บทที่ 4

วิธีการศึกษา

บทที่ 4 ประกอบด้วย 2 ส่วน คือ ส่วนที่ 1 ตัวแบบจำลองที่ใช้ในการศึกษา ส่วนที่ 2 เป็นวิธีการศึกษา

4.1 ตัวแบบจำลองที่ใช้ในการศึกษา

ตัวแบบจำลองที่ใช้ในการศึกษาเป็นระบบสมการ (System of Equations) เหตุผลที่พิจารณาระบบสมการเนื่องจาก การศึกษาทั้งระบบสมการทำให้สามารถพิจารณาตัวแปรได้หลายตัวไปพร้อมกัน ซึ่งสอดคล้องกับข้อความจริงที่ว่า การเคลื่อนไหวของตัวแปรทางการเงินในระบบเศรษฐกิจมีความสัมพันธ์ซึ่งกันและกัน มากกว่าความสัมพันธ์ในทิศทางเดียว ข้อความข้างต้นแสดงให้เห็นจริงจากการศึกษาของ Fama (1990) ที่พบความสัมพันธ์แบบ Contemporaneous Correlation ระหว่างตัวแปรทางการเงิน และการศึกษาที่ให้ผลสอดคล้องกับการศึกษาของ Fama ตัวอย่างเช่น การศึกษาของ Campbell and Shiller (1988) และ Campbell (1991) และอื่นๆ ดังนั้นการใช้ระบบสมการ จึงเป็นการเพิ่มข่าวสารข้อมูลลงไปในตัวแบบจำลอง ข้อมูลข่าวสารที่เพิ่มเข้าไปจะช่วยในการพรรณนาพฤติกรรมของตัวแปรที่กำลังศึกษา สมการที่ (4.1) แสดงถึงระบบสมการ Vector STAR เพื่อพรรณนาพฤติกรรมการเคลื่อนไหวเชิงสุ่มควบคู่กันระหว่างตัวแปรที่ใช้และอัตราผลตอบแทนส่วนเกินในพันธบัตรรัฐบาลในแต่ละอายุคงเหลือ

$$\mathbf{Y}_t = \left(\mu^1 + \sum_{j=1}^{P} \beta_j^1 \mathbf{Y}_{t-j} \right)\left(1 - G\left(s_t;\gamma,c\right)\right) + \left(\mu^2 + \sum_{j=1}^{P} \beta_j^2 \mathbf{Y}_{t-j} \right)\left(G\left(s_t;\gamma,c\right)\right) + \varepsilon_t \quad (4.1)$$

โดยที่

$\mathbf{Y}_t$ = $\left(exre\_iy_t,\ slope_t,\ exrre\_set_t\right)'_{3\times1}$

เวคเตอร์ของตัวแปรภายในระบบสมการ

$exre\_iy_t$ = อัตราผลตอบแทนส่วนเกินจากการลงทุนในพันธบัตรรัฐบาลที่มีอายุคงเหลือ i ปี : i = 1, 2, 5, 7 และ 10 ปี

$slope_t$ = ความชันของเส้นโครงสร้างอัตราดอกเบี้ย

$exre\_set_t$ = อัตราลตอบแทนส่วนเกินของการลงทุนตลาดหลักทรัพย์ของไทย

สมการที่ (4.1) แสดงให้เห็นว่าตัวแปรภายในระบบสมการมี 3 ตัวแปรคือ (1) อัตราผลตอบแทนส่วนเกินในพันธบัตรรัฐบาลในแต่ละอายุคงเหลือ (Excess Bond Return) (2) ความชันของเส้นโครงสร้างอัตราดอกเบี้ย (Slope of Term Structure Interest Rate) (3) อัตราผลตอบแทนส่วนเกินของการลงทุนในดัชนีหลักทรัพย์ (Excess Stock Return) ตัวแปรทั้ง 3 อยู่ภายใต้ทฤษฎีความต้องการเงินกู้และกรอบความคิดความต้องการเงินกู้และการศึกษาเชิงประจักษ์ การคำนวณตัวแปรในระบบสมการที่ (4.1) มีรายละเอียดดังนี้

(1) อัตราผลตอบแทนส่วนเกินจากการลงทุนในพันธบัตรรัฐบาลที่มีอายุคงเหลือ i ปี (Excess Bond Return) ใช้สัญลักษณ์ $exre\_iy_t$ สามารถคำนวณตามสมการที่ (4.2)

$$
\begin{aligned}
exre\_iy_T &= r_{iy,t} - r_{f,t} \\
r_{iy,t} &= \ln\left(\frac{B_{i,t}}{B_{i,t-1}}\right) \\
B_{iy,t} &= \exp\left\{-\text{spot rate}_{iy,t} * y\right\}
\end{aligned}
\tag{4.2}
$$

โดยที่

$exre\_iy_t$ = อัตราผลตอบแทนส่วนเกินจากการลงทุนในพันธบัตรรัฐบาลที่มีอายุคงเหลือ i ปี ณ วันที่ t โดยที่ i = 1 2 5 7 และ 10 ปี

$r_{iy,t}$ = อัตราผลตอบแทนจากการลงทุนในพันธบัตรรัฐบาลที่มีอายุคงเหลือ i ปี ณ วันที่ t โดยที่ i = 1 2 5 7 และ 10 ปี

$r_{f,t}$ = อัตราผลตอบแทนที่ปราศจากความเสี่ยง (Risk free rate) ในที่นี้จะใช้อัตราผลตอบแทนจากการลงทุนในตลาดซื้อคืนพันธบัตร 1 วัน (repurchase rate)[1]

$B_{iy,t}$ = ราคาพันธบัตรรัฐบาลที่มีอายุคงเหลือ i ปี ณ วันที่ t โดยที่ i = 1 2 5 7 และ 10 ปี

y = อายุคงเหลือของพันธบัตรรัฐบาล

$\text{spot rate}_{y,t}$ = อัตราคิดลดแบบสปอต (Spot rate) สำหรับตราสารหนี้ที่มีอายุคงเหลือ i ปี ณ วันที่ t โดยที่ i = 1 2 5 7 และ 10 ปี[2]

