



Panel data analysis

มนตรี พิริยะกุล

บทคัดย่อ

ในการวิเคราะห์การถดถอยจากหลายหน่วยสำรวจ (cross section) ที่แต่ละหน่วยสำรวจมีข้อมูลเรียงตามเวลา (time series) เรียกว่า panel data นั้นจำเป็นต้องใช้เทคนิคการวิเคราะห์ที่แตกต่างจาก OLS คือต้องมีการควบคุมปัจจัยนอกสมการการถดถอยเอาไว้ หากตัวแปรภายนอกไม่ผันแปรตามเวลาและสัมพันธ์กับตัวแปรอิสระให้ใช้วิธีวิเคราะห์แบบ fixed effect regression model (FEM) หากตัวแปรภายนอกไม่สัมพันธ์กับตัวแปรอิสระให้ใช้วิธีวิเคราะห์แบบ random regression model (REM หรือ ECM)

Abstract

In cases that data were collected in cross section and time series mixed, we need some special algorithm differ from OLS. In that context, controlling for time-invariant variables that correlate to independent variables the fixed effect regression model (FEM) is needed otherwise random effect regression model (REM) should be employed.

คำสำคัญ: ตัวแบบการถดถอยอิทธิพลคงที่ ตัวแบบการถดถอยอิทธิพลสุ่ม ตัวแปรที่มีอิทธิพลคงที่ตามเวลา

Keyword: Fixed Effect Regression Model, Random Effect Regression Model, time invariant variable

1.ความนำ

Panel data analysis เรียกอีกอย่างหนึ่งว่า longitudinal data analysis คือการติดตามบันทึกข้อมูลของหน่วยสำรวจ (individual หรือ entity หรือ subject) หลายหน่วยติดต่อกันไปหลายปี (หรือหน่วยเวลาใด ๆ) เป็นการวิเคราะห์ข้อมูลชนิดพหุมิติคือข้อมูลชนิดดังกล่าวอาจมีตัวแปรอิสระที่ตัวก็ได้แต่เป็นข้อมูลที่บันทึกเรียงตามเวลา (multiple time periods) ของแต่ละหน่วยสำรวจ หน่วยสำรวจอาจเป็นพื้นที่ ประเทศ จังหวัด เมือง หรือเขตปกครองระดับใดๆ องค์การ สถานประกอบการ คน สัตว์ สิ่งของ

¹ ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยรามคำแหง



Panel data เป็นการเก็บข้อมูลของหน่วยสำรวจ (เป็น cross section data) ซึ่งบันทึกค่า ตัวแปร ที่สนใจจากแต่ละหน่วย ณ จุดเวลา (point of time) หนึ่งที่กำหนด จากนั้นเก็บข้อมูลของหน่วยสำรวจ เดิมเพิ่มเติม ณ จุดเวลาถัดๆ ไป (คราวนี้เป็น time series) ซึ่งอาจเพิ่มในจำนวนที่เท่ากันหรือไม่เท่ากัน ในระหว่างหน่วยสำรวจก็ได้ ถ้าเพิ่มจำนวนจุดเวลาในจำนวนที่เท่ากันเรียกว่า balanced panel data เช่นข้อมูลทางการตลาดในตาราง 1 แต่ถ้าเพิ่มในจำนวนที่ไม่เท่ากันเรียกว่า unbalanced panel data เช่น ข้อมูลทางเศรษฐกิจสังคมของหน่วยสำรวจในตาราง 2 ต่อไปนี้

ตาราง 1 ข้อมูลทางการตลาด

ตำบล	ปี พ.ศ.	ราคา	ความต้องการซื้อต่อคน
ก	2555	75	2.0
ก	2556	85	1.8
ข	2555	50	1.0
ข	2556	48	1.1
ค	2555	60	1.5
ค	2556	65	1.4
ง	2555	55	0.8
ง	2556	60	0.7

ตาราง 2 ข้อมูลเศรษฐกิจสังคมของหน่วยวิเคราะห์

คนที่	ปี พ.ศ.	รายได้	เพศ	อายุ
1	2554	1600	1	23
1	2555	1500	1	24
2	2554	1900	2	41
2	2555	2000	2	42
2	2556	2100	2	43
3	2556	3300	1	34

จากตารางทั้ง 2 นักวิจัยคงสงสัยว่าจะทำการวิเคราะห์สมการถดถอยอย่างไร อาจมีข้อสงสัยในเรื่องวิธีการ คือจะวิเคราะห์ข้อมูลโดยใช้วิธีกำลังสองน้อยที่สุด (ordinary least square; OLS) ตรงๆ ได้หรือไม่ หรือว่าจะต้องทำอะไรก่อน

ความผันแปรที่เห็นในตารางเป็นความผันแปร 2 แบบคือความผันแปรภายในหน่วยสำรวจ (within group variation) และความผันแปรระหว่างหน่วยสำรวจ (across group variation) ความผันแปรภายในหน่วยสำรวจเป็นความผันแปรระหว่างข้อมูลภายในหน่วยสำรวจซึ่งแปรไปตามเวลา และความผันแปรระหว่างหน่วยสำรวจซึ่งแปรไปตามหน่วยสำรวจ โดยทั่วไปเมื่อพูดถึงความผันแปรระหว่างหน่วยสำรวจจะหมายถึงค่าเฉลี่ยของตัวแปรซึ่งมักจะผันแปรไปตามหน่วยสำรวจ



พิจารณาตาราง 1 เมื่อจะทำการศึกษาตามแบบการวิเคราะห์สมการถดถอยกับข้อมูลสำรวจ หรือข้อมูลภาคตัดขวาง (cross section) เฉพาะ พ.ศ. 2555 จะพบข้อมูลดังตาราง 3

ตาราง 3 ข้อมูลทางการตลาดภาคตัดขวางเฉพาะ พ.ศ. 2555 จากตาราง 1

ตำบล	ปี พ.ศ.	ราคา	ความต้องการซื้อต่อบุคคล
ก	2555	75	2.0
ข	2555	50	1.0
ค	2555	60	1.5
ง	2555	55	0.8

