บทที่ 3

ระเบียบวิธีการศึกษา

ในระเบียบวิธีที่จะใช้ทำการศึกษา เพื่อให้บรรลุวัตถุประสงค์ในข้อที่ 1 จึงได้ทำการวิเคราะห์ โดยใช้การแจกแจงร้อยละ พร้อมทั้งหาค่าสัมประสิทธิ์จินี่ (Gini Coefficient) เพื่อวัดการกระจายสินเชื่อให้แก่หน่วยธุรกิจในแต่ละสาขาเศรษฐกิจ และนำค่าสัมประสิทธิ์จินี่ที่ได้มาเปรียบเทียบกัน โดยทำการเปรียบเทียบค่าสัมประสิทธิ์จินี่ของการกระจายสินเชื่อโดยธนาคารพาณิชย์ที่อยู่ในเขตอำเภอเมืองและเขตอำเภอรอบนอก ตลอดจนเปรียบเทียบค่าสัมประสิทธิ์จินี่ของการกระจายสิน-เชื่อระหว่างสาขาเศรษฐกิจต่าง ๆ ที่อยู่ในเขตอำเภอเมืองและเขตอำเภอรอบนอก

สำหรับวัตถุประสงค์ข้อ 2 นั้น ได้ทำการศึกษาหลักเกณฑ์แผนงาน และนโยบายการให้สินเชื่อ โดยธนาคารพาณิชย์ทุก ๆ สาขาทั้งอำเภอเมือง และอำเภอรอบนอก โดยได้ทำการสัมภาษณ์บุคคลผู้มีอำนาจในการตัดสินใจอย่างละเอียดเช่น ผู้จัดการ ผู้ช่วยผู้จัดการ สมุหบัญชี หัวหน้าแผนกสินเชื่อที่เกี่ยวข้อง เป็นต้นและนำผลการดำเนินงานของธนาคารพาณิชย์ ที่ได้จากการวิเคราะห์มาเปรียบเทียบกับเป้า-หมายที่รัฐได้กำหนดเอาไว้เพื่อบรรลุวัตถุประสงค์ข้อ 3

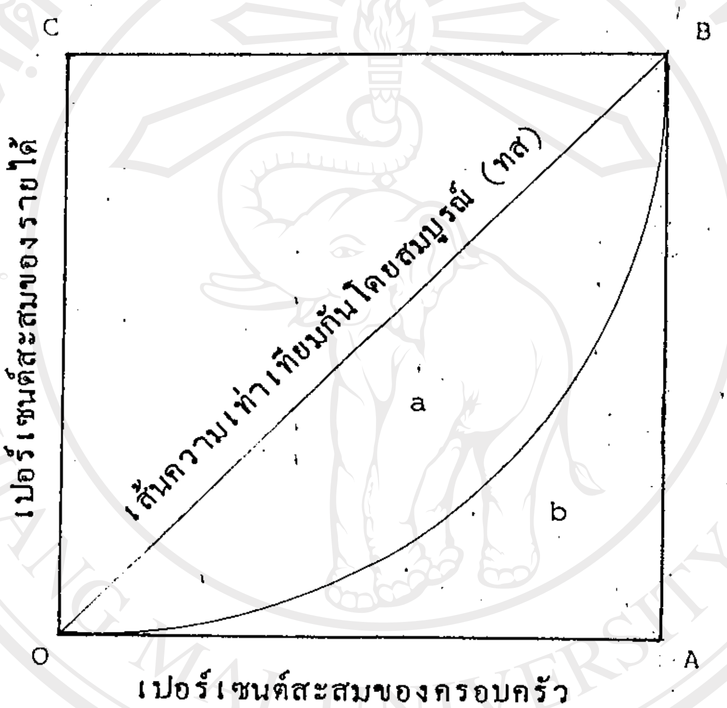
ในการศึกษาได้นำเอาวิธีการศึกษาอย่างละเอียด(intensive study) เข้ามาช่วยในการตรวจสอบข้อมูล และเพื่อทำความเข้าใจในกลไกอย่างแท้จริงเกี่ยวกับการดำเนินงานของธนาคารด้วย ทั้งนี้เพื่อให้ผลการวิเคราะห์ถูกต้องแม่นยำมากยิ่งขึ้น