---

[1] ข้อมูลอนุกรมเวลาของ Repurchase rate เก็บข้อมูลจาก www. thaibdc.or.th

[2] ข้อมูลอนุกรมเวลาของ อัตราคิดลดแบบสปอต เก็บข้อมูลจาก www. thaibdc.or.th

(2) ความชันของเส้นโครงสร้างอัตราดอกเบี้ย (Slope of Term Structure Interest Rate) ใช้สัญลักษณ์ $slope_t$ สามารถคำนวณได้ตามสมการที่ (4.3)

$$slope_t = Interpolation\ GB\ yield_{10y,t} - Interpolation\ GB\ yield_{1y,t} \quad (4.3)$$

โดยที่

$slope_t$ = ความชันของเส้นโครงสร้างอัตราดอกเบี้ย (Slope of Term Structure Interest Rate)

$Interpolation\ GB\ yield_{10y,t}$, $Interpolation\ GB\ yield_{1y,t}$ = อัตรา Interpolation of Government Bond Yield[3] ที่มีอายุคงเหลือ 10 ปี และ 1 ปี ณ วันที่ t

(3) อัตราผลตอบแทนส่วนเกินของการลงทุนตลาดหลักทรัพย์ของไทย (Excess Stock Return) ใช้สัญลักษณ์ $exre\_set_t$ สามารถคำนวณได้ตามสมการที่ (4.4)

$$exre\_set_t = r_{set,t} - r_{f,t}$$

$$r_{set,t} = \ln\left(\frac{set\ index_t}{set\ index_{t-1}}\right) \quad (4.4)$$

โดยที่

$exre\_set_t$ = อัตราผลตอบแทนส่วนเกินจากการลงทุนในตลาดหลักทรัพย์ของไทย ณ วันที่ t

$r_{set,t}$ = อัตราผลตอบแทนจากการลงทุนในตลาดหลักทรัพย์ของไทย ณ วันที่ t

$r_{f,t}$ = อัตราผลตอบแทนที่ปราศจากความเสี่ยง (risk free rate)

$set\ index_t$ = ดัชนีตลาดหลักทรัพย์ของไทย[4] ณ วันที่ t ทั้งนี้

เหตุผลที่พิจารณาเพียง 3 ตัวแปรข้างต้น ทั้งที่การศึกษาต่างประเทศใช้ตัวแปรตัวอื่น เช่น สภาพคล่องของตราสารหนี้ เนื่องจากเหตุผล 2 ประการ ประการแรก คือ มูลค่าการซื้อขายตราสารหนี้ในประเทศไทยแต่ละครั้งต้องใช้เงินเป็นจำนวนมาก ทำให้นักลงทุนที่เข้ามาลงทุนในตลาดตราสารหนี้ไทย จึงเป็นนักลงทุนสถาบันการเงินมากกว่านักลงทุนรายย่อยเหมือนกับตลาดหลักทรัพย์ เหตุผลประการที่ 2 คือ การซื้อขายพันธบัตรรัฐบาลในประเทศไทยเป็นการซื้อขายโดยตรงระหว่างผู้ลงทุนและผู้ระดมทุนแล้วจึงรายงานอัตราผลตอบแทนไปยังศูนย์ซื้อขายตราสาร

---

[3] ข้อมูลอนุกรมเวลาของ Interpolation of Government Bond Yield เก็บจาก www. bot.or.th

[4] ข้อมูลอนุกรมเวลาของ SET Index สามารถเก็บข้อมูลจาก www. set.or.th

หนี้แห่งประเทศไทย ดังนั้นข้อมูล bid-ask ซึ่งเป็นตัวแปรที่การศึกษาเชิงประจักษ์ต่างประเทศใช้เป็นตัวแทนสภาพคล่องการซื้อขาย จึงไม่เหมาะสมในกรณีประเทศไทย

4.2 วิธีการศึกษา

4.2.1 การทดสอบเสถียรภาพและรายงานค่าสถิติพรรณนาของตัวแปรที่ใช้ในการศึกษา

ก่อนการกำหนดค่าพารามิเตอร์ในตัวแบบจำลองควรจะทำการทดสอบลักษณะคุณสมบัติด้านความมีเสถียรภาพของข้อมูล (Stationary) เสียก่อน เนื่องจากข้อมูลอนุกรมเวลาทางการเงินส่วนใหญ่จะเปลี่ยนแปลงไปตามเวลาในลักษณะที่เพิ่มขึ้น จึงทำให้การกำหนดตัวแบบจำลองที่เหมาะสมนั้นเป็นไปได้ยากเพราะมีอิทธิพลของเวลามาเกี่ยวข้อง และข้อมูลที่ขาดคุณสมบัติ Stationary จะทำให้การวิเคราะห์โดยใช้เครื่องมือปกติทางสถิติได้ผลลัพธ์ที่ไม่ถูกต้อง

กระบวนการตัวแปรที่มีเสถียรภาพหรือ Stationary Process จะมีค่าเฉลี่ย (Mean) และค่าความแปรปรวน (Variance) มีขนาดเปลี่ยนแปลงไม่ขึ้นกับเวลาที่ใช้สุ่มตัวอย่าง ถ้าตัวแปรที่พิจารณนาขาดคุณสมบัติ Stationary หรือเป็นตัวแปรสุ่มแบบ $I(n)$ (Integrated Variable of Degree n) โดยที่ $n \geq 1$ ต้องพยายามปรับเปลี่ยนตัวแปรเชิงสุ่ม $I(n \geq 1)$ โดยการหาค่าความแตกต่าง n ครั้ง(n differencing) ให้เป็นตัวแปรเชิงสุ่มแบบ $I(0)$ ซึ่งมีคุณสมบัติ Stationary จึงสามารถนำไปวิเคราะห์โดยใช้เครื่องมือปกติทางสถิติได้