จากตารางจะพบว่าสมการถดถอยคือ Demand = -1.418 + 0.046 Price; $R^2 = 0.843$ จะสังเกตว่าอิทธิพลของราคาสินค้าที่มีต่อความต้องการซื้อเป็นไปในทางบวก คือราคาสินค้าสูงขึ้นความต้องการซื้อจะสูงขึ้นตาม เมื่อราคาลดลงความต้องการซื้อก็ลดลงตาม ซึ่งขัดกับหลักเกณฑ์ทางเศรษฐศาสตร์ คือไม่ควรเป็นไปได้ในบริบททั่วไป ที่เป็นเช่นนี้คงเป็นเพราะมีปัจจัยสำคัญอื่นอีกแต่ไม่ได้รับการกำหนดให้เป็นตัวแปรอิสระแอบส่งผลกระทบหรือสัมพันธ์กับตัวแปรอิสระ เราเรียกสถานการณ์นี้ว่า omission variable bias การจัดรูปแบบเพิ่มข้อมูลเสียใหม่ให้เป็นลักษณะ panel data อาจทำให้ความสัมพันธ์ระหว่างตัวแปรถูกต้องขึ้นและอาจแก้ปัญหา omission variable bias ได้ และการจัดให้มี within group variation ทำให้ตัวแปรอื่นที่อาจมีผลกระทบต่อตัวแปรอิสระถูกกักให้มีอิทธิพลเฉพาะภายในหน่วยสำรวจ ไม่ข้ามไปในระหว่างหน่วยสำรวจ

2. วิธีวิเคราะห์ข้อมูล

วิธีวิเคราะห์ panel data ทำได้หลาย 3 วิธีคือ 1) วิธี Pooled OLS Regression ซึ่งเป็นการวิเคราะห์โดยใช้ OLS ตรง ๆ 2) วิธี Fixed Effect Regression Model (FEM) กรณี FEM นี้ตัวแปรที่ถูกมองข้าม (omit) ต้องมีอิทธิพลคงที่อยู่เฉพาะภายในหน่วยสำรวจ ไม่แปรไปตามเวลา (time-invariant) เช่นคุณภาพการบริการ (SERVQUAL) เป็นปัจจัยที่แตกต่างกันในระหว่างองค์การและมีอิทธิพลที่คงที่ต่อค่าบริการในสมการ Demand = $\beta_0 + \beta_1 \text{Price} + u$ อยู่ตลอดเวลา 3) วิธี random effect model (REM)

ข้อตกลงเรื่อง time-invariant ของตัวแปรอิสระที่อยู่นอกสมการทำให้เราต้องวิเคราะห์สมการถดถอยที่เรียกว่า FEM โดย FEM จะสนใจความผันแปรตามเวลาภายในหน่วยสำรวจ คำว่าสนใจหมายถึงการหาทางควบคุมปัจจัยภายนอก คำว่าการควบคุมหมายถึงการกระทำใดๆที่ทำให้ปัจจัยภายนอกอยู่ในกรอบที่ไม่มาก่อวนผลการวิเคราะห์ข้อมูล อาจกระทำโดยการกำจัดอิทธิพลของ



ปัจจัยภายนอกหรือนำปัจจัยภายนอกมาร่วมเป็นตัวแปรอิสระหนึ่งทีอาจเสนอในรูปแบบอื่น ๆ คือ ผ่านกระบวนการตัวแปรหุ่น (dummy variable) ไม่เสนอตรง ๆ เพราะแม้ว่าตัวแปรภายนอกจะมีค่า มีอิทธิพลแต่ก็แฝงตัวอยู่ทำให้วัดค่าโดยตรงไม่ได้

1) วิธี Pooled OLS Regression

วิธี Pooled OLS Regression เป็นวิธีวิเคราะห์การถดถอยที่ไม่สนใจว่าหน่วยสำรวจ (cross section unit) จะได้รับผลกระทบจากปัจจัยภายนอกเฉพาะตัวที่ความแตกต่างกันหรือไม่และบันทึกข้อมูลตามเวลา (time series data) ใวันานแตกต่างกันเพียงใด เช่นการศึกษาผลประกอบการของ สายการบิน 10 สายการบิน แต่ละสายการบินบันทึกข้อมูลผลประกอบการ (total cost, C) ต้นทุนค่าน้ำมัน (fuel price, PF) น้ำหนักบรรทุก (loading factor, LF) และอื่นๆ เอาไว้ การบันทึกข้อมูลของแต่ละสายการบินอาจบันทึกด้วยระยะเวลาเท่าหรือไม่เท่ากันก็ได้ จากนั้นการวิเคราะห์กระทำตรง ๆ เสมือนเป็นข้อมูลชุดเดียวกัน คือมีสัมประสิทธิ์การถดถอยร่วมกัน ข้อตกลงการถดถอยยังคงเป็นแบบเดิมคือไม่มีความคลาดเคลื่อนจากการวัด (no error in variable) แล้ววิเคราะห์ข้อมูลด้วยวิธี OLS โดยตรง การวิเคราะห์ด้วยวิธี OLS โดยตรงเป็นการวิเคราะห์เราที่ไม่สนใจว่าหน่วยสำรวจแต่ละหน่วยจะได้รับผลกระทบจากปัจจัยภายนอกเฉพาะตัว (individual specific effect) ต่างกันหรือไม่ เช่นในสมการ (1) ต่อไปนี้ $C_{it} = \beta_{0i} + \beta_{1i} O_{it} + \beta_{2i} PF_{it} + \beta_{3i} LF_{it} + u_{it}$, $i = 1, 2, \dots, N$; $t = 1, 2, \dots, T_i$ พบว่าความเชี่ยวชาญของพนักงาน ความเป็นสถาบันขององค์กร คุณภาพบริการ คุณภาพผลิตภัณฑ์ของแต่ละสายการบินมีความแตกต่างกันและเป็นปัจจัยที่คงที่เฉพาะองค์กรอีกทั้งไม่แปรไปตามเวลา เช่น สายการบิน ก ย่อมมีพนักงานที่เชี่ยวชาญเฉพาะตัวและความเชี่ยวชาญก็ไม่แปรค่าไปตามเวลา เรียกว่า time invariant สายการบินอื่นก็มีความเชี่ยวชาญเฉพาะของตนเอง และแตกต่างกันไปในระหว่างสายการบิน การถือว่าสัมประสิทธิ์ β 's เป็นค่ารวม (pooled) จึงเป็นการมองว่าสายการบินทุกสายได้รับผลกระทบจากปัจจัยภายนอกเฉพาะตัว (individual specific effect) แบบเดียวกันซึ่งไม่น่าจะถูกต้อง สิ่งทีอาจเป็นภาพลวงตาที่พบจากผลการวิเคราะห์คือ R^2 มีค่าสูง ตัวแปรที่มีนัยสำคัญที่ระดับสูง (คือ t-statistics มีค่าสูงมาก หรือ p-value มีค่าต่ำมาก) และ DW (Durbin-Watson statistics) มีค่าต่ำ ซึ่งเป็นสัญญาณว่ากำลังมีปัญหา specification error คือตัวแปรอิสระอื่นที่ไม่ได้กำหนดไว้ในสมการถดถอยถูกรวมเข้ากับ ความคลาดเคลื่อนจากการวัดเกิดเป็นตัวแปรสุ่มอีกตัวหนึ่งเรียกว่า error term (คือ u) ผลที่ตามมาคือเกิดความเชื่อมโยงระหว่าง u กับตัวแปรอิสระเรียกว่า error in variable และตัวแปร u อาจสัมพันธ์กันเอง เรียกว่า serial correlation ผลของ error in



variable คือค่าประมาณของ β 's มีความเอนเอียงและความเอนเอียงก็จะไม่หายไปแม้จะใช้ตัวอย่างขนาดใหญ่ (inconsistency) (มนตรี พิริยะกุล, 2544)

