บทที่ 3

ระเบียบวิธีวิจัย

3.1 แบบจำลองที่ใช้ในการศึกษา

การทดสอบความสัมพันธ์ระหว่างดัชนีราคาผู้บริโภคจังหวัดเชียงใหม่และผลการจัดเก็บภาษีทางตรงและภาษีทางอ้อมของสำนักงานสรรพากรพื้นที่จังหวัดเชียงใหม่นี้ได้นำข้อมูลทุติยภูมิรายเดือนตั้งแต่ เดือนมกราคม 2547 จนถึงเดือนธันวาคม 2553 ประกอบด้วยสามตัวแปร ได้แก่ ดัชนีราคาผู้บริโภคจังหวัดเชียงใหม่ และรายได้จัดเก็บภาษีทางตรงและภาษีทางอ้อมของสำนักงานสรรพากรพื้นที่จังหวัดเชียงใหม่โดยนำมาหาอัตราการเปลี่ยนแปลงระหว่างเดือน แล้วนำมาทดสอบเพื่อหารูปแบบความสัมพันธ์โดยวิธีVector Autoregressive Model (VAR) ซึ่งใช้ข้อมูลจากสำนักดัชนีการเศรษฐกิจการค้า กระทรวงพาณิชย์ และสำนักงานคลังจังหวัดเชียงใหม่ โดยพิจารณาแบบจำลอง VAR ของระบบ Multivariate ที่มี n ตัวแปร

$$Ay_t = \Gamma_0 + \sum_{i=t}^{p} \Gamma_i y_{t-i} + u_t \qquad (45)$$

โดย

y_t	หมายถึง	Vector ขนาด n x 1 ของตัวแปร endogenous
A	หมายถึง	matrix ขนาด n x n ของค่าสัมประสิทธิ์ของตัวแปร endogenous โดยมี diagonal ประกอบด้วยค่าเท่ากับ 1
Γ_0	หมายถึง	Vector ขนาด n x 1 ของ intercept
Γ_i	หมายถึง	Matrix ขนาด n x n ของค่าสัมประสิทธิ์ของตัวแปร Lag endogenous
u_t	หมายถึง	Vector ขนาด n x 1 ของค่าความคลาดเคลื่อนหรือ shock ของแบบจำลอง

ในการศึกษาครั้งนี้ เลือกตัวแปรทั้งหมด 3 ตัวแปรในแบบจำลอง VAR ดังนี้

<u>ดัชนีราคาผู้บริโภคจังหวัดเชียงใหม่ (CMICPI)</u> ในการศึกษาครั้งนี้จะใช้ข้อมูลรายเดือนโดยเป็นอัตราการเปลี่ยนแปลงของดัชนีราคาผู้บริโภคจังหวัดเชียงใหม่ระหว่างเดือนซึ่งเป็นข้อมูลจากสำนักดัชนีเศรษฐกิจการค้า

<u>ภาษีทางตรง (CMIDRT)</u> ในการศึกษาครั้งนี้ จะใช้ผลรวมของข้อมูลรายเดือนของภาษีบุคคลธรรมดา และภาษีนิติบุคคล โดยเป็นข้อมูลจากตารางภาวการณ์คลังของสำนักงานคลังจังหวัด

เชียงใหม่ที่เป็น Nominal term เนื่องจากมีผลทางด้านการเพิ่มขึ้นของระดับราคาสินค้าอยู่ด้วย หลังจากนั้นนำมาหาอัตราการการเปลี่ยนแปลงของภาษีทางตรงระหว่างเดือน

ภาษีทางอ้อม (CMIIRT) ในการศึกษาครั้งนี้ จะใช้ผลรวมของข้อมูลรายเดือนของภาษีมูลค่าเพิ่ม ภาษีธุรกิจเฉพาะ และอากรแสตมป์ โดยเป็นข้อมูลจากตารางภาวการณ์คลังของสำนักงานคลังจังหวัดเชียงใหม่ที่เป็น Nominal term เนื่องจากมีผลทางด้านการเพิ่มขึ้นของระดับราคาสินค้าอยู่ด้วย หลังจากนั้นนำมาหาอัตราการการเปลี่ยนแปลงของภาษีทางตรงระหว่างเดือน