การทดสอบความมีเสถียรภาพของข้อมูลจะทำการทดสอบ Unit Root Test หรืออันดับความสัมพันธ์ของข้อมูล (Orders of Integration) ด้วยวิธีการศึกษาของ Dickey and Fuller (1979) โดยเริ่มต้นด้วยการประมาณแบบจำลอง Autoregression ลำดับหนึ่ง ซึ่งสามารถเขียนได้ดังสมการที่ (4.5)

$$X_t = \rho X_{t-1} + \varepsilon_t \quad ; t = 1, 2, 3, \ldots\ldots, T \tag{4.5}$$

โดยที่

X_t = ตัวแปรที่กำลังศึกษา

ρ = สัมประสิทธิ์ของตัวแปรความล่าช้า (Lagged) ของอนุกรมเวลา X_{t-1}

ε_t = ค่าความคลาดเคลื่อน (Error Term) โดยที่ $\varepsilon_t \sim N(0, \sigma^2)$

การทดสอบ Stationary จะพิจารณาค่าสัมประสิทธิ์ตัวแปรความล่าช้า (ρ) โดยมีเงื่อนไขคือ

ถ้า $|\rho| < 1$ แสดงว่า X อนุกรมเวลา X_t มีลักษณะที่เป็น Stationary

ถ้า $|\rho| \geq 1$ แสดงว่า X อนุกรมเวลา X_t มีลักษณะที่เป็น Non - Stationary

อนุกรมเวลาของตัวแปรส่วนใหญ่จะมีค่าเป็นบวกมากกว่าลบ ดังนั้นสมมติฐานแรก (Null Hypothesis) ที่เหมาะสม คือ $\rho = 1$ และสมมติฐานรอง (Alternative Hypothesis) คือ $\rho < 1$

การตั้งสมมติฐานสำหรับการทดสอบดังนี้

H_0 : $\rho = 1$ (Non - Stationary)

H_1 : $\rho < 1$ (Stationary)

ถ้าผลการทดสอบ Unit Root Test ปรากฏว่าการประมาณค่า ρ ไม่แตกต่างจาก 1 อย่างมีนัยสำคัญ แสดงว่าไม่สามารถปฏิเสธ Null Hypothesis และสรุปได้ว่าอนุกรมเวลา X_t เป็น Non – Stationary แต่ในทางตรงกันข้ามการประมาณค่า ρ มีค่าน้อยกว่า 1 อย่างมีนัยสำคัญ สรุปได้ว่าอนุกรมเวลา X_t เป็น Stationary

หลังจากปรับข้อมูลให้มีคุณสมบัติ Stationary เรียบร้อยแล้ว จะทำการพรรณนาค่าสถิติเบื้องต้น (Descriptive Statistics) เพื่อเป็นข้อมูลขั้นต้นในการประเมินรูปร่างของการแจกแจงของตัวแปรด้วยค่าสถิติเชิงพรรณนาประกอบด้วย ค่าเฉลี่ยตัวอย่าง (Sample Mean = $\hat{\mu}$) ค่าส่วนเบี่ยงเบนมาตรฐานตัวอย่าง (Sample Standard Deviation = $\hat{\sigma}$) ค่าความเบ้ตัวอย่าง (Sample Skewness = $\hat{S}$) ค่าสัมประสิทธิ์เคอโตซิสตัวอย่าง (Sample Kurtosis Coefficient = $\hat{K}$) และค่าสัมประสิทธิ์สหสัมพันธ์ตัวอย่าง (Sample Correlation = $\hat{\rho}$) ค่าสถิติดังกล่าวคำนวณได้จากสมการที่ (4.6)

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} X_t \qquad \hat{S} = \frac{1}{\hat{\sigma}^3}\frac{\sum_{t=1}^{T}(X_t - \hat{\mu})^3}{T-1}$$

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{T}(X_t - \hat{\mu})^2}{T-1} \qquad \hat{K} = \frac{1}{\hat{\sigma}^4}\frac{\sum_{t=1}^{T}(X_t - \hat{\mu})^4}{T-1} \qquad (4.6)$$

$$\hat{\rho}(X,Y) = \frac{\hat{\sigma}_{xy}}{\sqrt{\hat{\sigma}_x^2 * \hat{\sigma}_y^2}} \qquad \hat{\sigma}_{xy} = \frac{\sum_{t=1}^{T}(X_t - \hat{\mu}_x)(Y_t - \hat{\mu}_y)}{T-1}$$

ค่าเฉลี่ยตัวอย่าง (Sample Mean $=\hat{\mu}$) จะชี้ถึงอัตราผลตอบแทนที่คาดจากการลงทุน ในทางสถิติค่าเฉลี่ยตัวอย่างจะใช้เป็นค่าซึ่งระบุตำแหน่งที่ตั้งของการแจกแจง และเนื่องจากอัตราผลตอบแทนเป็นตัวแปรเชิงสุ่ม อัตราผลตอบแทนที่เกิดขึ้นจริงในแต่ละวันจึงอาจมีระดับแตกต่างไปจากอัตราผลตอบแทนที่คาดในลักษณะที่เป็นการกระจายตัวอยู่โดยรอบค่าเฉลี่ย ขนาดของการกระจายตัวสามารถชี้โดยค่าส่วนเบี่ยงเบนมาตรฐานตัวอย่าง (Sample Standard Deviation = $\hat{\sigma}$) ดังนั้นหากค่าส่วนเบี่ยงเบนมาตรฐานตัวอย่างมีค่ามากย่อมแสดงถึง การลงทุนในหลักทรัพย์มีความผันผวนไปจากค่าเฉลี่ยมาก ซึ่งเป็นการลงทุนที่มีความเสี่ยงสูง สำหรับค่าความเบ้ตัวอย่าง (Sample Skewness = $\hat{S}$) แสดงถึง การกระจายตัวของอัตราผลตอบแทนซึ่งเป็นตัวแปรเชิงสุ่มที่เกิดขึ้นโดยรอบค่าเฉลี่ยอาจจะเป็นการกระจายแบบอสมมาตร เบ้ไปทางซ้ายหรือไปทางขวา ความเบ้ของการกระจายสามารถใช้ค่าสกิวเนสเป็นเครื่องบ่งชี้ โดยค่าความเบ้ตัวอย่างที่เป็นลบชี้ว่า การแจกแจงมีความเบ้ไปทางซ้ายหรือมีส่วนหางชี้ไปทางซ้าย ค่าความเบ้ที่มีค่าเป็นบวกจะชี้ว่าการแจกแจงมีความเบ้ไปทางขวา หรือมีส่วนหางชี้ไปทางขวา ในขณะที่ถ้าการแจกแจงมีลักษณะสมมาตรค่าความเบ้จะมีค่าเป็นศูนย์ ค่าสัมประสิทธิ์เคอโตซิสตัวอย่าง (Sample Kurtosis Coefficient = $\hat{K}$) เป็นค่าที่ชี้ถึงขนาดของมวลบริเวณหางของการแจกแจง โดยค่าเคอโคซิสของการแจกแจงแบบปกติมีค่าเท่ากับ 3.00 แต่ถ้าการแจกงแจงที่มีหางอ้วน ค่าเคอโคซิสจะมีค่าสูงกว่า 3.00 และถ้าการแจกแจงที่มีหางผอม ค่าเคอโคซิสจะมีค่าต่ำกว่า 3.00 และเพื่อให้เกิดความมั่นใจ การศึกษาจึงจะทำการทดสอบว่า การแจกแจงของตัวแปรที่นำมาศึกษามีรูปร่างที่ต่างจากการแจกแจงแบบปกติอย่างมีนัยสำคัญหรือไม่ การทดสอบจะพิจารณาค่าสถิติ Wald Test (W) ซึ่งคำนวณได้จากสมการที่ (4.7)