ตัวแปรที่เรียกว่า time invariant variable คือตัวแปรที่มีค่าคงที่เสมอไม่ว่าเวลาจะแปรเปลี่ยนไปอย่างไรและวัดค่าไม่ได้เพราะแฝงตัวอยู่นอกสมการ อีกทั้งหน่วยสำรวจ (cross section unit) ที่ต่างกันอาจได้รับอิทธิพลจากตัวแปรที่แฝงตัวอยู่นอกสมการถดถอยคนละตัวกัน (unobserved heterogeneity) ตัวอย่างของตัวแปรอิสระที่แฝงตัวอยู่นอกสมการถดถอยนอกเหนือไปจากความเชี่ยวชาญของพนักงาน ความเป็นสถาบันขององค์กร คุณภาพของบริการ คุณภาพผลิตภัณฑ์ แล้วยังเป็นลักษณะของสถานประกอบการ (คือเป็นภาคเอกชนหรือภาครัฐ) ความเป็นธุรกิจครอบครัว หรือธุรกิจที่ไม่ใช่แบบธุรกิจครอบครัว ผู้นำสูงสุดเป็นชายหรือหญิง ผู้นำสูงสุดมีเชื้อชาติ สัญชาติใด ก่อตั้งมานานจนเป็นสถาบันหรือเพิ่งก่อตั้งได้ไม่นาน มีทรัพยากรที่หายากหรือไม่ (เช่น ความเชี่ยวชาญ เงินทุน ทุนมนุษย์ วัฒนธรรมด้านต่าง ๆ ความสามารถในการบริหาร เครือข่ายทางการค้าคือ คู่ค้าและลูกค้า) ประวัติศาสตร์ที่ยาวนานของการต่อสู้ฟันฝ่าในธุรกิจการบิน ตัวแปรเหล่านี้เป็นปัจจัยเฉพาะตัวของหน่วยวิเคราะห์ที่ไม่แปรค่าตามเวลา ทั้งยังสัมพันธ์กับตัวแปรอิสระ (Gujarati and Porter, 2009, p. 593) ตัวอย่างเช่น

ให้ C_{it} = ต้นทุนรวมการบินของสายการบินที่ i ; $i = 1, 2, \dots, N$ ในช่วงเวลา t ; $t = 1, 2, \dots, T_i$ (ตามตัวอย่าง $N = 6$ สายการบิน $t = 1970-1984$ คือบันทึกข้อมูลตามเวลา 15 ปี)

O_{it} = ผลผลิตในเทอมของรายได้จากค่าโดยสาร (output)

PF_{it} = ราคาน้ำมัน (fuel price)

LF_{it} = ค่าเฉลี่ยปริมาณสินค้าที่ขนส่ง (load factor)

สมการ panel analysis ของสายการบิน N สายแต่ละสายบันทึกข้อมูลสายละ T_i หน่วยเวลา คือ

$$C_{it} = \beta_{0i} + \beta_{1i}O_{it} + \beta_{2i}PF_{it} + \beta_{3i}LF_{it} + u_{it}, i=1, 2, \dots, N; t=1, 2, \dots, T_i \quad \dots (1)$$

เมื่อวิเคราะห์ด้วย pooled OLS regression จะได้สมการประมาณค่าเป็น

$$C_{it} = \hat{\beta}_0 + \hat{\beta}_1 O_{it} + \hat{\beta}_2 PF_{it} + \hat{\beta}_3 LF_{it} \quad \dots (2)$$

ซึ่งก็คือสมการที่มีสัมประสิทธิ์ β 's คงที่ แสดงว่าสายการบินแม้ต่างสายกันต่างก็มีปัจจัยด้านผลผลิต ราคาน้ำมัน ค่าเฉลี่ยบรรทุกทุกสินค้าไม่ต่างกันซึ่งไม่น่าจะจริง



ความสัมพันธ์ระหว่างตัวแปรภายนอกสมการกับตัวแปรอิสระภายในสมการส่งผลกระทบไปถึงตัวแปรตามด้วย เช่น เครือข่ายการค้าและความสามารถในการจัดการสามารถผูกพันกับคู่ค้านำมันให้จำหน่ายน้ำมันในราคามิตรภาพ ผูกพันลูกค้าให้ขนส่งผลิตภัณฑ์กับตนตลอดไป ทำให้มีต้นทุนต่างจากสายการบินอื่นที่มีความสามารถทางการบริหารหรือมีทรัพยากรที่แตกต่าง ตัวแปรพวกนี้แม้ทราบว่ามีอยู่แต่การจะนำมาเสนอให้เห็นตัวเป็นสิ่งที่ไม่ควรกระทำเนื่องจากวัดค่าไม่ได้โดยตรง เช่น ไม่ควรนำความสามารถในการจัดการ (M_i) เข้าบรรจุในสมการ(1) เป็นสมการ (3) ต่อไปนี้

$$C_{it} = \beta_0 + \beta_1 O_{it} + \beta_2 PF_{it} + \beta_3 LF_{it} + \beta_4 M_{it} + u_{it} \quad \dots (3)$$

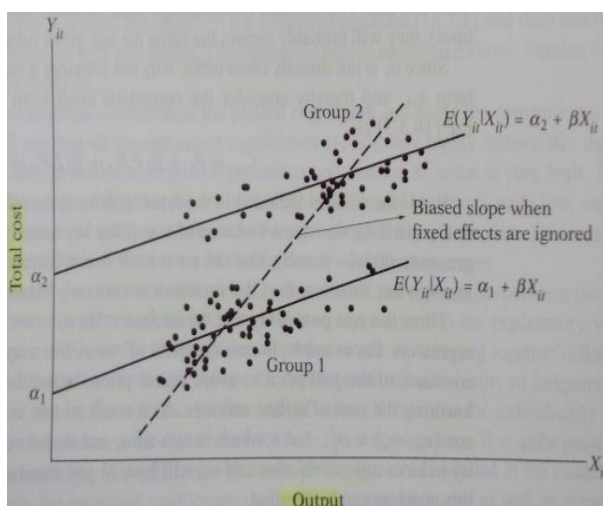
แต่ควรเสนอในรูปอิทธิพลแอบแฝงเฉพาะ(unobserved individual specific effect, α_i) เป็น

$$C_{it} = \beta_0 + \beta_1 O_{it} + \beta_2 PF_{it} + \beta_3 LF_{it} + \alpha_i + u_{it} \quad \dots (4)$$

