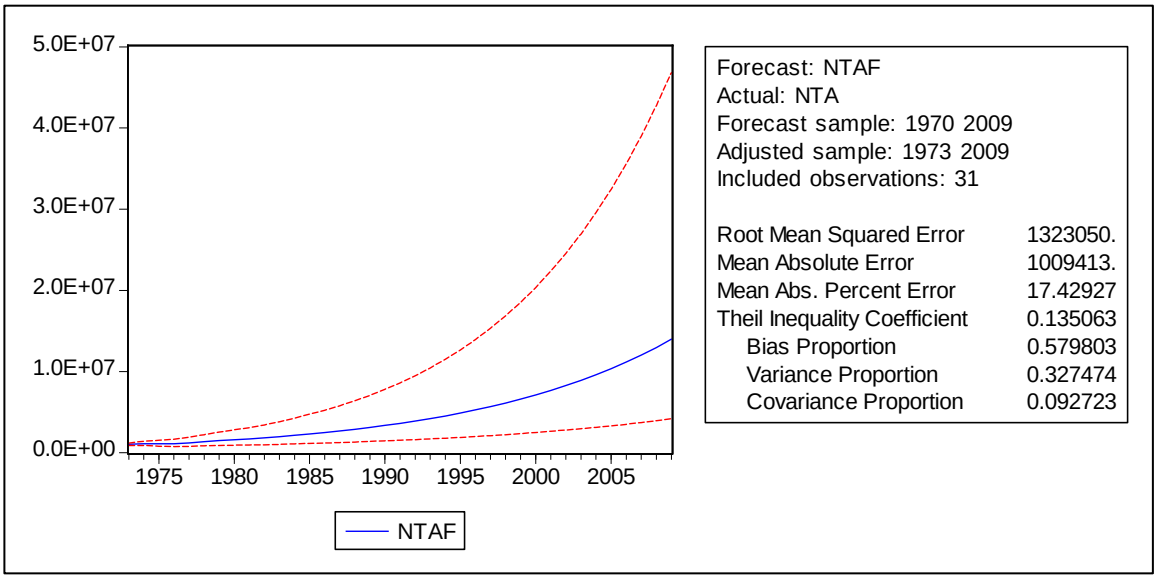
ขั้นตอนที่ 2 พิจารณา Correlogram และ Q-statistics พบว่า Correlogram of Residuals ของ Autocorrelation (ACF) ไม่มีลักษณะลดลงแบบ Exponential ในขณะเดียวกันค่า Q-statistics ที่คำนวณได้มีค่าต่ำกว่าค่าวิกฤตของ Chi-square ณ ระดับนัยสำคัญ 0.10 [Prob. $<$ 0.10] แสดงว่า ไม่มีสหสัมพันธ์ในตัวเองของค่าความคลาดเคลื่อน

จากการตรวจสอบความเหมาะสมของรูปแบบสมการพยากรณ์ สามารถสรุปได้ว่า แบบจำลอง ARIMA (2,1,2) เป็นแบบจำลองพยากรณ์ที่มีประสิทธิภาพและมีความเหมาะสมที่จะนำไปใช้ในการพยากรณ์

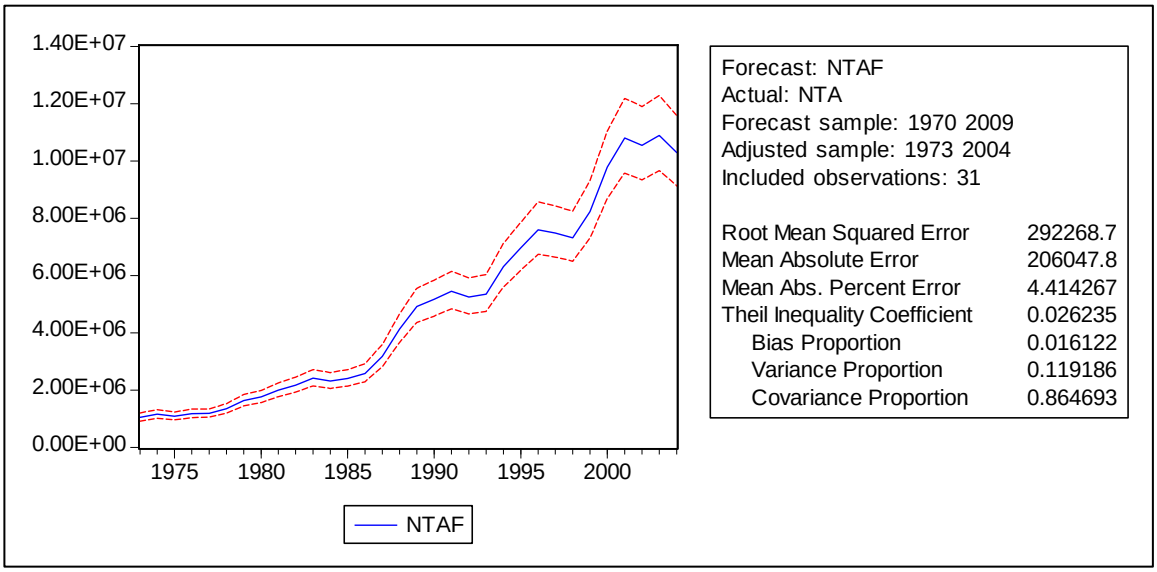
ง. การพยากรณ์ (Forecasting)