$$W = T\left\{\frac{\hat{S}^2}{6}+\frac{(\hat{K}-3)^2}{24}\right\} \qquad (4.7)$$

ถ้าการแจกแจงของตัวแปรที่นำมาศึกษาเป็นการแจกแจงแบบปกติ ค่าสถิติ W จะเป็นตัวแปรเชิงสุ่มที่มีการแจกแจงแบบไคสแควร์ ที่องศาความเป็นอิสระเท่ากับ 2 การศึกษาจะปฏิเสธข้อสมมติฐานว่า การแจกแจงของตัวแปรที่นำมาศึกษาเป็นการแจกแจงแบบปกติ ณ ระดับความเชื่อมั่นร้อยละ 99 เมื่อค่าสถิติ W มีค่าเกินกว่า 9.21

การรายงานค่าสถิติเบื้องต้นเป็นการตรวจสอบข้อมูลเบื้องต้นว่าข้อมูลที่นำมาศึกษาในครั้งนี้มีลักษณะคล้ายคลึงหรือแตกต่างกับการศึกษาเชิงประจักษ์อดีตมากน้อยเพียงใด เพื่อจะได้เปรียบเทียบผลการศึกษา สำหรับการตรวจสอบการกระจายตัวของตัวแปรที่นำมาศึกษา เพื่อตรวจสอบว่าควรจะใช้จำนวนข้อมูลมากน้อยเพียงใด กล่าวคือ ถ้าตัวแปรมีการกระจายตัวแบบไม่

ปกติ จำเป็นต้องใช้ข้อมูลที่มีขนาดใหญ่ เพื่อจะได้ค่าพารามิเตอร์ที่มีประสิทธิภาพ สำหรับค่าสัมประสิทธิ์สหสัมพันธ์ เป็นค่าสถิติเบื้องต้นที่ชี้ว่าชุดตัวแปรที่กำลังศึกษามีความสัมพันธ์เชิงเส้นกันหรือไม่ ถ้าการศึกษาไม่พบว่าชุดตัวแปรมีค่าสัมประสิทธิ์สหสัมพันธ์กัน ก็ไม่จำเป็นต้องทำการศึกษาโดยใช้ระบบสมการ

4.2.2 <u>การกำหนดพารามิเตอร์ในแบบจำลอง Vector STAR</u>

การศึกษาใช้ตัวแบบจำลอง Vector STAR ในการพรรณนาพฤติกรรมความเสี่ยงของอัตราผลตอบแทนส่วนเกินจากการลงทุนในพันธบัตรรัฐบาลในแต่ละจุดของเวลาแบบมีเงื่อนไขขึ้นกับ ค่าในอดีตของอัตราผลอตอบแทนส่วนเกินจากการลงทุนในพันธบัตรในแต่ละอายุคงเหลือ (exre_set) ความชันของเส้นโครงสร้างอัตราดอกเบี้ย (slope) และอัตราผลตอบแทนส่วนเกินของการลงทุนในดัชนีหลักทรัพย์ (exre_set) แสดงให้เห็นในสมการที่ (4.1)

ฟังก์ชันการเปลี่ยนแปลง (Transition Function ; $G(S_t;\gamma,C)$) ที่จะนำมาพิจารณาประกอบด้วย ฟังก์ชัน Logistic สมการที่ (4.8) และฟังก์ชัน Quadratic Logistic สมการที่ (4.9) ทั้งนี้จะต้องทำการทดสอบเพื่อเลือกฟังก์ชันที่เหมาะสมก่อนที่จะกำหนดค่าพารามิเตอร์ในตัวแบบจำลอง Vector STAR

$$G(S_t;\gamma,C) = \frac{1}{1+\exp\left\{\dfrac{-\gamma(S_t-C)}{\sigma_{S_t}}\right\}} \quad ; \gamma > 0 \tag{4.8}$$

หรือ

$$G(S_t;\gamma,C) = \frac{1}{1+\exp\left\{\dfrac{-\gamma(S_t-C_1)(S_t-C_2)}{\sigma_{S_t}}\right\}} \quad ; c_1 \leq c_2 \, , \gamma > 0 \tag{4.9}$$

ขั้นตอนการกำหนดค่าพารามิเตอร์ตัวแบบจำลอง Vector STAR จะเริ่มต้นจากการกำหนดค่าพารามิเตอร์ในตัวแบบจำลอง Vector Autoregressive Model (VAR Model) เสียก่อน สมการที่ (4.10) แสดงตัวแบบจำลอง VAR

$$\begin{bmatrix} exre\_iy_t \\ slope_t \\ exre\_set_t \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} \sum_{j=1}^{p} \begin{bmatrix} \beta_{11}^j & \beta_{12}^j & \beta_{13}^j \\ \beta_{21}^j & \beta_{22}^j & \beta_{23}^j \\ \beta_{31}^j & \beta_{32}^j & \beta_{33}^j \end{bmatrix} \begin{bmatrix} exre\_iy_{t-j} \\ slope_{t-j} \\ exre\_set_{t-j} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{bmatrix} \quad (4.10)$$

