ภาคผนวก

ภาคผนวก ก

การทดสอบความแปรปรวนของตัวแปรคลาดเคลื่อนที่ไม่คงที่ (heteroscedasticity)

สาเหตุของการเกิด heteroscedasticity

มีสาเหตุหลักที่สำคัญ 2 ประการคือ

1. การกำหนดตัวแบบที่ไม่ถูกต้อง โดยละเลยตัวแปรอิสระบางตัว ทำให้ค่าตัวรบกวนอาจจะมีความสัมพันธ์กับตัวแปรอิสระที่ถูกละเลยตัวนั้น

2. ข้อมูลประเภทภาคตัดขวาง (cross section data) ข้อมูลประเภทนี้ค่าความแปรปรวนของตัวรบกวนที่เกิดขึ้นในแต่ละค่าสังเกต (observation) อาจไม่คงที่ มีโอกาสมากเมื่อเทียบกับกรณีข้อมูลเป็นแบบอนุกรมเวลา (time series) ทั้งนี้เพราะค่าสังเกตของข้อมูลประเภทภาคตัดขวางจะมีความแตกต่างกันตามขนาดหรือลำดับ แต่ข้อมูลอนุกรมเวลาจะไม่มีความแตกต่างในประเด็นนี้

ในข้อเท็จจริงแล้วยังไม่มีข้อตกลงร่วมกันที่เป็นมาตรฐานสำหรับวิธีการทดสอบดังกล่าวนี้ว่าจะต้องใช้วิธีการใดวิธีการหนึ่งเท่านั้นเป็นมาตรฐาน (ไพฑูรย์ ไกรพรศักดิ์, 2546) ในที่นี้จะเสนอวิธีการ ดังต่อไปนี้

The White Test

การทดสอบเพื่อวิเคราะห์ว่าสมการประมาณการนั้นมีปัญหา heteroscedasticity หรือไม่ด้วยวิธีของ White จะอาศัยการประมาณการสมการถดถอยที่มีค่ากำลังสองเป็นตัวแปรอิสระของสมการที่จะทำการทดสอบ ที่จัดให้เป็นฟังก์ชั่นของตัวแปรอธิบายต่างๆที่ใช้อยู่ทั้งหมด ประกอบกับเพิ่มชุดของตัวแปรอธิบายเหล่านั้นที่อยู่ในรูปกำลังสองและค่าปฏิสัมพันธ์ของตัวแปรอธิบายเหล่านั้น กล่าวคือ

1. คำนวณหาค่าตัวแปรคลาดเคลื่อนจากสมการประมาณการที่ได้จากสมการตัวแบบ
2. ใช้ค่าตัวแปรคลาดเคลื่อนประมาณการนั้น มาสร้างสมการทดสอบโดยให้ค่ากำลังสองของตัวแปรคลาดเคลื่อนประมาณการที่ได้นั้นเป็นตัวแปรอิสระและใช้ตัวแปรอธิบายทั้งหมดรวมถึงค่ากำลังสองและค่าของพจน์ที่เป็นค่าปฏิสัมพันธ์ของตัวแปรอธิบายเหล่านั้นมาเป็นตัวแปรอธิบายในสมการทดสอบดังกล่าว คือ

$$(e_i)^2 = \beta_0 + \beta_1X_{1i} + \beta_2X_{2i} + \beta_3X_{3i} + \beta_4X_{1i}^{\ 2} + \beta_5X_{2i}^{\ 2} + \beta_6X_{3i}^{\ 2} + \beta_7X_{1i}X_{2i} + \beta_8X_{1i}X_{3i} + \beta_9X_{2i}X_{3i} + \mu_I$$