สำหรับการคำนวณค่าสัมประสิทธิ์จินี่นั้น ได้ใช้วิธีการอันเดียวกันกับการคำนวณรายได้ของครอบครัวไทย ปี 2515 โดยเมธี ครองแก้ว[1] โดยการใช้เส้นลอเรนซ์ (Lorenz Curve) เป็นตัวแสดงแบบแผนการกระจายรายได้

นั่นคือเมธีได้ใช้เส้นลอเรนซ์ (Lorenz Curve) เป็นตัวแสดงแบบการกระจายรายได้ด้วยการสร้างรูปสี่เหลี่ยมจตุรัสขึ้นมารูปหนึ่ง และกำหนดให้แกนนอนของรูป แสดงถึงสัดส่วนของความถี่สะสมของจำนวนครอบครัว โดยแบ่งตามชั้นของรายได้ และกำหนดให้แกนตั้งเป็นสัดส่วนความถี่สะสมของรายได้ทั้งหมด และได้เส้นลอเรนซ์จากการลากเส้นเชื่อมจุด Coordinate ของทั้งสองแกนดังกล่าว เส้นลอเรนซ์จะแสดงลักษณะการกระจายรายได้ โดยเปรียบเทียบกับเส้นทะแยงมุมของรูปสี่เหลี่ยมจตุรัส โดยที่ถ้าเส้นนี้ทับเส้นทะแยงมุมจะแสดงถึงความเท่าเทียมกันอย่างสมบูรณ์ (ทส) ในการกระจายรายได้ แต่ถ้าเส้นนี้ยิ่งออกห่างจากเส้นทะแยงมุมจะยิ่งแสดงถึงความไม่เท่าเทียมกันในการกระจายได้ และจากการศึกษาของเมธีได้ใช้สัมประสิทธิ์จินี่(Gini Coefficient)เป็นตัวชี้ให้เห็นถึงลักษณะการกระจายรายได้ โดยที่ถ้าหากว่าค่าของสัมประสิทธิ์จินี่ที่คำนวณได้เป็น 0 แสดงว่าการกระจายรายได้มีความเท่าเทียมกันอย่างที่สุด (absolute equality) และถ้าค่าเป็น1ก็แสดงว่าการกระจายรายได้ไม่เท่าเทียมกันอย่างที่สุด (absolute inequality)

[1] อ้างแล้ว, หน้า 138-152.

รูปแสดงเส้นลอเรนซ์



จากรูปสี่เหลี่ยมจตุรัสนี้ค่าสัมประสิทธิ์จินี่ คือ

$$G = \frac{\text{พื้นที่ } a}{\text{พื้นที่ } a+b}$$

โดยทั่วไปแล้วการคำนวณค่าสัมประสิทธิ์จินี่ ของการกระจายรายได้มักจะกระทำโดยการรวมเนื้อที่สามเหลี่ยม ซึ่งสร้างขึ้นมาจากจุด Coordinate ภายใต้เส้นลอเรนซ์จากนั้นก็คูณเนื้อที่นี้ด้วย 2 แล้วนำไปลบออกจากเนื้อที่ของสี่เหลี่ยมจตุรัส

ซึ่งมีค่าเท่ากับ 1 ส่วนที่เหลือจะเท่ากับสองเท่าของเนื้อที่ระหว่างเส้นเสมอภาคหรือเส้นความเท่าเทียมกันโดยสมบูรณ์ (ทส) กับเส้นลอเรนซ์ ซึ่งก็คือค่าสัมประสิทธิ์จินี่นั่นเองหรือในรูปของสูตร

<u>สูตรที่</u> 1

$$G = 1 - \sum_{i=1}^{n} (Z_i \; P_i)$$

โดย

G = ค่าสัมประสิทธิ์จินี่

$Z_i = Y_i + Y_{i-1}$ = ความถี่สะสมของสินเชื่อทั้งหมดของหน่วยธุรกิจซึ่งได้รับสินเชื่อระดับที่ i รวมกับความถี่สะสมของสินเชื่อทั้งหมดของหน่วยธุรกิจ ซึ่งได้รับสินเชื่อในระดับก่อนหน้า