ประเด็นปัญหาสำหรับการกำหนดค่าพารามิเตอร์ในตัวแบบจำลอง VAR คือ จำนวนความล่าช้า (p) ที่เหมาะสมก่อน (Optimum Lag) ควรจะเป็นเท่าไร LÜtkepohl (1991) เสนอให้ใช้การ Likelihood Ratio Test ค่าสถิติที่ใช้ทดสอบ LR แสดงให้เห็นดังที่สมการที่ (4.11)

$$LR = (T - C)\left(\log\left|\sum_r\right| - \log\left|\sum_u\right|\right) \quad \sim H_0 \ \chi^2_{\alpha,df} \quad (4.11)$$

โดยที่

T = จำนวนข้อมูล

C = จำนวนพารามิเตอร์ในแต่ละสมการของ Unrestricted Model

$\log\left|\sum_r\right| - \log\left|\sum_u\right|$ = Log ของ determinant ของ Variance - Covariance matrix ใน restricted และ unrestricted Model ตามลำดับ

การทดสอบมีสมมติฐาน ดังนี้

H_0 : จำนวนความล่าช้าเท่ากับจำนวนความล่าช้าในสมการ Restricted Model

H_1 : จำนวนความล่าช้าเท่ากับจำนวนความล่าช้าในสมการ Unrestricted Model

ภายใต้ H_0 ค่า LR Statistic มีการกระจายแบบ χ^2 และมีองศาความเป็นอิสระเท่ากับจำนวนสัมประสิทธิ์ใน Restricted Model ถ้าค่า LR Statistic มากกว่า $\chi^2_{\alpha,df}$ จะปฏิเสธสมมติฐานหลัก (Null hypothesis: H_0) การทดสอบทำโดยเปรียบเทียบ VAR Model ที่มีจำนวนความล่าช้ามากที่สุด (Unrestricted Model) แต่ไม่ทำให้เสียองศาความเป็นอิสระ Degree of Freedom มากเกินไป กับ VAR Model ที่มีจำนวนความล่าช้าน้อยกว่า 1 ช่วงเวลา (Restricted Model) ถ้าผลการทดสอบยอมรับสมมติฐานหลัก แบบจำลองที่มีจำนวนความล่าช้าน้อยกว่าเป็นแบบจำลองที่เหมาะสม

เมื่อทราบค่าความล่าช้าที่เหมาะสมของตัวแบบจำลอง VAR แล้ว จึงคำนวณค่าพารามิเตอร์ในตัวแบบจำลอง VAR ในแต่ละอายุคงเหลือได้ หลังจากนั้นต้องทำการทดสอบความสัมพันธ์เชิงเส้นตรงของระบบสมการเพื่อแสดงให้เห็นจริงถึงพฤติกรรมการเคลื่อนไหวของชุดตัวแปรภายในระบบสมการมีลักษณะเป็นเชิงเส้นตรงหรือไม่ (Linearity: against Vector STAR Model) และหาตัวแปรบ่งชี้ที่เหมาะสม ดังสมการที่ (4.12)

$$Y_t = \alpha_t + \sum_{j=1}^{P} \Gamma_j^0 \mathbb{Z}_{t-j}^0 + \sum_{j=1}^{P} \Gamma_j^1 \mathbb{Z}_{t-j}^1 + \sum_{j=1}^{P} \Gamma_j^2 \mathbb{Z}_{t-j}^2 + \sum_{j=1}^{P} \Gamma_j^3 \mathbb{Z}_{t-j}^3 + \varepsilon_t \qquad (4.12)$$

การประมาณค่าในสมการที่ (4.12) มีปัญหาอยู่ 2 ประเด็น ประเด็นแรกคือ ค่าความล่าช้าที่เหมาะสมควรจะเป็นเท่าไร คำตอบของคำถามนี้ จะใช้ค่าความล่าช้าที่เหมาะสมจะใช้ความล่าช้าที่เหมาะสมเดียวกับตัวแบบจำลอง VAR ส่วนประเด็นคำถามที่ 2 คือ ตัวแปรบ่งชี้หรือ (s_t) จะใช้ตัวแปรใด สำหรับตัวแปรที่คาดว่าจะเป็นตัวแปรบ่งชี้ในตัวแบบจำลอง Vector STAR คือ ค่าในอดีตของตัวแปรภายในระบบทั้ง 3 ตัว $\left[\text{exre\_y}_{t-1}, \text{slope}_{t-1}, \text{exrre\_set}_{t-1}\right]$ จากนั้นจะทำการทดสอบความเป็นเชิงเส้นตรง (Linearity: against Vector STAR Model) ในแต่ละครั้งโดยการแทนตัวแปรที่คาดว่าจะเป็นตัวแปรบ่งชี้ (s_t) ทั้ง 3 ตัว และใช้การทดสอบแบบ Wald Test โดยตั้งสมมติฐานดังนี้

$H_{01}: \Gamma^1 = \Gamma^2 = \Gamma^3 = 0$

$H_{11}: \Gamma^1 \neq \Gamma^2 \neq \Gamma^3 \neq 0$

ตัวแปรที่จะใช้เป็นตัวแปรบ่งชี้ (s_t) ในสมการ Vector STAR จะเป็นตัวแปรที่ปฏิเสธความเป็นเชิงเส้นตรงมากที่สุด ขั้นตอนต่อไปจะเป็นการกำหนดค่าพารามิเตอร์ของตัวแบบจำลอง Vector STAR ตามรายละเอียดในบทที่ 3

4.2.3 การทดสอบลักษณะตัวแบบจำลอง Vector STAR

เมื่อกำหนดพารามิเตอร์ในตัวแบบจำลอง Vector STAR แล้ว ขั้นตอนต่อไป เป็นการทดสอบความน่าเชื่อถือของค่าพารามิเตอร์ในตัวแบบจำลอง Vector STAR โดยมีสมมติฐานในการทดสอบ 2 สมมติฐาน