เมื่อผู้ใช้ต้องการให้โปรแกรม EViews พยากรณ์ ให้ทำตามขั้นตอนในหัวข้อ “การสร้างค่าพยากรณ์ หน้าที่ 30 ของคู่มือฉบับนี้” ผู้ใช้สามารถเลือกการพยากรณ์ได้ 2 แบบ คือ แบบ Dynamic forecast และแบบ Static forecast โดยโปรแกรม EViews ก็จะทำการพยากรณ์และเสนอค่าสถิติต่างๆ ดังรูป

ผลการพยากรณ์แบบ Dynamic Forecast



ผลการพยากรณ์แบบ Static Forecast



เอกสารอ้างอิง

ถวิล นิลใบ. **เศรษฐมิติ 2**. กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยรามคำแหง, 2544.

ทรงศิริ แต้สมบัติ. **การพยากรณ์เชิงปริมาณ.** กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์, 2549.

ทรงศิริ แต้สมบัติ. **เทคนิคการพยากรณ์เชิงปริมาณ.** กรุงเทพฯ: หจก. สำนักพิมพ์ฟิสิกส์เซ็นเตอร์, 2539.

ไพฑูรย์ ไกรพรศักดิ์. **เศรษฐมิติเบื้องต้น.** กรุงเทพฯ: คณะเศรษฐศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2546.

มุกดา แม้นมินทร์. อนุกรมเวลาและการพยากรณ์. กรุงเทพฯ: สำนักพิมพ์ประกายพรึก, 2549

วิชิต หล่อจีระชุณห์กุล และคณะ. **เทคนิคการพยากรณ์.** กรุงเทพฯ : โครงการส่งเสริมเอกสารวิชาการ สถาบันบัณฑิตพัฒนบริหารศาสตร์, 2539

ศิริลักษณ์ สุวรรณวงศ์. **การวิเคราะห์อนุกรมเวลา.** กรุงเทพฯ : สุวีริยาสาส์น, 2535.

อัครพงศ์ อั้นทอง. "**การประเมินจำนวนนักท่องเที่ยวต่างชาติที่ลดลงของไทยจากผลของโรค SARS: โดยใช้แบบจำลอง SARIMA.**" เอกสารประกอบการนำเสนอ ณ งานประชุมสัมมนานักเศรษฐศาสตร์รุ่นเยาว์ ครั้งที่ 1 คณะเศรษฐศาสตร์ มหาวิทยาลัยเชียงใหม่, 2547.

อัครพงศ์ อั้นทอง. **คู่มือการใช้โปรแกรม EViews เพื่อการวิเคราะห์ Unit Root, Cointegration Correction Model (ตามวิธีการของ Engel and Granger)**. สถาบันวิจัยสังคม มหาวิทยาลัยเชียงใหม่, 2546.

EViews 5.1 Command and Programming Reference. Quantitative Micro Software, LLC, 2005.

EViews 5.1 User's Guide. Quantitative Micro Software, LLC, 2005.

Enders, W. **Applied Econometric Time Series.** Second edition, New York: John Wiley & Sons, 2004.

Greene, W.H. **Econometric Analysis.** 5th ed. Printice Hall, 2003.

Gujarati, D. **Basic Econometrics**. 3rd ed. McGraw–Hill, 1995.

Johnston, J. and J, Dinardo.. **Econometric Methods**. 4th ed. McGraw–Hill, 1997.

R.R. Johnson. **A Guide to Using EViews with Using Econometrics: A Practical Guide.** University of San Diego, 2000.

Ramanathan, Ramu. **Introductory Econometrics with Application**. 3rd ed. Harcourt Brace & Company, 1995.

Studenmund, A.H. **Using Econometrics: A Practical Guide**. 5th e.d. Pearson Education, Inc., 2006.