P_i = ความถี่สะสม หรือสัดส่วนของหน่วยธุรกิจที่ได้รับสินเชื่อระดับที่ i

i = ระดับชั้นของสินเชื่อแต่ละระดับ

สำหรับการศึกษาครั้งนี้จะไม่นำเอาวิธีการคำนวณค่าสัมประสิทธิ์จินี่ โดยสูตรที่ 1 (วิธีลัด) มา เพราะเห็นว่าเป็นวิธีการที่ไม่ละเอียดพอ และได้นำสูตรที่ 2 มาใช้คำนวณ

สูตรที่ 2

$$G = 1-2\left\{ \sum_{i=1}^{n} (f_i - f_{i-1})(Y_{i-1}) + \frac{1}{2}\sum_{i=1}^{n} (f_i - f_{i-1})(Y_i - Y_{i-1}) \right\}$$

โดย

G = ค่าสัมประสิทธิ์จินี่

f_i = ความถี่สะสมของจำนวนครอบครัว ซึ่งมีรายได้ที่ระดับ i

Y_i = ความถี่สะสมของรายได้ทั้งหมดของครอบครัว ซึ่งมีรายได้ที่ระดับ i

i = 1, 2, ..., n = จำนวนชั้นของรายได้ที่ทราบหรือเป็นระดับชั้นของรายได้แต่ละช่วง

สำหรับการศึกษา เรื่องการกระจายสินเชื่อโดยธนาคารพาณิชย์จะกำหนด ให้แกนตั้งของรูปสี่เหลี่ยมจตุรัส คือ เปอร์เซนต์สะสมของสินเชื่อ และแกนนอน คือ เปอร์เซนต์สะสมของหน่วยธุรกิจ และให้

G = ค่าสัมประสิทธิ์จินี่

f_i = ความถี่สะสมของจำนวนหน่วยธุรกิจที่ได้รับสินเชื่อระดับที่ i

Y_i = ความถี่สะสมของสินเชื่อทั้งหมดที่หน่วยธุรกิจได้รับ ณ ระดับที่ i

i = ระดับชั้นของสินเชื่อ

n = จำนวนระดับชั้นของสินเชื่อ

ค่าสัมประสิทธิ์จินี่ที่คำนวณได้ด้วยวิธีการดังกล่าวนี้เป็นที่น่าสงสัยว่าอาจจะมากกว่าหรือน้อยกว่าค่าที่ควรจะเป็นจริงก็ได้ ทั้งนี้เพราะการคำนวณหาเนื้อที่ภายใต้เส้นลอเรนซ์ คือเนื้อที่ b นั้น เป็นการคำนวณหาค่าของพื้นที่สามเหลี่ยม ซึ่งเป็นพื้นที่ครึ่งหนึ่งของรูปสี่เหลี่ยมจตุรัสที่ถูกแบ่งด้วยเส้นทะแยงมุม ตามรูปก็คือ สามเหลี่ยม OAB และพื้นที่สี่เหลี่ยมจตุรัส OABC ซึ่งจะเห็นว่าได้รวมเอาเนื้อที่เหนือเส้นลอเรนซ์เข้าไปด้วย ดังนั้นถ้าหากว่าการคำนวณหาเนื้อที่ภายใต้เส้นลอเรนซ์ คือเนื้อที่ b ดังกล่าวนี้สามารถจะหาโดยให้ได้ค่าที่ใกล้เคียงกับความเป็นจริงให้มากที่สุดเท่าที่จะทำได้โดยไม่ใช้วิธีประมาณค่าตามวิธีลัด หรือตามสูตร

$$G = 1-2\left[\sum_{i=1}^{n}(f_i-f_{i-1})(Y_{i-1}) + \frac{1}{2}\sum_{i=1}^{n}(f_i-f_{i-1})(Y_i-Y_{i-1})\right]$$