เมื่อกำหนดตัวแปรทั้ง 3 จะได้แบบจำลอง VAR ในกรณีของแบบจำลองอันแรก (first-order) ในรูปแบบสมการข้างล่างนี้

$$\begin{bmatrix} \mathrm{CMICPI_t} \\ \mathrm{CMIDRT_t} \\ \mathrm{CMIIRT_t} \end{bmatrix} = \begin{bmatrix} \mathrm{b_{10}} \\ \mathrm{b_{20}} \\ \mathrm{b_{30}} \end{bmatrix} + \begin{bmatrix} b_{11(L)} & b_{12(L)} & b_{13(L)} \\ b_{21(L)} & b_{22(L)} & b_{23(L)} \\ b_{31(L)} & b_{32(L)} & b_{33(L)} \end{bmatrix} \begin{bmatrix} \mathrm{CMICPI_{t-1}} \\ \mathrm{CMIDRT_{t-1}} \\ \mathrm{CMIIRT_{t-1}} \end{bmatrix} + \begin{bmatrix} \mathrm{e_{CMICPI,t}} \\ \mathrm{e_{CMIDRT,t}} \\ \mathrm{e_{CMIIRT,t}} \end{bmatrix} \qquad (46)$$

3.2 วิธีการศึกษา

การวิเคราะห์ข้อมูลจะใช้โปรแกรม E-view 7 ในการทดสอบความสัมพันธ์ระหว่างดัชนีราคาผู้บริโภคจังหวัดเชียงใหม่และผลการจัดเก็บภาษีทางตรงและทางอ้อมของสำนักงานสรรพากรพื้นที่เชียงใหม่ ซึ่งใช้เทคนิคทางเศรษฐมิติที่เรียกว่า Vector Autoregression Model (VAR) โดยข้อมูลที่ใช้ในแบบจำลองนั้นเป็นตัวแปรในลักษณะของอนุกรมเวลา (Time Series) ด้วยเหตุที่ว่าการสร้างแบบจำลองของ VAR นั้นไม่ได้ยึดตามทฤษฎีที่เป็นโครงสร้าง(Structure) เท่าใดนัก เช่นแบบจำลองระบบสมการที่เกี่ยวพันกัน(Simultaneous Equation Model) อีกทั้งตามทฤษฎีของแบบจำลองแล้ว VAR ยังให้ผลการประมาณการหรือทำนาย (Forecast) ที่ดีกว่าวิธีของแบบจำลองที่เป็นโครงสร้าง เช่น แบบจำลองระบบสมการที่เกี่ยวพันกันที่ยุ่งยาก และเป็นโมเดลที่สามารถจัดการกับปัญหา Simultaneity Bias ได้ดี (Gujarati, 2003) ซึ่งการใช้แบบจำลอง VAR นั้น มีความได้เปรียบในแง่ของในกรณีที่เราอาจจะไม่ทราบความสัมพันธ์ที่แท้จริงในระหว่างตัวแปรทั้งหมดที่เกี่ยวข้องกัน หรืออาจจะไม่ทราบว่าตัวแปรใดเป็น Endogenous Variable หรือ Exogenous Variable กันแน่ แต่ทราบว่าโดยรวมแล้วตัวแปรทุกตัวในแบบจำลอง VAR มีผลต่อกัน ดังนั้นจึงสามารถใช้แบบจำลอง VAR ในการศึกษาถึงผลกระทบหรือความสัมพันธ์ของตัวแปรตัวใดตัวหนึ่งในแบบจำลองต่อตัวแปรอื่นในแบบจำลองได้โดยวิธีการวิเคราะห์ปฏิกิริยาตอบสนองต่อความแปรปรวน (Impulse Response Function) โดยมิต้องกังวลกับการตัดสินใจในการสร้างสมการในแบบ Structural Model เพราะใน VAR จะให้ตัวแปรทุกตัวเป็น Endogenous Variable อีกทั้งในการใช้ Simultaneous Equation Model ในกรณีที่ไม่ทราบความสัมพันธ์ที่แท้จริงของตัวแปรต่างๆในแบบจำลองอาจเกิดปัญหาเมื่อทำการ