ภายใต้ข้อสมมติฐาน Null Hypothesis การมีความแปรปรวนของตัวแปรคลาดเคลื่อนที่มีค่าคงที่ (homoscedasticity) แล้วตัวสถิติที่ใช้ในที่นี้สามารถคำนวณได้มีค่าเท่ากับ nR^2 เมื่อค่า R^2 ในที่นี้เป็นค่าที่ยังไม่ได้มีการปรับอันเนื่องมาจากค่าองศาความเป็นอิสระ (unadjusted R^2) และคำนวณได้จากสมการ โดยมีองศาของความเป็นอิสระเท่ากับจำนวนสัมประสิทธิ์ของตัวแปรอธิบายทั้งหมด (สัมประสิทธิ์ที่เป็นความชัน)

การแก้ปัญหา heteroscedasticity โดยวิธีการกำลังสองถ่วงน้ำหนัก (weighted least square : WLS)

วิธีการ WLS นี้ สามารถดำเนินการตามขั้นตอนดังนี้ หารสมการตัวแบบตลอดด้วยค่าของตัวแปร proportional factor (Z) ซึ่งเป็นตัวแปรอิสระที่ปรากฏว่ามีความสัมพันธ์ที่จะมีผลต่อการเกิดความแปรปรวนของตัวแปรคลาดเคลื่อนไม่คงที่นั้น จากนั้นประมาณการสมการที่หารตลอดด้วย Z นั้น ด้วยวิธีกำลังสองน้อยที่สุด (OLS) เมื่อได้ค่าสัมประสิทธิ์ของตัวแปรที่ถูกหารด้วยตัวแปร Z เหล่านั้นแล้ว ก็สามารถเทียบนำเอาสัมประสิทธิ์เหล่านี้ไปใช้เขียนสมการดังเดิมที่เราต้องการได้ เนื่องจากจากข้อเท็จจริงแล้วการหารด้วยตัวแปรใดๆ ก็ตามจะไม่ก่อให้เกิดการเปลี่ยนแปลงสัมประสิทธิ์ของสมการแต่อย่างใด

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i \tag{1.1}$$

ในที่นี้ สมมติให้ค่าความแปรปรวนของตัวแปรคลาดเคลื่อนมีค่าไม่คงที่ ดังแสดงในสมการที่ (1.2)

$$Var(\varepsilon_i) = \sigma^2 = \sigma^2 Z^2_i \tag{1.2}$$

จากนั้น ทำการหารตลอดด้วยค่าตัวแปร Z

$$Y_i / Z_i = \beta_0/ Z_i + \beta_1 X_{1i} / Z_i + \beta_2 X_{2i}/ Z_i + u_i \tag{1.3}$$

ด้วยวิธีการดังกล่าวข้างต้นนี้ เราสามารถจะทำการประมาณการสมการ (1.3) ได้ด้วยวิธีการกำลังสองน้อยที่สุด (OLS) แต่อย่างไรก็ตาม จะพบว่า สมการที่ (1.3) ที่ได้มาดังกล่าว ไม่มีตัวค่าคงที่รวมอยู่ในสมการด้วย ซึ่งโดยหลักการประมาณการด้วยวิธีการกำลังสองน้อยที่สุดนี้ หากไม่มีตัวคงที่รวมอยู่ในสมการด้วยแล้ว จะก่อให้เกิดปัญหาหรือผลเสียคือ อาจก่อให้เกิดการเอนเอียง (bias) หรือเกิดการคลาดเคลื่อนจากการวัด (measurement error) เกิดขึ้น

ในทางปฏิบัติ เรามักจะใส่เพิ่มตัวคงที่ในสมการที่จะประมาณการด้วยวิธี WLS เสมอ (ดังสมการที่ 1.4) จะเป็นการปลอดภัยกว่า และหากว่าตัวคงที่นั้นๆ ไม่มีอยู่จริงแล้ว ก็จะไม่ก่อให้เกิดปัญหาเอนเอียงหรือการวัดคลาดเคลื่อน แต่อย่างไรอีกด้วย

$$Y_i / Z_i = \alpha_0 + \beta_0/ Z_i + \beta_1 X_{1i} / Z_i + \beta_2 X_{2i}/ Z_i + u_i \tag{1.4}$$