แล้ว คาดว่าค่าสัมประสิทธิ์จินี่น่าจะถูกต้อง และใกล้เคียงความจริงมากกว่า

โดยวิธีการศึกษาเช่นเดียวกันกับที่ N.C. Kakwani และ N. Podder ได้ใช้ทำการศึกษา กรณีการกระจายรายได้ในกลุ่มครัวเรือน หรือผู้บริโภคชาวออสเตรเลีย ในช่วงปี 1967-1976[2] ซึ่ง เมธี ครองแก้ว ได้นำเอาวิธีการเดียวกันนี้มาทำการศึกษา เพื่อทดสอบผลของการวิเคราะห์ด้วยวิธี แรก หรือวิธีลัดที่กล่าวมาแล้ว เพื่อหาค่าสัมประสิทธิ์จินี่ที่ใกล้เคียงค่าที่เป็นจริง และถูกต้องให้มากยิ่งขึ้น โดยใช้สูตรดังต่อไปนี้

สูตรที่ 3

$$G = 2a(\sqrt{2})^{1+\alpha+\beta} B(1+\alpha, 1+\beta)$$

โดยกำหนดให้:

G = ค่าสัมประสิทธิ์จินี่

$B(1+\alpha, 1+\beta)$ = ค่าปฏิพันธ์เบต้า (Beta function) ซึ่งมีค่าเท่ากับ $\dfrac{\Gamma(1+\alpha)\,\Gamma(1+\beta)}{\Gamma(2+\alpha+\beta)}$

α และ β = ค่าสัมประสิทธิ์ (Coefficient) ของเส้นลอเรนซ์

ซึ่งในกรณีของการศึกษาเรื่องการกระจายสินเชื่อ โดยธนาคารพาณิชย์ครั้งนี้ถือว่าข้อมูลเกี่ยวกับจำนวนสินเชื่อและจำนวนลูกค้าที่ได้รับสินเชื่อนั้นเป็น ประชากรไม่จำกัด(unlimited population) ซึ่งไม่สามารถจะรวบรวมมาได้ทั้งหมด

---

[2] N.C. Kakani and N. Podder, "Efficience Estimation of The Lorenz Curve and Associated Inequality Measures from Grouped observations", Econometrica, Vol. 44, No.1 (January 1976), pp.137-148.

ดังนั้นจึงไม่สามารถจะคำนวณค่า α และ β โดยตรงได้ ด้วยเหตุนี้จึงทำการสุ่มตัวอย่างข้อมูลดังกล่าวบางส่วนมา แล้วคำนวณหาค่า $\hat{\alpha}$ และ $\hat{\beta}$ จากข้อมูลตัวอย่างนั้นและถือว่า $\hat{\alpha}$ และ $\hat{\beta}$ คือ ตัวประมาณค่า (estimators) ของ α และ β

ดังนั้นจึงได้แทนค่า $\hat{\alpha}$ และ $\hat{\beta}$ ลงในสมการถดถอยของ Kakwani และ Podder และใช้สูตรว่า

$$G = 2a(\sqrt{2})^{1+\hat{\alpha}+\hat{\beta}} B(1+\hat{\alpha}, 1+\hat{\beta})$$

และค่าปฏิพันธ์เบต้า

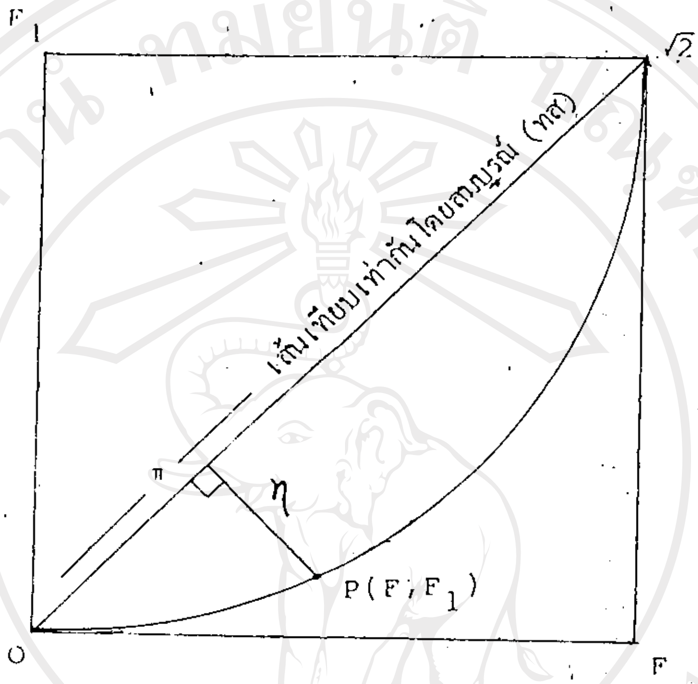
$$B(1+\hat{\alpha}, 1+\hat{\beta}) = \frac{\gamma(1+\hat{\alpha})\,\gamma(1+\hat{\beta})}{\gamma(2+\hat{\alpha}+\hat{\beta})}$$