ตัดทิ้งหรือเพิ่มตัวแปรบางตัวในระบบสมการ ซึ่งอาจเกิดปัญหาเช่น Identification Error ได้โดยมีขั้นตอนดังนี้

3.2.1 การทดสอบความนิ่งของข้อมูลหรือยูนิทรูท (Unit Root Test)

ทำการทดสอบว่าตัวแปรที่จะนำมาทำการศึกษามีลักษณะนิ่งหรือไม่ Augmented Dickey-fuller (ADF) Test โดยมีสมการในการทดสอบดังนี้

$$\Delta X_t = \theta X_{t-1} + \sum_{i=1}^{p} \emptyset_i \Delta X_{t-i} + e_t \quad (47)$$

$$\Delta X_t = \alpha + \theta X_{t-1} + \sum_{i=1}^{p} \emptyset_i \Delta X_{t-i} + e_t \quad (48)$$

$$\Delta X_t = \alpha + \beta_t + \theta X_{t-1} + \sum_{i=1}^{p} \emptyset_i \Delta X_{t-i} + e_t \quad (49)$$

สมมติฐานที่ใช้ในการทดสอบ คือ

$H_0 : \theta = 0$ (X_t เป็นข้อมูลที่มีลักษณะไม่นิ่ง Non-stationary)

$H_1 : \theta < 0$ (X_t เป็นข้อมูลที่มีลักษณะนิ่ง Stationary)

จากนั้นทำการเปรียบเทียบค่าสถิติที่ได้จาก ADF test ถ้าปฏิเสธสมมติฐานหลักแสดงว่าข้อมูลที่นำมาทดสอบมีลักษณะนิ่งที่ order of integration Zero [I(0)] แต่ถ้ายอมสมมติฐานหลักแสดงว่าข้อมูลที่ทดสอบมีลักษณะไม่นิ่งที่ order of integration Zero [I(d);d > 0]

3.2.2 การเลือกความล่าช้า (Lag) ที่เหมาะสม

ในการศึกษานี้ใช้เกณฑ์ Akaike Information Criteria (AIC) และ Schwarz's Bayesian Information Criterion (SC, BIC หรือ SBC) เป็นเกณฑ์ในการพิจารณาความเหมาะสมของจำนวนความล่าช้าหรือ Lag ของแบบจำลองมีสูตรดังนี้

$$AIC = log\widehat{\sigma^2} + 2\frac{p+q}{T} \quad (50)$$

โดยที่ $\widehat{\sigma^2}$ คือ ค่าประมาณของความแปรปรวนของ e_t

$$SC = log\widehat{\sigma^2} + 2\frac{p+q}{T}logT \quad (51)$$

เกณฑ์ทั้งสองเป็นเกณฑ์ที่อาศัยความควรจะเป็น (likelihood-based) และแสดงให้เห็นถึงความสมดุล (ที่มีผลในทางตรงกันข้าม) (trade off) ระหว่าง "fit" ซึ่งวัดโดยค่าของความควรจะเป็นและ "ตระหนี่ (parsimony)" ซึ่งวัดโดยจำนวนของพารามิเตอร์อิสระp+q ถ้าค่าคงที่ถูกนำไปรวมอยู่ในแบบจำลองด้วยจำนวนของพารามิเตอร์ดังกล่าวก็จะเพิ่มขึ้นเป็น p+q+1 สำหรับหลักเกณฑ์ในการตัดสินใจเลือกแบบจำลองก็คือเราจะเลือกแบบจำลองที่มีค่า AIC หรือ SC ที่มีค่าน้อยที่สุด ค่า AIC และ SC จะน้อยจากสาเหตุดังต่อไปนี้คือ มีความแปรปรวน และความแปรปรวนร่วมน้อย มีจำนวนของตัวแปรและจำนวน Lag น้อย และสุดท้ายมีจำนวนข้อมูลในการประมาณค่ามาก