อย่างไรก็ตาม ยังมีปัญหาบางประการเกี่ยวกับการประมาณการด้วยวิธี WLS ดังนี้

1. ปัญหาว่าจะใช้ตัวแปรอะไรเป็น Z ในที่นี้ซึ่งเป็นปัญหาที่ไม่ง่ายนักในการประมาณการแต่ละครั้ง

2. ปัญหาด้านรูปแบบของความสัมพันธ์ระหว่างความแปรปรวนของตัวแปรคลาดเคลื่อนกับตัวแปร Z ดังกล่าวว่าที่แท้จริงเป็นอย่างไร เป็นแบบยกกำลังสองหรือไม่

3. หากว่าเป็นกรณีความแปรปรวนของตัวแปรคลาดเคลื่อนที่ไม่คงที่นั้นเป็นแบบไม่แท้จริง (impure heteroscedastcity) แล้ว การประมาณการแบบ WLS แม้ว่าจะช่วยในการลดปัญหาความเอนเอียงที่เกิดจากการละเลยตัวแปรอธิบายลงไปได้ แต่วิธีการ WLS ดังกล่าวยังด้อยกว่าผลจากกรณีที่สามารถประมาณการจากสมการที่ได้มีการระบุรูปแบบความสัมพันธ์ที่ถูกต้อง

ภาคผนวก ข

การทดสอบอัตสหสัมพันธ์ข้ามเวลาของตัวแปรความคลาดเคลื่อน (autocorrelation)

เมื่อข้อมูลเก็บรวบรวมตามเวลา ควรทำการศึกษาดูผลกระทบของเวลาว่าจะถูกรวมเข้าไปในความคลาดเคลื่อนหรือไม่ ซึ่งผลกระทบอาจจะมีมากขึ้น ลดลง หรือหมดไปก็ได้ การศึกษาความสัมพันธ์ระหว่างเวลาภายในข้อมูลชุดนั้นจะเรียกความสัมพันธ์นั้นว่า อัตสหสัมพันธ์ ความสัมพันธ์นี้เกิดขึ้นตามระดับต่างๆ ของตัวแปรตามในช่วงเวลาหนึ่งกับเวลาถัดไป ดังนั้นส่วนที่เหลือในช่วงเวลาหนึ่งจะสัมพันธ์กับส่วนที่เหลือในช่วงเวลาหนึ่งกับเวลาถัดไป ดังนั้นอัตสหสัมพันธ์จะเกิดขึ้นระหว่างคู่ของส่วนที่เหลือและคำนวณสัมประสิทธิ์ของส่วนที่เหลือคู่นั้นๆ อัตสหสัมพันธ์จะมีค่าเป็นได้ทั้ง “บวก” และ “ลบ”

จากสมการ $X_t = \rho X_{t-1} + e_t$ โดยที่ $-1 < \rho < 1$

กำหนดโครงสร้างของ X_t ให้เป็นฟังก์ชันของ e_t โดยอาศัยการแทนซ้ำเนื่องจากอิทธิพลของ X มีต่อกันและกัน โดยการย้อนกาลเวลาไปในอดีต 1 วาระ จะพบว่า

$$X_{t-1} = \rho X_{t-2} + e_{t-1}$$

$$r_{t,t-1} = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=2}^{n} e_{t-1}^2}$$