สำหรับการศึกษาโดย Kakwani และ Podder ได้ทำการสมมุติให้ระดับรายได้ที่ x ของหน่วยครัวเรือนหนึ่ง ซึ่งเป็นตัวแปรสุ่ม (random Variable) ซึ่งมีปฏิพันธ์การกระจายความน่าจะเป็นไปได้ (Probability distribution function) เท่ากับ F(x) และสมมุติให้ค่าเฉลี่ยของการกระจายรายได้คือ η และค่าระดับรายได้เป็นบวกเท่านั้น นั่นคือหมายความว่าจะไม่มีครัวเรือนไหนที่จะมีรายได้ติดลบดังนั้นจึงกำหนดโมเมนท์ที่หนึ่งของปฏิพันธ์การกระจายรายได้ x เท่ากับ

$$F_1(x) = \frac{1}{\mu}\int_0^x Xg(x)dx \qquad (1)$$

โดยให้ g(x) คือปฏิพันธ์ความหนาแน่น (density function)

รูปแสดงเส้นลอเรนซ์



เส้นลอเรนซ์ที่แสดงในรูปนี้ คือเส้นแสดงถึงความสัมพันธ์ระหว่าง $F(x)$ กับ $F_1(x)$ สมการของเส้น $F_1 = F$ หรือเส้นทแยงมุมล่างซ้ายไปยังมุมบนขวาของรูปสี่เหลี่ยมจตุรัส ซึ่งมีพื้นที่เท่ากับหนึ่ง ก็คือเส้นความเท่าเทียมกันโดยสมบูรณ์ (ทส)

กำหนดให้ P เป็นจุด ๆ หนึ่งบนเส้นลอเรนซ์ โดยมี Co-ordinate (F,F_1) และ

$$\pi = \frac{1}{\sqrt{2}}(F+F_1) \text{ และ } \eta = \frac{1}{\sqrt{2}}(F-F_1)$$

โดย η คือความยาวของเส้นที่ลากจาก P ไปยังเส้นเท่าเทียมกันโดยสมบูรณ์ (ทส) และ π คือระยะทางจุดเริ่มต้นไปตามเส้น ทส. สมการของเส้นลอเรนซ์ในรูปของ π และ η สามารถเขียนได้ดังนี้คือ

$$\eta = f(\pi)$$

โดยค่าของ π จะเปลี่ยนจาก 0 ถึง $\sqrt{2}$ และค่าของ η ที่ได้จะบอกถึงส่วนโค้งของเส้นลอเรนซ์ตลอดความยาวของเส้นทแยงมุม π

$F = f$ = ความถี่สะสมของจำนวนหน่วยธุรกิจที่ได้รับสินเชื่อตามชั้นของสินเชื่อต่าง ๆ กัน

$F_1 = y$ = ความถี่สะสมของสินเชื่อทั้งหมดตามชั้นของสินเชื่อต่าง ๆ กัน

จากนั้น Kakwani และ Podder ได้สร้างสมการในรูปเฉพาะ (explicit form) เพื่อใช้ประมาณความสัมพันธ์ของ η กับ π คือ

$$\eta = a\pi^{\alpha}(\sqrt{2} - \pi)^{\beta}, \; a>0, \; \alpha>0 \text{ และ } \beta>0 \qquad (2)$$