แต่ α_i ไม่อาจวัดค่าได้จึงกำหนดให้ $v_{it} = \alpha_i + u_{it}$ จะได้สมการใหม่เป็น

$$C_{it} = \beta_0 + \beta_1 O_{it} + \beta_2 PF_{it} + \beta_3 LF_{it} + v_{it} \quad \dots (5)$$

คราวนี้เราจะเห็นว่าเราทำการพิจารณาถี่ถ้วนแล้วก็นำ individual specific effect มาบรรจุในสมการด้วยแล้ว แต่ α_i คือลักษณะเฉพาะตัวขององค์การซึ่งสัมพันธ์กับตัวแปรอิสระดังกล่าวแล้ว เมื่อมาจัดรวมอยู่กับ u_{it} กลายเป็นตัวแปรสุ่มใหม่คือ v_{it} ทำให้ได้สมการ (5) ที่ขัดแย้งกับข้อตกลงเรื่อง error in variable คือ $E(v_{it}O_{it}) \neq 0$, $E(v_{it}PF_{it}) \neq 0$, $E(v_{it}LF_{it}) \neq 0$ และเรายังพิสูจน์ได้อีกว่า $E(v_{it}v_{is}) \neq 0$; $t \neq s$ อีกทั้ง α_i ยังไปเชื่อมโยง u_{it} เข้ากับ u_{is} ; $t \neq s$ จึงเกิดปัญหา autocorrelation ตามมาอีกด้วย การใช้



pooled OLS Regression ทำให้ได้ภาพสมการคล้ายภาพด้านซ้ายมือนี้ สมการที่เสนอในรูปเส้นประคือสมการที่ประมาณค่ามาจากวิธี pooled OLS Regression แม้จะพุ่งผ่านไปในกลุ่มข้อมูลแต่ก็ยังมองเห็นว่าผ่านได้ไม่แนบเนียนเพราะแต่ละกลุ่มของข้อมูลตามหน่วยสำรวจที่ต่างกันนั้นจะสามารถนำเสนอสมการตามข้อมูลภาคตัดขวางได้ แต่วิธีวิเคราะห์แบบ pooled OLS Regression ก็มีได้ใส่ใจเรื่องนี้

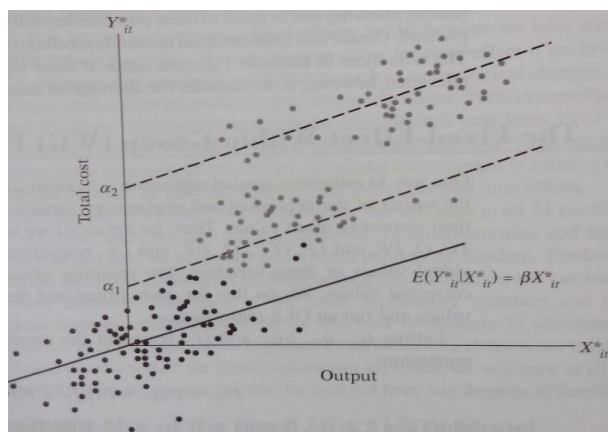
ทำให้ผลการศึกษาเอนเอียง เรียกว่า heterogeneity bias

2) Fixed Effect Least Square Dummy Variable (LSDV)

จากการที่ pooled OLS มีข้อโต้แย้งเรื่องความเหมาะสมเพราะไม่พิจารณาความผันแปรตามเวลาและความผันแปรตามหน่วยสำรวจ เราจึงต้องให้ความสนใจประเด็นโต้แย้งทั้ง 2 ด้วย ในกรณีที่มีตัวแปรที่แสดงลักษณะเฉพาะของหน่วยสำรวจ (คือคงที่ภายในหน่วยสำรวจ แต่ก็แปรไป ในระหว่างหน่วยสำรวจ เรียกว่า subject effect หรือ individual effect) สามารถนำเสนอสมการถดถอยได้ดังนี้ จากสมการ

$$C_{it} = \beta_{it} + \beta_2 Q_{it} + \beta_3 PF_{it} + \beta_4 LF_{it} + u_{it}; i = 1, 2, \dots, N; t = 1, 2, \dots, T_i \quad \dots (6)$$

สังเกตว่าค่าคงที่ในสมการนี้คือ β_{it} ซึ่งแสดงให้เห็นว่าหน่วยสำรวจมีลักษณะเฉพาะที่ต่างกันไปตามเวลา (แต่ถ้าใช้ β_{it} จะเป็น time variant) แต่สัมประสิทธิ์ $\beta_2, \beta_3, \beta_4$ คงที่แสดงว่าตัวแปร Q, PF และ LF มีอิทธิพลร่วมกันในทุกหน่วยสำรวจแสดงให้เห็นได้ด้วยภาพเส้นตรง



(หรืออาจเป็น surface) วางขนานกัน โดยวางสูงต่ำจากแกนนอนหรือพื้นต่างกันไปตามคล้าย ๆ กับภาพนี้ จะเห็นว่าแต่ละสมการต่างก็มี intercept ของตน แสดงว่าแต่ละหน่วยสำรวจภาคตัดขวางมีอิทธิพลเฉพาะตัวที่ไม่แปรตามเวลา (unobserved time-invariant) เช่น ปรัชญาการบริหาร องค์กร โครงสร้างของคณะผู้บริหาร

บุคลิกภาพของ CEO เพศของ CEO ซึ่งกรณีนี้เรียกว่า Fixed Effect Regression Model (FEM) วิเคราะห์สมการถดถอยที่เหมาะสมสำหรับกรณีนี้คือการควบคุมอิทธิพลเฉพาะตัวที่ไม่แปรตามเวลานั้นเอาไว้ ซึ่งอาจทำได้ 2 วิธีคือ การกำหนดตัวแปรหุ่นแทนหน่วยสำรวจหรือการกำจัดอิทธิพลคงที่เฉพาะตัวนั้นเสียก่อน