เมื่อ $r_{t,t-1}$ เป็นสหสัมพันธ์ของส่วนที่เหลือตามช่วงเวลา t และ $t-1$

การทดสอบสมมติฐาน

เมื่อนำข้อมูลมาสร้างสมการถดถอยย่อมจะเกิดความสัมพันธ์ระหว่างส่วนที่เหลือรวมอยู่ด้วย ดังนั้นจึงจำเป็นต้องทดสอบดูความสัมพันธ์นั้นว่าจะกระทบข้อมูลตามช่วงเวลานั้นหรือไม่ การทดสอบวิธีดังกล่าวนี้จะใช้วิธีของเดอร์บิน–วัตสัน (Durbin-Watson) สถิติที่ใช้คือ “d” การทดสอบสมมติฐานจะทดสอบอัตสหสัมพันธ์ที่เป็นทั้ง “บวก” และ “ลบ”

ให้ $\rho_{t,t-1}$ เป็นสัมประสิทธิ์อัตสหสัมพันธ์ของประชากรที่ช่วงเวลา t และ t-1 ดังนั้นการทดสอบสมมติฐานเมื่อเป็นอัตสหสัมพันธ์เชิงบวก การตั้งสมมติฐานดังนี้

$H_0 : \rho_{t,t-1} = 0$ $\qquad$ $H_1 : \rho_{t,t-1} > 0$

สถิติที่ใช้ทดสอบคือ

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$$

เมื่อเป็นอัตสหสัมพันธ์เชิงลบ การตั้งสมมติฐานจะเป็น

$$H_0 : \rho_{t,t-1} = 0 \qquad\qquad H_1 : \rho_{t,t-1} < 0$$

สถิติที่ใช้ทดสอบคือ

$$\bar{d} = 4 - d$$

การยอมรับหรือปฏิเสธสมมติฐานจะใช้ตารางเดอร์บิน–วัตสัน (Durbin-Watson Table)

ภายใต้คำว่า "ρ" ดังกล่าวจะมีค่า d_L และ d_u ในตารางซึ่งค่านี้จะหมายถึงขอบเขตต่ำสุด (lower limit) และขอบเขตสูงสุด (upper limit) ตามลำดับ ดังนั้นถ้าตั้งสมมติฐานเป็นอัตสหสัมพันธ์เชิงบวก การสรุปผลจะพิจารณาดังนี้

1. ปฏิเสธ (reject) H_0 : ถ้า $d < d_L$
2. ยอมรับ (accept) H_0 : ถ้า $d > d_u$
3. ถ้า $d_L \le d \le d_u$ จะสรุปผลไม่ได้

และถ้าตั้งสมมติฐานเป็นอัตสหสัมพันธ์เชิงลบ การสรุปผลจะพิจารณาดังนี้

4. ปฏิเสธ (reject) H_0 : ถ้า $\bar{d} < d_L$
5. ยอมรับ (accept) H_0 : ถ้า $\bar{d} > d_u$
6. ถ้า $d_L \le \bar{d} \le d_u$ จะสรุปผลไม่ได้

การแก้ปัญหา autocorrelation โดยวิธีการกำลังสองน้อยที่สุดแบบทั่วไป (generalized least squares : GLS)

วิธีการ GLS เป็นวิธีการที่นิยมกันวิธีหนึ่งที่สามารถแก้ปัญหาผลเสียที่เกิดขึ้นจากการประมาณการสมการที่เกิดปัญหาสหสัมพันธ์ข้ามเวลาของตัวแปรความคลาดเคลื่อนได้ ซึ่งวิธีการนี้

จะพยายามไปแก้ไข Matrix Variance-Covariance Matrix ของตัวความคลาดเคลื่อนในสมการนั้นให้กลายเป็น Matrix ที่มีคุณสมบัติถูกต้องตามข้อสมมติฐานแบบดั่งเดิม (ไพฑูรย์ ไกรพรศักดิ์ , 2546 : 8-13) ดังนี้