ข้อจำกัดที่ว่าให้ $a > 0$ หมายความว่า $\eta > 0$ นั่นคือ เส้นลอเรนซ์จะต้องอยู่ใต้เส้นทแยงมุม ส่วน $\alpha > 0$ และ $\beta > 0$ นั้น ก็หมายความว่า η จะมีค่าเท่ากับ 0 เมื่อ $\pi = 0$ หรือ $\pi = \sqrt{2}$

ตามลักษณะการกระจายรายได้ที่แท้จริง ค่า F ก็คือ เปอร์เซนต์ หรือความถี่สะสม (cumulative percentages or frequencies) ของจำนวนครัวเรือนที่ได้รับรายได้ตามชั้นของรายได้ต่างๆกัน (f) และค่า F_1 คือเปอร์เซนต์หรือความถี่สะสมของรายได้ของครัวเรือนทั้งหมดตามชั้นของรายได้ต่าง ๆ กัน (Y) ดังนั้นค่าของ η และ π ตามสมการ (2) สามารถแทนค่าได้จากข้อมูลการกระจายรายได้ได้ดังนี้ คือ

$$\left(\frac{f-y}{\sqrt{2}}\right) = a\left(\frac{f+y}{\sqrt{2}}\right)^{\alpha}\left[\sqrt{2} - \left(\frac{f+y}{\sqrt{2}}\right)\right]^{\beta} \qquad (3)$$

หรือในรูปของ logarithms เพื่อคำนวณหาค่า parameters a, α และ β ตามวิธี least squares method ทำได้ดังนี้

$$\log\left(\frac{f-y}{\sqrt{2}}\right) = \log a + \alpha \log\left(\frac{f+y}{\sqrt{2}}\right) + \beta \log\left[\sqrt{2} - \left(\frac{f+y}{\sqrt{2}}\right)\right] \quad (4)$$

สมการ (4) ไม่ได้รวมเอาความผันแปรสุ่ม (random disturbance) เข้าไว้ด้วย โดยสมมุติว่ามีการกระจายปกติ

โดยที่ค่า $\propto$ และ β สามารถจะประมาณค่าได้ โดยวิธีกำลังสองน้อยที่สุด (Least Square Estimation) โดยสมมุติค่าตัวแปรในสมการว่า

$$\log\left(\frac{f-y}{\sqrt{2}}\right) = y_1 \quad (5)$$

$$\log\left(\frac{f+y}{\sqrt{2}}\right) = y_2 \quad (6)$$

$$\log\left[\sqrt{2}-\left(\frac{f+y}{\sqrt{2}}\right)\right] = y_3 \quad (7)$$

เมื่อได้ทราบค่าสัมประสิทธิ์ (Coefficient) $\propto$ และ β ของเส้นลอเรนซ์แล้วการคำนวณหาค่าสัมประสิทธิ์จินี่ก็ทำได้ โดยการอินทิเกรทสมการเส้นลอเรนซ์ จากค่า 0 ถึง $\sqrt{2}$ ก็จะได้เนื้อที่ระหว่างเส้นทะแยงมุมกับเส้นลอเรนซ์ สองเท่าของเนื้อที่นี้ก็คือค่าสัมประสิทธิ์จินี่ โดยใช้สูตรว่า

$$G = 2\int_0^{\sqrt{2}} a\pi^{\alpha}(\sqrt{2} - \pi)^{\beta} d\pi \quad (8)$$

$$= 2a(\sqrt{2})^{1+\alpha+\beta} B(1+\alpha, 1+\beta) \quad (9)$$

โดย $B(1+\propto, 1+\beta)$ คือค่าปฏิพันธ์เบต้า (Beta function) ซึ่งคำนวณขึ้นใหม่ได้โดยผ่าน gamma function คือ

$$B(1+\propto, 1+\beta) = \frac{\gamma(1+\propto)\ \gamma\ (1+\beta)}{\gamma\ (2+\propto+\beta)}$$