ตามตัวอย่างเรามีหน่วยสำรวจทั้งสิ้น 6 หน่วย ให้เรากำหนดตัวแปรหุ่น 6 ตัว คือ $D_1 = 1$ ถ้าค่าสังเกตที่ปรากฏอยู่นั้นเป็นค่าสังเกตที่มาจากหน่วยสำรวจที่ 1 และ $D_1 = 0$ ถ้าค่าสังเกตค่าสังเกตที่ปรากฏอยู่นั้นเป็นค่าสังเกตที่มาจากหน่วยสำรวจอื่น $D_2 = 1$ ถ้าค่าสังเกตที่ปรากฏอยู่นั้นเป็นค่าสังเกตที่มาจากหน่วยสำรวจที่ 2 และ $D_2 = 0$ ถ้าค่าสังเกตค่าสังเกตที่ปรากฏอยู่นั้นเป็นค่าสังเกตที่มาจากหน่วยสำรวจอื่น ดังนี้เรื่อยไปจนถึง D_6 สมการ (6) จะเปลี่ยนเป็น



$$C_{it} = \beta_2 Q_{it} + \beta_3 PF_{it} + \beta_4 PL_{it} + \alpha_1 D_{1it} + \alpha_2 D_{2it} + \alpha_3 D_{3it} + \alpha_4 D_{4it} + \alpha_5 D_{5it} + u_{it} + \alpha_6 D_{6it} + u_{it}$$

$$; i = 1, 2, \dots, N; t = 1, 2, \dots, T_i \quad \dots (7)$$

ในที่นี้เราใช้แนวความคิดว่าให้นำ $D_1, D_2, \dots, D_6$ มาแทนความหมายของ β_{1i} ผลการตรวจสอบสมมติฐานว่า $\alpha_1, \alpha_2, \dots, \alpha_6$ ต่างจาก 0 หรือไม่ดูได้จากนัยสำคัญของ F-test คือ

$$F_c = \frac{(R_2^2 - R_1^2)/(p_2 - p_1)}{(1 - R_2^2)/[n - (p_2 - 1)]} \quad \dots (8)$$

R_2^2 = ค่า R^2 จากสมการ LSDV ในสมการ (7) R_1^2 = ค่า R^2 จากสมการ pooled OLS regression ในสมการ (1) p_2 = จำนวนพารามิเตอร์ใน LSDV จากสมการ (7) p_1 = จำนวนพารามิเตอร์ใน pooled OLS regression จากสมการ (1) และ n = จำนวนค่าสังเกต ถ้า F มีนัยสำคัญก็แสดงว่า $\alpha_i \neq \alpha_j; i \neq j$ หรือหน่วยสำรวจมีความแตกต่างกัน แสดงให้เราทราบว่าอย่าใช้ pooled OLS เพราะเป็นวิธีวิเคราะห์ที่ไม่ใส่ใจความแตกต่างระหว่างหน่วยสำรวจ

จากสมการ (7) ถ้าเราพิจารณาเห็นว่าสมการนี้แปรไปตามเวลาด้วย (เรียกว่า time effect) เพราะมีปัจจัยภายนอกที่วัดค่าสังเกตไม่ได้ส่งผลให้สมการแปรไปตามเวลา เช่นด้านเทคโนโลยี พบว่าการเปลี่ยนแปลงด้านเทคโนโลยีช่วยลดต้นทุนรวมของการบินลงทุกปี ทำให้ค่าใช้จ่ายด้านน้ำมันลดลงเรื่อยๆ หรือการเปลี่ยนด้านกฎหมายและการเปลี่ยนแปลงด้านนโยบายภาษีส่งผลกระทบต่อรายได้จากค่าโดยสาร ทำให้รายได้ของการบินลดลงอย่างต่อเนื่อง ให้เพิ่มตัวแปรหุ่นของเวลาเข้าไปในสมการ (7) สมมุติมีการบันทึกข้อมูล 15 ปี ก็ให้กำหนดตัวแปรหุ่น $T_1 = 1$ ถ้าเป็นปีที่ 1, $T_1 = 0$ ถ้าเป็นปีอื่น, $T_2 = 1$ ถ้าเป็นปีที่ 2, $T_2 = 0$ ถ้าเป็นปีอื่น ดังนี้เรื่อยไปจนถึง T_{15} สมการ (7) จะเปลี่ยนเป็น

$$C_{it} = \beta_2 Q_{it} + \beta_3 PF_{it} + \beta_4 PL_{it} + (\alpha_1 D_{1it} + \alpha_2 D_{2it} + \alpha_3 D_{3it} + \alpha_4 D_{4it} + \alpha_5 D_{5it} + \alpha_6 D_{6it}) + (\gamma_1 T_{1it} + \gamma_2 T_{2it} + \gamma_3 T_{3it} + \dots + \gamma_{15} T_{15it}) + u_{it} \quad \dots (9)$$

สมการ (7) เรียก one way FEM ขณะที่สมการ(9) เรียกว่า two way FEM โดยดูได้จากการเพิ่มจำนวนกลุ่มของตัวแปรหุ่นว่าเพิ่ม 1 กลุ่มหรือเพิ่ม 2 กลุ่ม

หากพิจารณาเห็นว่าหน่วยสำรวจต่าง ๆ นั้นล้วนมีลักษณะเฉพาะตน การกำหนดให้มีค่าสัมประสิทธิ์ถดถอยคงที่เท่ากันเป็น $\beta_2, \beta_3, \beta_4$ เสมอไปนั้นไม่ถูกต้อง สัมประสิทธิ์การถดถอยควรผันแปรไปได้ในระหว่างหน่วยสำรวจ ถ้าเป็นดังนี้ก็ให้กำหนดตัวแปรหุ่นให้แก่ตัวแปรอิสระ ดังนี้



$A_{1t} = 1$ ถ้า Q_{1t} เป็น ค่าสังเกตจากหน่วยสำรวจที่ 1 $A_{1t} = 0$ ถ้าเป็นค่าสังเกตจากหน่วยสำรวจอื่น

$A_{2t} = 1$ ถ้า Q_{2t} เป็นค่าสังเกตจากหน่วยสำรวจที่ 2 $A_{2t} = 0$ ถ้าเป็นค่าสังเกตจาก หน่วยสำรวจอื่น ดังนี้เรื่อยไปจนถึง A_{6t}

$B_{1t} = 1$ ถ้า PF_{1t} เป็นค่าสังเกตจากหน่วยสำรวจที่ 1 $B_{1t} = 0$ ถ้าเป็นค่าสังเกตจากหน่วยสำรวจอื่น

$B_{2t} = 1$ ถ้า PF_{2t} เป็นค่าสังเกตจากหน่วยสำรวจที่ 2 $B_{2t} = 0$ ถ้าเป็นค่าสังเกตจากหน่วยสำรวจอื่น ดังนี้เรื่อยไปจนถึง B_{6t}

