Return Predictability

ผศ.ดร. วีระชาติ กิเลนทอง
มหาวิทยาลัยหอการค้าไทย

Efficient Market Hypothesis

- ทฤษฎีประสิทธิภาพของตลาด (Efficient Market Hypothesis of Fama 1970) คือ การที่ราคา $p_{j,t}$ สะท้อนถึง ``ข้อมูล'' ที่มีอยู่ทั้งในอดีตและปัจจุบันอย่างสมบูรณ์
- เราสามารถแปลความหมายให้อยู่ในรูปของสมการคณิตศาสตร์ได้ว่า ค่าผลต่างระหว่างราคาที่ $t+1$, $p_{j,t+1}$, และค่าความคาดหวังของราคาที่ $t+1$, $E[p_{j,t+1}|I_t]$ ภายใต้ข้อมูลในอดีตทั้งหมดจนถึงเวลา t, I_t, จะต้องเป็นตัวแปรสุ่มที่มีค่าความคาดหวังเท่ากับศูนย์

$$x_{j,t+1} = p_{j,t+1} - E[p_{j,t+1}|I_t] \tag{1}$$

และ

$$E(x_{j,t+1}|I_t) = 0 \tag{2}$$

- ปรากฏการณ์นี้ถูกเรียกว่า ``fair game'' ซึ่งหมายความว่า หากเราซื้อสินทรัพย์นี้ที่ราคาที่จุดเริ่มต้นของเวลา t เท่ากับค่าความคาดหวังของราคาที่ $t+1$, $E[p_{j,t+1}|I_t]$, และขายสินทรัพย์ที่จุดเริ่มของเวลา $t+1$ และก็ซื้อสินทรัพย์นี้ใหม่ที่ราคา $E[p_{j,t+2}|I_{t+1}]$ และทำอย่างนี้ไปเรื่อยๆ Efficient Market Hypothesis บอกว่า เราควรจะได้กำไรสุทธิเท่ากับศูนย์ ดังสมการข้างต้น
- ไม่มีใครมีอำนาจผูกขาดในการใช้ข้อมูลใดๆ เพื่อสร้างกำไรได้

Efficient Market Hypothesis

เงื่อนไขที่เพียงพอ (sufficient condition) ของทฤษฎีประสิทธิภาพของตลาด

- ไม่มี transection cost ในการซื้อขาย
- ไม่มีค่าใช้จ่ายในการเข้าถึงข้อมูลที่เกี่ยวกับตลาด
- นักลงทุนทุกคนเห็นด้วยกับบทบาทของข้อมูลที่มีต่อราคาของสินทรัพย์ในปัจจุบันและการกระจายตัวของความคาดหวังของราคาในอนาคตของแต่ละหลักทรัพย์

Efficient Market Hypothesis

ประสิทธิภาพของตลาดขึ้นอยู่กับข่าวสารข้อมูล (I_t) ซึ่งแบ่งได้เป็น 3 ระดับดังนี้

from of Market Efficiency	Market prices Reflect		
	past Market data	Public Information	Private Information
weak form	x		
semi-strong form	x	x	
strong form	x	x	x

Predictability Test

- เราใช้ข้อมูลราคาในอดีตทดสอบประสิทธิภาพของตลาดแบบอ่อน (weak form test) โดยการทดสอบ สมมติฐานการเดินแบบสุ่ม (Random Walk Hypothesis) ดังต่อไปนี้

$$\mathrm{p}_{t+1} = \mu + \mathrm{p}_t + \epsilon_{t+1} \tag{3}$$

โดยที่

- $\mathrm{p}_{t+1} = \ln \mathrm{P}_{t+1}$ คือค่าของ ล็อกการิทึมของราคาหลักทรัพย์ในช่วงเวลา $t+1$
- μ คือ ค่าความคาดหวังของการเปลี่ยนแปลงของราคาในหลักทรัพย์
- $\epsilon_{t+1} \sim \mathrm{IID}(0, \sigma^2)$ หมายความว่า ค่าความคาดเคลื่อนนั้นมีการกระจายตัวแบบอิสระและเหมือนกันทุกช่วงเวลา (independently and identically distribution)

Random Walk Hypothesis

หากเราสมมติให้ $\mu = 0$

$$p_{t+1} = p_t + \epsilon_{t+1} \tag{4}$$

ดังนั้น ค่าความคาดหวังของราคาในวันพรุ่งนี้จะเท่ากับราคาในวันนี้

$$E[p_{t+1}|p_t, p_{t-1}, \cdots] = E[p_t|p_t, p_{t-1}, \cdots] + E[\epsilon_{t+1}|p_t, p_{t-1}, \cdots] = p_t \tag{5}$$

ซึ่งเราเรียกแบบจำลองนี้ว่า martingale model

Random Walk Hypothesis

อัตราผลตอบแทน (r_t)

$$E[p_{t+1} - p_t | p_t, p_{t-1}, \cdots] = E[r_{t+1} | p_t, p_{t-1}, \cdots] = E[\epsilon_{t+1} | p_t, p_{t-1}, \cdots] = 0 \quad (6)$$

โดย ที่ r_t คือ อัตราผลตอบแทนสุทธิ (net return) หมายความว่า

$$r_{t+1} = \epsilon_{t+1} \quad (7)$$

ในกรณีที่ มี drift term ดังนั้น เขียนได้เป็น

$$r_{t+1} = \mu + \epsilon_{t+1} \quad (8)$$

ดังนั้น สมมุติฐานที่เราต้องการทดสอบในส่วนนี้ คือ การทดสอบว่า อัตราผลตอบแทนสุทธิ (net return) มีลักษณะของการเดินแบบสุ่มหรือไม่?