จากตารางเบต้าที่คำนวณขึ้นมาได้นั้น ค่า α และ β ที่คำนวณได้ไม่ตรงกับค่าในตารางที่มีอยู่ทั่วไป ดังนั้นจึงได้ใช้วิธีการอินทิเกรท gamma function ด้วยวิธีการดังต่อไปนี้

$$\Gamma(x) = \left.\frac{t^{x}.e^{-t}}{x}\right|_{0}^{\infty} + \frac{1}{x}\int_{0}^{\infty} e^{-t}.t^{x}dt$$ [3]

และ

$$\Gamma(x) = \frac{\Gamma(x+1)}{x}$$

หรือ

$$x\Gamma(x) = \Gamma(x+1)$$

โดย $0 \leqslant x < 2$

เมื่อได้ใช้สูตรที่ 2 ทำการวิเคราะห์ โดยวิธีลัด และสูตรที่ 3 โดยวิธี Regression ก็สามารถหาค่าสัมประสิทธิ์จินี่ต่อไปได้

อย่างไรก็ตามในการวิเคราะห์ข้อมูลแบบตัดขวาง (Cross Sectional data) ตามที่ได้ผลดังกล่าวมาแล้วทั้งหมดนี้บางครั้ง ข้อสมมุติอาจไม่เป็นจริงก็ได้ เพราะความแปรปรวนของตัวรบกวน (disturbance terms) อาจไม่เท่ากัน ซึ่งกล่าวอีกนัยหนึ่งเรียกว่าเกิด heteroscedasticity ขึ้นมาก็เป็นได้ การเกิด heteroscedasticity นี้จะทำให้ผลของการวิเคราะห์โดยใช้วิธี Ordinary Least Squares Method นั้นทำให้ได้ค่าของตัวสัมประสิทธิ์ที่ประมาณออกมานั้นไม่มี efficiency ซึ่งเป็นคุณสมบัติที่ไม่เป็นที่ต้องการ แม้ว่าผลที่ได้จะยังคงเป็น

---

[3] C.Ray Wylie and Louis C.Barret, Advanced Engineering Mathematics, 5^{th} edition, Mc.graw-Hill International Book Company, 1982, chapter 8 pp.401-421.

Unbiased อยู่ก็ตาม และนอกจากนี้แล้วเมื่อเกิด heteroscedasticity แล้ว และถ้ายังคงใช้สูตรการคำนวณหาค่าความแปรปรวน ของค่าประมาณตัวสัมประสิทธิ์ตามสูตรเดิมที่ปรากฏ ก็มักจะทำให้ค่าประมาณของความแปรปรวนของค่าตัวสัมประสิทธิ์เหล่านี้มีค่าที่ต่ำกว่าความเป็นจริงได้และจะมีผลทำให้ค่า t-ratios สูงไปกว่าความเป็นจริงด้วย และทำให้เกิดผลของการสรุปที่ผิดพลาดได้ และค่า t-ratios ดังกล่าวก็จะเป็นค่าt-ratios ที่ไม่ถูกต้องอีกด้วยดังนั้นการใช้ตารางt-students กับค่า t-ratios ที่คำนวณมาโดยวิธีดังกล่าวข้างต้นก็จะเป็นเรื่องที่ไม่น่าจะถูกต้อง

ด้วยเหตุผลทั้งหมดนี้จึงควรจะมีการตรวจสอบข้อสมมุติเกี่ยวกับ homoscedasticity อย่างละเอียด ถ้าหากมี heteroscedasticity เกิดขึ้นก็ควรจะใช้วิธี Estimate Generalized Least Squares มากกว่าใช้วิธีOrdinary Least Squares Method ที่กล่าวมาแล้วนั้น[4]

การตรวจสอบ Heteroscedasticity

ในการคำนวณหาค่าตัว parameter a, α, β ตามวิธีกำลังสองน้อยที่สุด (Ordinary Least Square Method) นั้น มีข้อสมมุติอยู่ประการหนึ่งว่าความแปรปรวน (variance) ของตัวรบกวนทุกตัว (disturbance terms) จะต้องมีค่าคงที่ หรือจะต้องเป็น homoscedasticity นั่นคือให้