ในขณะที่เกณฑ์ทั้งสองดังกล่าวมีความแตกต่างกันให้เลือกใช้ SC ไว้ก่อนเพราะว่า SC มีคุณสมบัติว่า SC จะเลือกแบบจำลองที่ถูกต้องเกือบแน่นอน สำหรับ AIC นั้นมีแนวโน้มที่จะเป็นลักษณะเชิงเส้นกำกับในแบบจำลองที่มีพารามิเตอร์มากเกินไป นอกจากนั้นในการศึกษานี้ จะทำการเปรียบเทียบผลการเลือก Lag กับเกณฑ์อื่นด้วยคือ Final Prediction Error (FPE) และHannan-Quinn Information Criterion (HQIC) ซึ่งให้ความหมายในลักษณะใกล้เคียงกัน

3.2.3 การทดสอบหา Cointegration วิธีการของ Johansen

Johansen (1988) และ Stock and Watson (1988) ได้เสนอตัวประมาณค่าแบบ maximumlikelihood (maximum likelihood estimator) ซึ่งทำให้สามารถหลีกเลี่ยงการใช้ตัวประมาณค่า 2ขั้นตอนได้ (two-step estimators) และสามารถที่จะประมาณค่าและทดสอบการมีอยู่จริงของcointegrating vectors หลาย vectors ได้ นอกจากนี้แล้วการทดสอบดังกล่าวยังทำให้เราสามารถทดสอบการใส่ข้อจำกัดของพารามิเตอร์ของ cointegrating vectors และความเร็วของการปรับตัว (speed of adjustment) ได้อีกด้วยอย่างไรก็ตามทั้งวิธีการของ Johansen (1988) และ Stock and Watson (1988) ต่างก็อาศัยความสัมพันธ์ระหว่าง rank ของเมทริกซ์และ characteristic roots ของเมทริกซ์ดังกล่าวอย่างมากและเพื่อที่จะเข้าใจขั้นตอนของวิธีการของ Johansen (1988) จึงเป็นการสรุปวิธีการและขั้นตอนของ Johansen (1988) ดังนี้

พิจารณา autoregressive process

$$y_t = A_1y_{t-1} + A_2y_{1-2} + \ldots + A_py_{t-p} + \varepsilon_t \tag{52}$$

จากสมการ (52) เอา y_{t-1} ไปลบออกทั้งสองข้างจะได้

$$\Delta_{y1} = (A_1 - I)\, y_{t-1} + A_2y_{t-3} + \ldots + A_py_{t-p} + \varepsilon_1 \tag{53}$$

จากสมการ (53) บวกเข้าและลบออกทางขวามือด้วย $(A - I)y_{t-2}$ จะได้

$$\Delta y_1 = (A_1 - I)\Delta y_{t-1} + (A_2 + A_1 - I)y_{t-2} + A_3y_{t-3} + \ldots + A_p + y_{t-p} + \varepsilon_t \tag{54}$$

ทำเช่นนี้ไปเรื่อย ๆ จะได้

$$\Delta y_t = \sum_{i=1}^{p-1} \pi_1 \Delta y_{t-1} + \pi y_{t-p} + \varepsilon_t \tag{55}$$

โดยที่ $\pi = -\left[I - \sum_{i=1}^{p} A_i\right]$

สิ่งสำคัญในสมการ (55) ก็คือ ค่าลำดับชั้น (rank) ของเมทริกซ์ π นั่นคือ ค่าลำดับชั้น (rank) ของ π จะเท่ากับจำนวนของ cointegrating vector ซึ่งสามารถแสดงได้ในรายละเอียดดังนี้

1. ถ้าต่างลำดับชั้น (rank) เท่ากับศูนย์ เมทริกซ์ π จะเป็นเมทริกซ์ศูนย์ และสมการ (55) ก็คือแบบจำลอง VAR ในรูปของผลต่างลำดับที่หนึ่ง (first difference)

2. ถ้าค่าลำดับชั้น (rank) ของ π เท่ากับ n (ซึ่งก็คือ มีค่าลำดับชั้น (rank)) เต็มที่หรือที่เรียกว่า full rank ซึ่ง vector process จะมีลักษณะนิ่ง (stationary) และเป็น VAR ใน level

3. ถ้าค่าลำดับชั้น (rank) ของ π เท่ากับ 1 เราก็จะมี cointegrating vector เพียง vector เดียว และ πy_{t-p} ก็คือ ปัจจัยการปรับตัวของความคลาดเคลื่อน (error-correction factor)

4. ในกรณีซึ่ง $1 < \text{rank } (\pi) < n$ เราก็จะมี cointegrating vector หลาย cointegrating vector