(1) ตัวแบบจำลอง Vector STAR เป็นแบบจำลอง ที่ประกอบด้วยแบบจำลองเชิงเส้นตรง 2 เส้น รวมกันแบบถ่วงน้ำหนักด้วยฟังก์ชันที่มีค่าระหว่าง 0 ถึง 1 ดังนั้น ถ้าพารามิเตอร์ของสมการเชิงเส้นตรงทั้ง 2 เท่ากัน แสดงว่าไม่มีความแตกต่างกันระหว่าง 2 regimes ทำให้ตัวแบบจำลอง Vector STAR ลดรูปเป็น ตัวแบบจำลอง VAR ดังนั้นในสมมติฐานแรกจึงต้องทำการทดสอบว่า ค่าพารามิเตอร์ของตัวแบบจำลองเชิงเส้นตรงทั้ง 2 สมการในตัวแบบจำลอง Vector STAR ในสมการที่ (4.1) มีค่าเท่ากันอย่างมีนัยสำคัญทางสถิติหรือไม่ โดยตั้งสมมติฐานหลักว่า

$H_0: \beta_{jk}^1 = \beta_{jk}^2$; สำหรับทุกๆ j k

$H_1: \beta_{jk}^1 \neq \beta_{jk}^2$; สำหรับทุกๆ j k

โดยที่

β_{jk}^i = สัมประสิทธิ์ของ $\mathbf{Y}_{k,t-j}$; i = 1 และ 2 ตามลำดับ

การทดสอบใช้ Wald Test มีการกระจายแบบ $\chi^2(h)$ โดยที่ h คือจำนวนเงื่อนไขในสมมติฐานหลัก ถ้าปฏิเสธสมมติฐานหลัก สามารถกล่าวได้ว่า ค่าพารามิเตอร์ของสมการเชิงเส้นตรงทั้งสองไม่เท่ากันอย่างมีนัยสำคัญทางสถิติ

(2) ค่าพารามิเตอร์ γ ในตัวแบบจำลอง Vector STAR เป็นค่าที่ชี้ถึงความเร็วของการเปลี่ยนแปลงจากฟังก์ชันหนึ่งไปอีกฟังก์ชันหนึ่ง ถ้า $\gamma = 0$ ทำให้ $G(s_t) = \frac{1}{2}$ ตัวแบบจำลอง Vector STAR จะกลายเป็นเพียงตัวแบบจำลอง VAR ดังนั้น จึงจำเป็นต้องทำการทดสอบว่าพารามิเตอร์ γ มีค่าเท่ากับศูนย์อย่างมีนัยสำคัญทางสถิติหรือไม่ โดยทำการทดสอบ t-Statistic.

4.2.4 <u>เปรียบเทียบความสามารถในการพยากรณ์ระหว่างตัวแบบจำลอง Vector STAR และตัวแบบจำลอง VAR</u>

ความสามารถในการพยากรณ์ของตัวแบบจำลอง สามารถดูจากค่าเฉลี่ยความคลาดเคลื่อนจากการพยากรณ์ คือ ค่าเฉลี่ยความแตกต่างระหว่างค่าคาดการณ์ที่ได้จากตัวแบบจำลองกับค่าที่แท้จริงของตัวแปร ณ เวลาเดียวกัน โดยตัวแบบจำลองที่ให้ค่าเฉลี่ยความคลาดเคลื่อนต่ำ ย่อมแสดงถึง ตัวแบบจำลองที่มีสามารถพยากรณ์ได้ดี ทั้งนี้ตัวแบบจำลองที่ดีควรจะมีความสามารถในการพยากรณ์ทั้งในช่วง In-Sample และ Out-of Sample ดังนั้นเพื่อเป็นการตรวจสอบแบบ Robustness Check จึงทำการตรวจสอบความแม่นยำของตัวแบบจำลองใน 2 กรณี คือ In-Sample Criteria และ Out-of Sample Criteria เพราะบางครั้ง ตัวแบบจำลองพยากรณ์ได้ดีในช่วง In-Sample แต่ในช่วง Out-of Sample อาจจะพยากรณ์ได้ไม่ดี ตัวอย่างเช่น การศึกษาตัวแบบจำลองที่พรรณนาพฤติกรรมการเคลื่อนไหวของอัตราแลกเปลี่ยนในปีค.ศ. 1983 ของ Meese and Rogoff (1983) จนถึง ปีค.ศ. 2003 ของ Kilian and Taylor (2003) ให้ข้อสรุปที่ใกล้เคียงกันคือ พบว่าตัวแบบจำลองที่ศึกษามีความสามารถในการพยากรณ์ที่ดีในช่วง In-Sample แต่กลับไม่ให้ค่าสถิติที่แสดงถึงสามารถพยากรณ์ช่วง Out-of Sample ไม่ดี

4.2.4.1 ช่วงข้อมูล In-Sample คือ การใช้ข้อมูลทั้งหมดที่มีเพื่อคำนวณหาค่าพารามิเตอร์ในตัวแบบจำลอง ค่าสถิติที่ชี้ถึงความแม่นยำของตัวแบบจำลอง ในช่วง In-Sample คือ ค่าเฉลี่ยของค่าสัมบูรณ์ของความคลาดเคลื่อน (Mean absolute Error: MAE) และ รากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนยกกำลังสอง (Root Mean Squared Error: RMSE)[5] ค่าสถิติทั้งสองแสดงถึงความคลาดเคลื่อนในการพยากรณ์ของตัวแบบจำลอง โดยที่ตัวแบบจำลองใดที่ให้ค่าสถิติทั้งสองต่ำ คือตัวแบบจำลองที่สามารถพยากรณ์ตัวแปรได้ใกล้เคียงกับค่าที่เกิดขึ้นจริงของตัวแปร ค่าสถิติทั้งสองได้แสดงสูตรคำนวณในสมการที่ (4.14) และ (4.15)

$$\text{Mean Absolute Error} = \frac{\sum_{t=1}^{T}|e_t|}{T} \qquad (4.14)$$

$$\text{Root Mean Squared Error} = \sqrt{\frac{\sum_{t=1}^{T}e_t^2}{T}} \qquad (4.15)$$