$W_{1t} = 1$ ถ้า PL_{1t} เป็นค่าสังเกตจากหน่วยสำรวจที่ 1 $W_{1t} = 0$ ถ้าเป็นค่าสังเกตจากหน่วยสำรวจอื่น

$W_{2t} = 1$ ถ้า PL_{2t} เป็น ค่าสังเกตจากหน่วยสำรวจที่ 2 $W_{2t} = 0$ ถ้าเป็นค่าสังเกตจากหน่วยสำรวจอื่น ให้กำหนดดังนี้เรื่อยไปจนถึง W_{6t} เรียกสัมประสิทธิ์ของตัวแปรหุ่นเหล่านี้ว่า differential slope dummy coefficient และเมื่อผนวกตัวแปรหุ่นเหล่านี้ลงในสมการ(9) จะได้สมการ (10)

$$C_{it} = (a_1A_{1t}Q_{1t} + a_2A_{2t}Q_{2t} + a_3A_{3t}Q_{3t} + a_4A_{4t}Q_{4t} + a_5A_{5t}Q_{5t}) + (b_1B_{1t}PF_{1t} + b_2B_{2t}PF_{2t} + b_3B_{3t}PF_{3t} + b_4B_{4t}PF_{4t} + b_5B_{5t}PF_{5t}) + (w_1W_{1t}PL_{1t} + w_2W_{2t}PL_{2t} + w_3W_{3t}PL_{3t} + w_4W_{4t}PL_{4t} + w_5W_{5t}PL_{5t}) + (\alpha_1D_{2it} + \alpha_2D_{3it} + \alpha_3D_{3it} + \alpha_4D_{4it} + \alpha_5D_{5it} + \alpha_6D_{6it}) + (\gamma_1T_{2it} + \gamma_2T_{2it} + \gamma_3T_{3it} + \dots + \gamma_{15}T_{15it}) + u_{it} \quad \dots (10)$$

สังเกตจากสมการ LSDV จะพบว่ามีตัวแปรอิสระเพิ่มขึ้นมากกว่าเดิม มากน้อยเท่าไรขึ้นอยู่กับว่ามีเวลาและหน่วยสำรวจมากน้อยเพียงใด ผลที่ตามมาคือ degree freedom ลดลงมาก ค่า SEE จะใหญ่ขึ้นเพราะตัวหารเล็กลงทำให้ค่าของ t-statistics เล็กลงอาจมีผลให้สมมุติฐานมีแนวโน้มไม่มีนัยสำคัญ นอกเหนือการที่มีตัวแปรมากเกินไปย่อมเสี่ยงต่อปัญหาภาวะร่วมเส้นตรงพหุ (multicollinearity)

3) Fixed Effect within Group (WG) วิธี Demean

เนื่องจากมีปัจจัยที่คงที่ตามเวลา (time invariant) มีค่าแฝงอยู่ในสมการดังนั้นเมื่อเรานำค่าตัวแปรในแต่ละหน่วยสำรวจลบด้วยค่าเฉลี่ยของตัวแปรในแต่ละหน่วยสำรวจ ผลที่ได้ รับคือปัจจัยคงที่ ถูกกำจัด การลบด้วยค่าเฉลี่ยของตัวแปรในหน่วยสำรวจนั้นเรียกว่า demean นั่นคือจากสมการ (1) คือ

$$C_{it} = \beta_{1i} + \beta_2 Q_{it} + \beta_3 PF_{it} + \beta_4 PL_{it} + u_{it}; i = 1, 2, \dots, N; t = 1, 2, \dots, T_i \quad \dots (1)$$

$$\text{ดังนั้น } \bar{C}_i = \beta_{1i} + \beta_2 \bar{Q}_i + \beta_3 \bar{PF}_i + \beta_4 \bar{PL}_i + \bar{u}_i; i = 1, 2, \dots, N \quad \dots (11)$$



นำสมการ (11) ไปลบออกจากสมการ (1) จะได้สมการ (12) และ (13) ดังนี้

$$C_{it} - \bar{C}_i = \beta_2(Q_{it} - \bar{Q}_i) + \beta_3(PF_{it} - \bar{PF}_i) + \beta_4(PL_{it} - \bar{PL}_i) + (u_{it} - \bar{u}_i) \quad \dots (12)$$

$$\text{หรือ } c_{it} = \beta_2 q_{it} + \beta_3 pf_{it} + \beta_4 pl_{it} + u_{it}; i = 1, 2, \dots, N; t = 1, 2, \dots, T_i \quad \dots (13)$$

สมการ (12) และ (13) เป็นสมการที่ unobserved time-invariant variable ถูกกำจัด เมื่อนำไปวิเคราะห์จะได้ผลการศึกษาที่ค่าประมาณความชันไม่เอนเอียง และหากมีความเอนเอียงก็จะหายไปเมื่อขนาดตัวอย่างของค่าสังเกตใหญ่ขึ้น (เรียกว่า consistent)

อย่างไรก็ตามหากมี observed time-invariant ปนอยู่ในสมการที่ (1) เช่นในสมการการบริโภค (Keynesian consumption function) ซึ่งประกอบด้วย รายได้และการลงทุน หากมีตัวแปรอัตราดอกเบี้ยและอัตราการเติบโตทางเศรษฐกิจซึ่งค่อนข้างคงที่ในระยะสั้นร่วมเป็นตัวแปรอิสระอยู่ด้วยการดำเนินการตามวิธี demean หรือ differencing ที่จะกล่าวถึงต่อไปจะทำให้ตัวแปรทั้งสองคืออัตราดอกเบี้ยและอัตราการเติบโตที่เศรษฐกิจ (GDP growth) ถูกกำจัดออกไปจากสมการด้วย

4) Fixed Effect within Group วิธี Differencing

วิธีนี้ดำเนินการเพื่อกำจัด unobserved subject-specific characteristic ที่ไม่แปรตามเวลา ซึ่งก็คือ unobserved time-invariant variable เหมือนในวิธีที่ 3 โดยการลบกันระหว่างค่าสังเกตในเวลา $t+1$ กับเวลา t คือ

$$\Delta C_{it} = C_{it+1} - C_{it}, \Delta Q_{it} = Q_{it+1} - Q_{it}, \Delta PF_{it} = PF_{it+1} - PF_{it}, \Delta PL_{it} = PL_{it+1} - PL_{it} \\ ; i = 1, 2, \dots, N, t = 1, 2, \dots, T_i$$

แล้ววิเคราะห์สมการถดถอย

$$\Delta C_{it} = \beta_1 + \beta_2 \Delta Q_{it} + \beta_3 \Delta PF_{it} + \beta_4 \Delta PL_{it} + (u_{it+1} - u_{it}) \quad \dots (14)$$