สำหรับการทดสอบ cointegration หรือการทดสอบความสัมพันธ์ระยะยาวระหว่างตัวแปรเพื่อใช้ในการเลือกแบบจำลองที่ใช้ในการประมาณค่าระหว่าง VAR และ VEC ในการศึกษานี้ได้ใช้การทดสอบ Johansen Trace ของ Johansen and Juselius (1990) เพื่อหาจำนวนของความสัมพันธ์ cointegration ได้ ด้วยการใช้การทดสอบ Likelihood Ratio test statistic ภายใต้ข้อสมมติฐานหลัก คือ

$$H_0 : \text{rank}(\Pi) = r$$

และ $$H_1 : \text{rank}(\Pi) \geq r + 1$$

โดยที่ Π คือ เมตริกซ์สัมประสิทธิ์ของความสัมพันธ์ระหว่าง ΔY_t และ ใน ΔY_{t-1} ในแบบจำลอง VEC

r คือ จำนวน rank ของเมตริกซ์ Π

โดยเมื่อค่าทดสอบ Trace มากกว่าค่าวิกฤต ทำให้สามารถปฏิเสธสมมติฐานหลัก (null hypothesis) หมายความว่า ตัวแปรใน Y_t ไม่มีความสัมพันธ์กัน หากค่าทดสอบ Trace มีค่าน้อยกว่าค่าวิกฤต จะยอมรับสมมติฐานหลัก หมายความว่า ตัวแปรใน Y_t มีความสัมพันธ์กันอย่างน้อยหนึ่งความสัมพันธ์ ลำดับต่อไปก็จะเป็นการทดสอบซ้ำ โดยใช้สมมติฐาน คือ

$$H_0 : \text{rank}(\Pi) = r$$

และ $$H_1 : \text{rank}(\Pi) \geq r + 1$$

โดยที่ Π คือ เมตริกซ์สัมประสิทธิ์ของความสัมพันธ์ระหว่าง ΔY_t และ ใน ΔY_{t-1} ในแบบจำลอง VEC

r คือ จำนวน rank ของเมตริกซ์ Π

ซึ่งในกรณีที่สามารถปฏิเสธสมมติฐานครบ จนกระทั่ง Full Rank เราสามารถใช้แบบจำลอง VAR ในการประมาณค่าได้ หากไม่ใช่ Full Rank มีความสัมพันธ์ระหว่างตัวแปรทั้งสอง ซึ่งทำให้สามารถหาความสัมพันธ์ในระยะสั้นและระยะยาวได้ เราจะใช้แบบจำลอง VEC แทน

3.2.4 แบบจำลอง VAR

การศึกษานี้ได้กำหนดแบบจำลอง VAR เป็นแบบจำลองที่เหมาะสมเพื่อใช้ในการศึกษา เนื่องจากลักษณะและความสัมพันธ์ของตัวแปรอาจไม่ชัดเจนและเป็นความสัมพันธ์ในเชิงพลวัตร ประกอบกับข้อสมมติให้ตัวแปรแต่ละตัวไม่ส่งผลต่อตัวแปรอื่นๆในช่วงเวลาเดียวกัน อีกทั้งการศึกษาครั้งนี้ส่วนหนึ่งเพื่อตอบคำถามถึงผลกระทบของเงินทุนเคลื่อนย้ายที่มีต่อดัชนีค่าเงินที่แท้จริง ถึงขนาด ทิศทาง ระยะเวลา ความคงอยู่ (Persistence) และสัดส่วนของผลกระทบที่มีต่อดัชนีค่าเงินที่แท้จริง

เนื่องจากความสัมพันธ์ของตัวแปรแต่ละตัวมีความสัมพันธ์ที่ไม่แน่นอน และส่งผลกระทบระหว่างกันทั้งทางตรงและทางอ้อม ข้อสมมติประการหนึ่งที่จำเป็นและเหมาะสมต่อการศึกษาในครั้งนี้ คือ ตัวแปรแต่ละตัวจะไม่ส่งผลกระทบต่อตัวแปรตัวอื่นในช่วงเวลาเดียวกันหรือไม่ส่งผลกระทบอย่างทันทีเมื่อตัวแปรหนึ่งเปลี่ยนแปลง เพราะการตอบสนองต่อ Shock ที่เกิดขึ้นและที่มีผลต่อตัวแปรต่างๆในระบบเศรษฐกิจนั้นยังมีความล่าช้า (Non-Contemporaneous Effect)