โดยที่

e_t = ค่าคลาดเคลื่อนการพยากรณ์ ณ เวลา t ; $e_t = (Y_t - \hat{Y}_t)$

Y_t = ข้อมูลที่แท้จริงของ Y ณ เวลา t

$\hat{Y}_t$ = ข้อมูลที่ได้จากการประมาณข้อมูล Y_t โดยใช้ตัวแบบจำลอง

T = จำนวนข้อมูล

4.2.4.2 ค่าสถิติที่แสดงความแม่นยำของแบบจำลองใน Out of Sample ซึ่งเป็นช่วงของข้อมูลที่มิได้นำมาใช้ในการกำหนดค่าพารามิเตอร์ในตัวแบบจำลอง ดังนั้นการทดสอบความสามารถของตัวแบบจำลองในช่วงข้อมูล Out of Sample จึงเป็นการทดสอบความสามารถในการพยากรณ์ตัวแบบจำลองอย่างแท้จริง อย่างไรก็ตามการทดสอบความสามารถในช่วงข้อมูล Out of Sample จะตรวจสอบ 2 วิธี

(1) Appending window sample squared prediction error เป็นการเปรียบเทียบความสามารถในการพยากรณ์ของตัวแบบจำลองด้วยการใช้ข้อมูลข่าวสารที่เกิดขึ้นใหม่ทุกครั้ง จึงทำให้จำนวนข้อมูลข่าวสารที่ใช้ในตัวแบบจำลองไม่เท่ากันในแต่ละครั้ง ค่าสถิติ Appending

---

[5] ความแตกต่างระหว่างค่าสถิติ Mean absolute Error กับ Root Mean Squared Error คือ Mean absolute Error จะให้น้ำหนักกับค่าความคลาดเคลื่อนทุกตัวเท่ากัน แต่ค่า Root Mean Squared Error จะให้น้ำหนักกับค่าความคลาดเคลื่อนไม่เท่ากัน ขึ้นอยู่กับขนาดของความคลาดเคลื่อน โดยจะให้น้ำหนักกับบริเวณปลายหางมากกว่า

window Sample squared prediction error มีวิธีการคำนวณดังนี้ คือใช้ข้อมูลตั้งแต่ตัวที่ 1 ถึง R (เรียกว่า in sample) เพื่อกำหนดชุดพารามิเตอร์ที่ AR ใช้สัญลักษณ์ (β_{AR}) หลังจากนั้นจะนำชุดพารามิเตอร์ที่ β_{AR} เพื่อประมาณค่า $\hat{Y}_{AR+1}$ และค่า $Error^2_{AR+1} = (Y_{R+1} - \hat{Y}_{AR+1})^2$ จากนั้นจะใช้ข้อมูลตั้งแต่ตัวที่ 1 ถึง R+1 กำหนดค่าชุดพารามิเตอร์ที่R+1 ใช้สัญลักษณ์ (β_{AR+1}) และได้ใช้ชุดพารามิเตอร์ β_{AR+1} นำไปประมาณค่า $\hat{Y}_{AR+2}$ และค่า $Error^2_{AR+2} = (Y_{AR+2} - \hat{Y}_{AR+2})^2$ ทำอย่างนี้ต่อไปเรื่อยๆจนถึงข้อมูลตัวที่ T และหาค่าเฉลี่ยของ $Error^2 = \left(Error^2_{AR+1} + Error^2_{AR+2} + \ldots\ldots + Error^2_T\right) \div (T-R)$ ค่าสถิติที่ได้เรียกว่า Appending window Sample squared prediction error ตัวแบบจำลองที่สามารถพยากรณ์ที่ดี จะต้องให้ค่าสถิติ Appending window Sample squared prediction error ที่ต่ำ

(2) Rolling mean squared prediction error เป็นการเปรียบเทียบความสามารถในการพยากรณ์ของตัวแบบจำลองด้วยการใช้จำนวนข้อมูลข่าวสารที่เท่ากันทุกครั้ง ในการพยากรณ์โดยมีวิธีการคำนวณ ดังนี้คือใช้ข้อมูลตั้งแต่ตัวที่ 1 ถึง R (เรียกว่า in sample) เพื่อกำหนดชุดพารามิเตอร์ที่ RR ใช้สัญลักษณ์ (β_{RR}) หลังจากนั้นจะนำชุดพารามิเตอร์ที่ β_{RR} ไปประมาณค่า $\hat{Y}_{RR+1}$ และคำนวณค่า $Error^2_{RR+1} = (Y_{R+1} - \hat{Y}_{RR+1})^2$ จากนั้นจะใช้ข้อมูลตั้งแต่ตัวที่ 2 ถึง R+1 กำหนดค่าชุดพารามิเตอร์ที่ R+1 ใช้สัญลักษณ์ (β_{RR+1}) นำชุดพารามิเตอร์ที่ β_{RR+1} ประมาณค่า $\hat{Y}_{RR+2}$ และค่า $Error^2_{R+2} = (Y_{R+2} - \hat{Y}_{R+2})^2$ ทำอย่างนี้ต่อไปเรื่อยๆ จนถึงข้อมูลตัวที่ T และหาค่าเฉลี่ยของ $Error^2 = \left(Error^2_{RR+1} + Error^2_{RR+2} + \ldots\ldots + Error^2_T\right) \div (T-R)$ ค่าสถิติที่ได้เรียกว่า Rolling mean squared prediction error ตัวแบบจำลองที่สามารถพยากรณ์ที่ดี จะต้องให้ค่าสถิติ Rolling mean squared prediction error ที่ต่ำ