ผลการวิเคราะห์สมการนี้แม้จะแก้ปัญหาได้ แต่ก็ยังมีปัญหาตามมาคือปัญหา autocorrelation และปัญหาที่ตัวแปรอิสระที่สังเกตได้และไม่แปรค่าตามเวลา (observed time-invariant) เช่น เพศ เชื้อชาติ ระยะทาง ขนาดพื้นที่ หรืออื่นๆที่ค่อนข้างคงที่และปรากฏเป็นตัวแปรอิสระร่วมอยู่ในสมการ ถูกกำจัดออกไปด้วยเช่นเดียวกับที่พบในวิธี Demean

5) Random Effect Model (REM) หรือ Error Component Model (ECM)

พิจารณาสมการต่อไปนี้

$$C_{it} = \beta_1 + \beta_2 Q_{it} + \beta_3 PF_{it} + \beta_4 PL_{it} + u_{it} \quad \dots (15)$$



พบว่าเป็นสมการที่มีค่า β_{1i} แปรไปตามหน่วยสำรวจ ($i = 1, 2, \dots, N$) ซึ่งเป็นสัมประสิทธิ์ที่แสดงถึงความเป็นตัวของตัวของหน่วยสำรวจเรียกว่า subject-specific characteristic ที่อาจสะท้อนมาจากศิลปะการบริหารจัดการองค์การ องค์ประกอบของคณะผู้บริหาร เพศผู้นำ นวัตกรรมของผู้นำ ทำให้ β_{1i} คงที่ภายในหน่วยสำรวจแต่ก็แปรไปในระหว่างหน่วยสำรวจ

ถ้าหากเรากล่าวว่า β_{1i} เป็นตัวแปรสุ่มที่สุ่มมาจากกลุ่มประชากรของหน่วยสำรวจที่มีค่าเฉลี่ยเท่ากับ β_1 ผลที่ตามมาคือ

$$\beta_{1i} = \beta_1 + \varepsilon_i \quad \dots (16)$$

ε_i คืออิทธิพลของตัวแปรเฉพาะหน่วยสำรวจดังกล่าวแล้ว โดยที่ ε_i เป็นตัวแปรสุ่มที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ σ_ε^2 แปลว่าองค์การที่ใช้เป็นตัวอย่งนั้นเป็นเพียงตัวอย่างที่สุ่มจากประชากรมีลักษณะร่วมกันคือ β_1 และผิดแผกกันไปโดยสุ่มเพราะ ε_i ดังนั้น สมการ (15) จึงเปลี่ยนเป็น

$$\begin{aligned} C_{it} &= (\beta_1 + \varepsilon_i) + \beta_2 Q_{it} + \beta_3 PF_{it} + \beta_4 PL_{it} + u_{it} \\ &= \beta_1 + \beta_2 Q_{it} + \beta_3 PF_{it} + \beta_4 PL_{it} + w_{it} \end{aligned} \quad \dots (17)$$

โดยที่ $w_{it} = \varepsilon_i + u_{it}$ และ ε_i กับ u_{it} มีคุณสมบัติตามข้อตกลงการถดถอยทุกประการ และที่สำคัญ w_{it} ต้องไม่เกี่ยวข้องกับใดๆ กับตัวแปรอิสระตามข้อตกลงเรื่อง error in variable แต่ ε_i อาจเกี่ยวข้องกับตัวแปรอิสระได้เพราะเป็นสิ่งที่เป็ผลของ unobserved subject specific characteristic หากเป็นเช่นนั้นก็ต้องแปลว่า w_{it} เกี่ยวข้องกับตัวแปรอิสระด้วย จึงเกิดเป็นปัญหา error in variable ผลก็คือค่าประมาณเอนเอียงและไม่หมดไปแม้จะใช้ตัวอย่างใหญ่ขึ้น (inconsistent) เรียกว่า w_{it} ว่า idiosyncratic term คือเป็นส่วนเหลือลักษณะเฉพาะคือผันแปรไปตามทั้งในภาคตัดขวางและในอนุกรมเวลาเพราะเกิดจากความคลาดเคลื่อนด้านอนุกรมเวลา (u_{it}) และความ คลาดเคลื่อนด้านตัดขวาง (ε_i)

การที่เราเรียกสมการ (17) ว่า ECM ก็เพราะ error term ในสมการเป็นสิ่งที่ประกอบขึ้นด้วยความคลาดเคลื่อนจากภาคตัดขวาง ε_i และความคลาดเคลื่อนจากอนุกรมเวลา u_{it} ดังกล่าว

ข้อสังเกตคือ เมื่อมาถึงจุดนี้ควรเข้าใจแล้วว่าใน FEM นั้นย่อมมีค่า β_{1i} ของตนเอง (fixed) เพราะอิทธิพลของปัจจัยเฉพาะตัว แต่ใน ECM (หรือ REM) นั้นเทอมคงที่จะเท่ากับ β_1 ซึ่งเป็นค่าเฉลี่ยจากประชากรของหน่วยสำรวจและมี ε_i ที่ทำให้แต่ละหน่วยสำรวจมีค่าเบี่ยงไปจากค่าเฉลี่ย (คือ β_1)

$$\text{(พิจารณา } w_{it} \text{ พบว่า } E(w_{it}) = E(\varepsilon_i + u_{it}) = E(\varepsilon_i) + E(u_{it}) = 0$$

$$\text{และ } V(w_{it}) = V(\varepsilon_i) + V(u_{it}) = \sigma_\varepsilon^2 + \sigma_u^2 \text{ (เนื่องจาก } E(\varepsilon_i u_{it}) = 0)$$



ซึ่งหาก $\sigma^2_{\varepsilon_i} = 0$ สมการ (17) จะลดรูปเป็นสมการ (15) $\sigma^2_{\varepsilon_i} = 0$ แปลว่าปัจจัยเฉพาะตัวของ หน่วยสำรวจ ไม่แตกต่างกัน ทำให้เราวิเคราะห์สมการ (17) ได้โดยวิธี pooled OLS regression โดยไม่ต้องสนใจเรื่อง fixed effect หรือ random effect เลย และเมื่อพิจารณา $\rho_{w_{it}w_{is}}$ จะพบว่า

$$\rho_{w_{it}w_{is}} = \frac{E(w_{it}w_{is})}{\sqrt{V(w_{it})}\sqrt{V(w_{is})}} = \frac{E[(\varepsilon_i + u_{it})(\varepsilon_i + u_{is})]}{V(w_{it})} = \frac{\sigma_{\varepsilon}^2}{\sigma_{\varepsilon}^2 + \sigma_u^2}$$