เราสร้างแบบจำลองของเวกเตอร์นี้ในรูปของค่าที่ผ่านมาในอดีตของเวกเตอร์ดังกล่าวนี้ ผลที่ได้ก็คือ Vector autoregression (VAR) สามารถเขียนได้ดังนี้

$$Ay_t=\Gamma_0+\sum_{i=t}^{p}\Gamma_i y_{t-i}+u_t \qquad (56)$$

โดย

y_t	หมายถึง	Vector ขนาด n x 1 ของตัวแปร endogeneous
A	หมายถึง	matrix ขนาด n x n ของค่าสัมประสิทธิ์ของตัวแปร endogeneous โดยมี diagonal ประกอบด้วยค่าเท่ากับ 1
Γ_0	หมายถึง	Vector ขนาด n x 1 ของ intercept
Γ_i	หมายถึง	Matrix ขนาด n x n ของค่าสัมประสิทธิ์ของตัวแปร Lag endogenous
u_t	หมายถึง	Vector ขนาด n x 1 ของค่าความคลาดเคลื่อนหรือ shock ของแบบจำลอง

3.2.5 การวิเคราะห์ปฏิกิริยาตอบสนองต่อความแปรปรวน (Impulse Response Function : IRF)

เนื่องจากการวิเคราะห์แบบจำลอง VAR ไม่สามารถวิเคราะห์จากค่าสัมประสิทธิ์ที่ได้จากการประมาณค่า จึงต้องอาศัยวิธีการอื่นในการช่วยวิเคราะห์ Impulse Response Function (IRF) เป็นอีกหนึ่งวิธีการ ที่อาศัยแนวคิด Moving Average เพื่อพิจารณาการเคลื่อนไหวของตัวแปรที่เป็นอนุกรมเวลา โดยแบบจำลอง VAR จะอาศัยคุณสมบัติ Stability ของแบบจำลอง ในการเขียนแบบจำลองให้อยู่ในรูปของ Vector Moving Average (VMA) ดังนี้

$$\begin{bmatrix} \mathrm{CMICPI_t} \\ \mathrm{CMIDRT_t} \\ \mathrm{CMIIRT_t} \end{bmatrix} = \begin{bmatrix} \overline{\mathrm{CMICPI_t}} \\ \overline{\mathrm{CMIDRT_t}} \\ \overline{\mathrm{CMIIRT_t}} \end{bmatrix} + \sum_{i=0}^{n} \begin{bmatrix} \phi_{11(i)} & \phi_{12(i)} & \phi_{13(i)} \\ \phi_{21(i)} & \phi_{22(i)} & \phi_{23(i)} \\ \phi_{31(i)} & \phi_{32(i)} & \phi_{33(i)} \end{bmatrix} \begin{bmatrix} \mathrm{u_{CMICPI,t}} \\ \mathrm{u_{CMIDRT,t}} \\ \mathrm{u_{CMIIRT,t}} \end{bmatrix} \quad (57)$$

จากนั้นทำการหาตัวคูณ Multiplier ($\phi_{ij}(i)$) ของค่าความผิดพลาด (ε_i) ในแบบจำลอง VMA ในแต่ละช่วงเวลา และนำตัวคูณนั้นมา Plot กราฟเทียบกับเวลาจะได้ IRF หลังจากที่ได้ IRF จะสามารถวิเคราะห์ความสัมพันธ์ของตัวแปรหนึ่งต่ออีกตัวแปรหนึ่งในแต่ละช่วงเวลา ซึ่งในการศึกษานี้ IRF สามารถบอกทิศทางแนวโน้มการเปลี่ยนแปลงและขนาดของผลกระทบในแต่ละช่วงเวลาได้ โดยตัวแปรที่มีผลต่อดัชนีค่าเงินที่แท้จริงที่สำคัญ คือ ความหนืดของดัชนีค่าเงินที่แท้จริง (Persistence) และตัวแปรอื่น