ค่าสถิติ Appending window Sample squared prediction error และ Rolling mean squared prediction error มีข้อดีและข้อเสียแตกต่างกัน กล่าวคือ หากมีการเปลี่ยนแปลงโครงสร้างเศรษฐกิจแล้วการใช้วิธี Appending window Sample squared prediction error จะบรรจุข้อมูลข่าวสารทั้งช่วงก่อนและหลังการเปลี่ยนแปลงโครงสร้างเศรษฐกิจ อาจจะทำให้มีการใช้ข้อมูลที่ไม่จำเป็น ส่งผลให้ความสามารถในการพยากรณ์ของตัวแบบจำลองลดลง แต่หากใช้วิธี Rolling mean squared prediction error จะให้ผลที่ดีกว่า เพราะวิธี Rolling mean squared prediction error ได้มีการตัดข้อมูลในอดีตที่ไม่จำเป็นต่อการพยากรณ์ออกไปได้เนื่องจากพฤติกรรมของตัวแปรได้เปลี่ยนแปลงไปแล้วเมื่อมีการเปลี่ยนแปลงโครงสร้างเศรษฐกิจ ในทางตรงกันข้าม ถ้าไม่มีการเปลี่ยนแปลงโครงสร้างเศรษฐกิจการใช้วิธี Appending window Sample squared prediction error ย่อมจะให้ผลดีกว่าการใช้วิธี Rolling mean squared prediction

error เพราะในการพยากรณ์แต่ละครั้งได้ใช้ข้อมูลที่เพิ่มขึ้นซึ่งเป็นข้อมูลที่ช่วยในการพรรณนาพฤติกรรมการเคลื่อนไหวของตัวแปร

อย่างไรก็ตามทั้ง 2 วิธี ต่างมีประเด็นคำถามเดียวกันคือ จำนวนข้อมูล R ที่ใช้ในการพยากรณ์ควรจะเป็นเท่าไร คำตอบของคำถามนี้ คือ เนื่องจากการศึกษาครั้งนี้ในระบบสมการและจำนวนพารามิเตอร์ที่มากในตัวแบบจำลอง Vector STAR จึงจำเป็นต้องใช้จำนวนข้อมูลที่ใหญ่ในการกำหนดค่าพารามิเตอร์ที่มีประสิทธิภาพ ดังนั้นจึงเลือกจำนวนข้อมูล R = 1,000 ข้อมูลในการพยากรณ์ โดยจะทำการพยากรณ์ไปข้างหน้า 300 ครั้ง

ขั้นต่อไปจะทำการทดสอบความแม่นยำในการพยากรณ์ระหว่างตัวแบบจำลอง Vector STAR และตัวแบบจำลอง VAR ในช่วง Out of Sample (out of sample test of predictive ability) ถ้าตัวแบบจำลอง VAR และ Vector STAR ให้ค่าความคลาดเคลื่อนที่ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ย่อมหมายความว่า ตัวแบบจำลองทั้งสองมีความแม่นยำในการพยากรณ์ไม่แตกต่างกันอย่างมีนัยสถิติเช่นกัน การทดสอบความแม่นยำในการพยากรณ์จะทำทั้ง Appending window sample squared prediction errors และ Rolling mean squared prediction errors โดยมีหลักสมมติฐานดังนี้[6]

H_{10} : Appending window Sample SPE_{VAR} = Appending window Sample $SPE_{Vector\ STAR}$

H_{11} : Appending window Sample SPE_{VAR} ≠ Appending window Sample $SPE_{Vector\ STAR}$

H_{20} : Rolling $MSPE_{VAR}$ = Rolling $MSPE_{Vector\ STAR}$

H_{21} : Rolling $MSPE_{VAR}$ ≠ Rolling $MSPE_{Vector\ STAR}$

Diebold and Mariano (1995) และ West (1996) เสนอให้ทำการทดสอบความสามารถของตัวแบบจำลองในช่วง Out of Sample โดยใช้ MSE ค่าสถิติในการทดสอบแสดงให้เห็นในสมการที่ (4.16)

$$MSE = P^{1/2} \times \frac{P^{-1}\sum_{t=1}^{P}\left(\hat{e}^2_{STAR,t} - \hat{e}^2_{VAR,t}\right)}{\sqrt{P^{-1}\sum_{t=1}^{P}\left(\hat{e}^2_{STAR,t} - \hat{e}^2_{VAR,t}\right)^2}} \sim H_0\ N(0,1) \qquad (4.16)$$

---

[6] ไม่สามารถทดสอบความสามารถในการพยากรณ์ระหว่างวิธี Appending window sample squared prediction errors และ Rolling mean squared prediction errors ได้ เนื่องจากจำนวนข้อมูลที่ใช้ในการกำหนดค่าพารามิเตอร์ในแต่ละวิธีไม่เท่ากัน

โดยที่

$e^2_{VAR,t}$ = ความคลาดเคลื่อนในการพยากรณ์ของตัวแบบจำลอง VAR ณ เวลา t (Forecast Error); $e^2_{VAR,t} = Y_t - \hat{Y}_{VAR,t}$

$e^2_{STAR,t}$ = ความคลาดเคลื่อนในการพยากรณ์ของตัวแบบจำลอง STAR ณ เวลา t (Forecast Error); $e^2_{STAR,t} = Y_t - \hat{Y}_{STAR,t}$

Y_t = ข้อมูลที่แท้จริงของ Y ณ เวลา t

$\hat{Y}_{VAR,t}$ = ข้อมูลที่ได้จากการประมาณค่าตัวแปร Y_t โดยใช้ตัวแบบจำลอง VAR

$\hat{Y}_{STAR,t}$ = ข้อมูลที่ได้จากการประมาณค่าตัวแปร Y_t โดยใช้ตัวแบบจำลอง STAR

P = จำนวนข้อมูลที่มีการพยากรณ์ในช่วง Out of Sample (the number of 1-step of ahead prediction out of sample)

ภายใต้ Null Hypothesis (H_0) ที่ว่าตัวแบบจำลอง VAR และ Vector STAR มีความสามารถในการพยากรณ์ไม่แตกต่างกัน ค่าสถิต MSE จะ มีการกระจายแบบปกติ ถ้าการทดสอบสามารถปฏิเสธสมมติฐานหลักได้ จะสรุปว่า ความสามารถในการพยากรณ์ในช่วง Out of Sample ของตัวแบบจำลอง VAR และ Vector STAR มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