$$Y_t = \beta_0 + \beta_1 X_{1t} + \varepsilon_t \quad (2.1)$$

ในที่นี้สมมติให้เกิดปัญหาสหสัมพันธ์ข้ามเวลา

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

ดังนั้นสมการที่ 2.1 สามารถจะเขียนได้ ดังสมการที่ 2.2 ดังนี้

$$Y_t = \beta_0 + \beta_1 X_{1t} + \rho\varepsilon_{t-1} + \mu_t \quad (2.2)$$

ในสมการที่ 2.2 ข้างต้นนั้นจะพบว่า ε เป็นตัวแปรความคลาดเคลื่อนที่เกิดปัญหาสหสัมพันธ์ข้ามเวลา และ μ เป็นตัวแปรความคลาดเคลื่อนที่ไม่มีปัญหาสหสัมพันธ์ข้ามเวลาแล้วคูณสมการที่ 2.1 ข้างต้นด้วยค่า ρ (ค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรความคลาดเคลื่อน) ตลอดทั้งสองข้าง จากนั้นทำการเลื่อนเวลา (lag) ถอยหลังไปหนึ่งช่วงเวลา จะได้สมการที่ 2.3

$$\rho Y_{t-1} = \rho\beta_0 + \rho\beta_1 X_{1t-1} + \rho\varepsilon_{t-1} \quad (2.3)$$

ดำเนินการโดยลบสมการที่ 2.2 ด้วยสมการที่ 2.3 ผลลัพธ์ที่ได้ แสดงดังสมการที่ 2.4

$$Y_t - \rho Y_{t-1} = \beta_0(1-\rho) + \beta_1(X_{1t} - \rho X_{1t-1}) + \mu_t \quad (2.4)$$

มีข้อสังเกตว่าตัวแปรความคลาดเคลื่อนในสมการที่ 2.4 นั้น ไม่มีปัญหาสหสัมพันธ์ข้ามเวลาอีกต่อไป สมการที่ 2.4 ข้างต้นนั้น เขียนใหม่ให้อยู่ในรูปอย่างง่ายที่สามารถจะนำไปใช้ในการประมาณการด้วยวิธีกำลังสองน้อยที่สุด (OLS) ได้ดี เนื่องจากในสมการนี้นั้น ไม่มีปัญหาสหสัมพันธ์ข้ามเวลาของตัวแปรความคลาดเคลื่อนแล้ว ดังนี้

$$Y_t^* = \beta_0^* + \beta_1 X^*_{1t} + \mu_t \quad (2.5)$$

สมการที่ 2.5 นี้ใช้เป็นสมการประมาณการด้วยวิธีการกำลังสองน้อยที่สุด (OLS) โดยที่สมการที่ 2.5 นี้ก็คือการประมาณการตามวิธีการ GLS นั่นเอง การใช้วิธีการประมาณการแบบกำลังสองน้อยที่สุด (OLS) สำหรับสมการที่ 2.5 นี้ไม่เกิดปัญหาสหสัมพันธ์ข้ามเวลาของตัวแปรความคลาดเคลื่อนแต่อย่างไร และจะให้ค่าประมาณการที่มีคุณสมบัติที่พึงประสงค์ทุกประการ

ประวัติผู้เขียน

ชื่อ	นางสาวหัตทยา บุญธรรม	
วัน เดือน ปี เกิด	16 ตุลาคม 2524	
ประวัติการศึกษา	ปีการศึกษา 2538	สำเร็จการศึกษามัธยมศึกษาตอนต้น โรงเรียนวัชรวิทยา จังหวัดกำแพงเพชร
	ปีการศึกษา 2541	สำเร็จการศึกษามัธยมศึกษาตอนปลาย โรงเรียนกำแพงเพชรพิทยาคม จังหวัดกำแพงเพชร
	ปีการศึกษา 2545	สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต (เศรษฐศาสตร์เกษตร) มหาวิทยาลัยเชียงใหม่ (เกียรตินิยมอันดับสอง)
ประสบการณ์	ปี 2547 – ปัจจุบัน	นักวิชาการส่งเสริมการเกษตร ระดับ 3 สำนักงานเกษตรอำเภอบางกระทุ่ม จังหวัดพิษณุโลก