แสดงว่าเกิดปัญหา autocorrelation ขึ้น

3. การทดสอบ Hausman test และ Breusch-Pagan Lagrange Multiplier test

การทดสอบว่าสมการควรเป็น FEM หรือ REM สามารถกระทำได้ 2 วิธี คือทดสอบ ตาม Hausman test หรือทดสอบตาม Breusch-Pagan Lagrange Multiplier test ดังนี้

1) Hausman test ตัวทดสอบนี้กำหนดสมมติฐานไว้เป็น

H_0 : FEM = REM vs H_1 : ควรใช้ FEM หรือ

H_0 : w_{it} ไม่สัมพันธ์กับตัวแปรอิสระและตัวแปร time-invariant (REM)

vs H_1 : w_{it} สัมพันธ์กับตัวแปรอิสระและตัวแปร time-invariant (FEM)

ถ้ายอมรับ H_0 แสดงว่าเราอาจใช้ FEM หรือ REM ก็ได้ แต่ REM จะมี Efficient คือ variance ต่ำกว่า แต่ถ้าปฏิเสธ H_0 ให้ใช้ FEM

ให้ $\hat{\beta}_{FE}$ = เวกเตอร์ของค่าประมาณการถดถอยตามวิธี FEM

ให้ $\hat{\beta}_{RE}$ = เวกเตอร์ของค่าประมาณการถดถอยตามวิธี REM

ให้ $V(\hat{\beta}_{FE})$ = อะเรย์ของความผันแปรของค่าประมาณการถดถอยตามวิธี FEM

ให้ $V(\hat{\beta}_{RE})$ = อะเรย์ของความผันแปรของค่าประมาณการถดถอยตามวิธี REM

$Q = \hat{\beta}_{FE} - \hat{\beta}_{RE}$ และ T = ช่วงเวลา (ต้องเท่ากันทุกหน่วยสำรวจ)

$HT = TQ' [V(\hat{\beta}_{FE}) - V(\hat{\beta}_{RE})]^{-1} Q \sim \chi_k^2$ โดยที่ k = จำนวนพารามิเตอร์ในเวกเตอร์ Q

2) Breusch-Pagan Lagrange Multiplier (BP) ตัวทดสอบนี้ใช้ทดสอบว่าในสมการ

$\sigma_w^2 = \sigma_{\varepsilon}^2 + \sigma_u^2$ นั้น $\sigma_u^2 = 0$ หรือไม่คือทดสอบสมมติฐาน

H_0 : $\sigma_u^2 \neq 0$ vs H_1 : $\sigma_u^2 = 0$ หรือ H_0 : ไม่มี Random Effect ในสมการ vs H_1 : มี Random

Effect ในสมการ



4. สรุปข้อแนะนำการเลือกใช้สมการ

จาก $\beta_{it} = \beta_i + \varepsilon_{it}$ โดยที่ ε_{it} คือ อิทธิพลของปัจจัยเฉพาะหน่วยสำรวจ ถ้า ε_{it} เกี่ยวข้องกับตัวแปรอิสระให้เลือกใช้ FEM แต่ถ้า ε_{it} ไม่เกี่ยวข้องกับตัวแปรอิสระให้ใช้ ECM (REM) นอกจากนี้ยังต้องพิจารณาถึงจำนวนหน่วยสำรวจ (N) และจำนวนวาระ (T) ด้วย กล่าวคือ

1. ถ้า N เล็กคือมีเพียงไม่กี่หน่วยสำรวจและ T มีค่าสูงคือมีจำนวนวาระมาก กรณีนี้จะใช้ FEM หรือ ECM ก็ได้เพราะค่าประมาณไม่ต่างกันมาก แต่ FEM เหมาะสมกว่า
2. ถ้า N ใหญ่คือมีจำนวนหน่วยสำรวจมาก (เช่นจำนวนองค์กร คน วัตถุ พื้นที่ จังหวัด ประเทศ) แต่ T มีค่าน้อยคือมีหน่วยสำรวจละไม่กี่วาระ FEM และ ECM จะให้ค่าประมาณต่างกันมาก กรณีนี้ให้พิจารณาว่าหน่วยสำรวจที่มีอยู่นั้นเป็นสิ่งที่มีอยู่เพียงเท่านั้นหรือเป็นสิ่งที่ได้สุมมาจากกลุ่มประชากร ถ้ามีเพียงเท่านั้นคือทั้งหมดมีเพียงเท่านั้น เช่น สายการบินในประเทศ จังหวัดที่มีพื้นที่ปลูกปาล์มน้ำมัน ให้ใช้ FEM มิเช่นนั้นให้ใช้ ECM

อย่างไรก็ตามถ้า ε_{it} สัมพันธ์กับตัวแปรอิสระการใช้ FEM จะเหมาะสมมาก แต่ถ้าไม่สัมพันธ์กับตัวแปรอิสระการใช้ ECM ก็จะเหมาะสมกว่า หากไม่เป็นดังนี้คือถ้า ε_{it} สัมพันธ์กับตัวแปรอิสระแต่ใช้ ECM หรือถ้า ε_{it} ไม่สัมพันธ์กับตัวแปรอิสระ แต่ใช้ FEM ผลการประมาณค่าจะเอนเอียง (bias)

3. ECM สามารถประมาณค่าอิทธิพลของ observed time-invariant variable ได้ขณะที่ FEM จะใช้วิธีควบคุมตามวิธี LSDV



เอกสารอ้างอิง

มนตรี พิริยะกุล (2544), *เทคนิคการวิเคราะห์สมการถดถอย*, มหาวิทยาลัยรามคำแหง

Blumenstock, J. , *Fixed Effect Model (very important stuff)*, ค้นเมื่อ 30 ธันวาคม 2556 จาก

[findpdf.net/reader/Fixed-Effect -Models-very-important-stuff.html](http://findpdf.net/reader/Fixed-Effect-Models-very-important-stuff.html) หรือจาก [http:// www.](http://www.jblumenstock.com/courses/econ174/FEModels.pdf)

[jblumenstock.com/courses/econ174/FEModels.pdf](http://www.jblumenstock.com/courses/econ174/FEModels.pdf)

Wikipedia, Fixed Effect Model, ค้นเมื่อ 30 พฤศจิกายน 2556 จาก [en.Wikipedia.org/ wiki/fixed effect model](http://en.Wikipedia.org/wiki/fixed_effect_model)

Wikipedia, Panel analysis, ค้นเมื่อ 30 พฤศจิกายน 2556 จาก [en.Wikipedia.org/wiki/ panel analysis](http://en.Wikipedia.org/wiki/panel_analysis).

Gujarati, D. N. and Porter, D. C. (2009). *Basic Econometrics*, McGraw-Hill, Singapore.