---

[4] Johnston, J.Econometric Method 2nd ed. New York: Mc-Graw-Hill, 1972. and Judge, G.G., R.C.Hill, W.E Griffiths, H.Lutkepohl, and T-C.Lee, Introduction to the Theory and practice of Econometrics, New York:John Willey and Son, 1982

$$\sigma_i^2 = \sigma_o^2$$

โดย σ_o^2 = ค่าความแปรปรวนที่คงที่ ค่าความแปรปรวนที่คงที่ของตัวรบกวน

σ_i^2 = ความแปรปรวนของตัวรบกวนของตัวแปรตัวที่ i

i = 1, 2,..., n

n = จำนวนตัวอย่าง

ข้อมูลแบบตัดขวาง (Cross-section data) เป็นข้อมูลที่เก็บในเวลาเดียวกันแต่ต่างสถานที่กัน ดังนั้นข้อสมมุติดังกล่าวมักจะไม่ค่อยเป็นจริง นั่นคือความแปรปรวนของตัวรบกวนมักจะไม่เท่ากัน หรือเรียกว่าเกิด heteroscedasticity ขึ้น ผิดกับข้อมูล time series data ซึ่งเก็บจากสถานที่เดียวกันหรือตัวแปรต่างๆ เป็นชุดเดียวกันแต่ต่างเวลากัน ดังนั้นแม้เวลาจะผ่านไปแต่ตัวแปรต่าง ๆ เหล่านี้ก็ยังมีระบบ หรือระเบียบคล้าย ๆ กันอยู่ ค่าการกระจายของแต่ละกลุ่มจึงไม่ต่างกันมาก นั่นคือความแปรปรวนของตัวรบกวนทุกตัวจะมีค่าคงที่

ในกรณีของการใช้ข้อมูลแบบตัดขวางทำการวิเคราะห์ในกรณีศึกษานี้ ถ้าเกิด heteroscedasticity และยังคงใช้Ordinary Least Square Method อยู่ ก็จะทำให้ค่าที่ประมาณได้ไม่มี efficiency อันเป็นคุณสมบัติที่ไม่เป็นที่ต้องการ แม้ค่าจะ unbiased อยู่ก็ตาม และนอกจากนี้แล้ว เมื่อเกิด heteroscedasticity ขึ้น และยังใช้สูตรการคำนวณหาความแปรปรวนของค่าประมาณการหาสัมประสิทธิ์ตามสูตรเดิมอยู่มักจะทำให้ค่าประมาณการของความแปรปรวนของค่าสัมประสิทธิ์ต่ำไปกว่าความเป็นจริง ซึ่งจะมีผลทำให้ค่า t-ratios สูงไปกว่าความเป็นจริง และทำให้ได้ข้อสรุปที่ผิดพลาด และค่า t-ratios ที่ได้ก็เป็นค่าที่ไม่ถูกต้อง ดังนั้นการใช้ตาราง t-student กับค่า t-ratios ที่คำนวณได้โดยวิธีนี้จึงไม่น่าจะถูกต้องด้วย

ในการที่จะพิจารณาว่า สมการถดถอยที่ถูกสร้างขึ้นมาใช้นั้นเกิดปัญหา heteroscedasticity หรือไม่ทำได้หลายวิธี เช่น ใช้วิธีการง่าย ๆ โดยการเขียนแผนภาพการกระจาย (Scatter diagram) ของค่าความคลาดเคลื่อน (residual) ยกกำลังสอง ตามค่าของตัวแปรอิสระบางตัว ที่สงสัยว่าจะมีความสัมพันธ์กับ residual term ถ้าลักษณะของกราฟของ residual ไม่เป็นเหมือนเส้นขนานแต่อาจเป็นรูปตัว V หรือ กลมรี ซึ่งแสดงให้เห็นถึงความสัมพันธ์อย่างมีระบบ ก็แสดงว่าอาจเกิดปัญหา heteroscedasticity ขึ้นได้ หรือใช้วิธีทดสอบวิธีต่างๆ เช่น Spearman's rank correlation, Glejser, Bartlett, Ramsey's Bamset, Breusch-Pagan, หรือ Goldfeld and Quandt test ก็ได้

สำหรับกรณีศึกษาครั้งนี้จะใช้วิธีทดสอบของ Glejser และ Goldfeld and Quandt Test