Random Walk Hypothesis I

- สมมุติฐานการเดินสุ่มแบบที่ 1 (Random Walk Hypothesis I)
  - $\mathrm{r}_{t+1} \sim \mathrm{IID}(0, \sigma^2)$ มีการกระจายตัวแบบอิสระและเหมือนกันทุกช่วงเวลา (independently and identically distribution)
- การทดสอบ สมมุติฐานการเดินสุ่มแบบที่ 1 คือการทดสอบความเป็น IID ของอัตราผลตอบแทนสุทธิ ซึ่งมีอยู่หลายวิธี เช่น
  - วิธีการใช้อนุกรมและส่วนกลับ (sequence and reversals)
  - การทดสอบโดยจำนวนอนุกรม (Runs Test)

Sequence and Reversals

- จากแบบจำลอง

$$p_{t+1} = \mu + p_t + \epsilon_{t+1} \tag{9}$$

เราสามารถสร้างตัวแปรใหม่เพื่อใช้ในการทดสอบสมมุติฐานดังต่อไปนี้

$$I_t = \begin{cases} 1 & \text{if} \quad r_t = p_t - p_{t-1} > 0 \quad \text{ด้วยความน่าจะเป็น} \quad \pi = \Phi(\frac{\mu}{\sigma}) \\ 0 & \text{if} \quad r_t = p_t - p_{t-1} < 0 \quad \text{ด้วยความน่าจะเป็น} \quad 1 - \pi \end{cases} \tag{10}$$

และ จำนวนครั้งที่มีจำนวนคู่ที่มีทิศทางการเปลี่ยนแปลงของราคาเหมือนกัน

$$T_s = \sum_{t=1}^{n} Y_t, \tag{11}$$

โดยที่

$$Y_t = I_t * I_{t+1} + (1 - I_t)(1 - I_{t+1}) = \begin{cases} 1, & \text{with } \pi_s = \pi^2 + (1-\pi)^2; \\ 0, & \text{with } 1 - \pi_s \end{cases}$$

Sequence and Reversals

- Cowles 3rd and Jones (1937) คำนวณค่าสถิติ (statistics)

$$\widehat{\mathrm{CJ}} = \frac{\mathrm{T_s}}{\mathrm{T} - \mathrm{T_s}}$$

เพื่อใช้ในการทดสอบว่า ϵ_{t+1} เป็น IID หรือไม่ โดยที่ $\mathrm{T}+1$ คือจำนวนข้อมูล (ช่วงเวลา) ที่มีอยู่ทั้งหมด

- นอกจากนี้ Cowles 3rd and Jones (1937) ได้คำนวณว่า การกระจายตัวของ $\widehat{\mathrm{CJ}}$ ภายใต้ null Hypothesis จะอยู่ในรูปของ

$$\widehat{\mathrm{CJ}} \sim \mathrm{N}\left(\frac{\pi_s}{1-\pi_s}, \frac{\pi_s(1-\pi_s) + 2(\pi^3 + (1-\pi)^3 - \pi_s^2)}{\mathrm{T}(1-\pi_s^4)}\right) \tag{12}$$

- ยกตัวอย่างเช่น หากเราประมาณค่า $\mu = 0$ (ด้วยการประมาณค่าสมการ (9)) เราจะสามารถคำนวณได้ว่า

$$\pi = \Phi(0) = \frac{1}{2} = \pi_s$$

$$\widehat{\mathrm{CJ}} \sim \mathrm{N}\left(1, \frac{4}{\mathrm{T}}\right)$$

Runs Test: Mood (1940)

- การทดสอบนี้คล้ายๆ กับแบบแรก ที่เราจะทำการนับตัวที่มีทิศทางที่เหมือนกัน แต่ต่างกันตรงที่ไม่ได้สนใจเป็นคู่ แต่จะสนใจเป็นกลุ่มๆ แทน โดยเปรียบเทียบกับอนุกรมที่คาดหวังภายใต้ null hypothesis
- ตัวอย่างการนับ ``11001100'' มี $N_{runs} = 4$ และ 11110000 มี $N_{runs} = 2$
- ค่าสถิติสำหรับการทดสอบเขียนได้ดังต่อไปนี้

$$z = \frac{N_{runs} - 2T\pi(1-\pi)}{2\sqrt{(T\pi(1-\pi)(1-3\pi(1-\pi))}} \sim N(0,1) \tag{13}$$

โดยที่ $\pi = \Phi\left(\frac{\mu}{\sigma}\right)$

- Wallis and Roberts(1956) ปรับปรุงสมการข้างต้นเพื่อแก้ไขความเบี่ยงเบนที่เกิดจากการประมาณเป็น

$$z = \frac{N_{runs} + \frac{1}{2} - 2T\pi(1-\pi)}{2\sqrt{(T\pi(1-\pi)(1-3\pi(1-\pi))}} \sim N(0,1) \tag{14}$$

ตัวอย่างการแสดง การทดสอบ โดยใช้ run test

FEDR	runtest		
	$\hat{z}$	P-value	test-hypothesis
Weekly(2002-2014)	-0.45	0.65	fail to reject
monthly(1975-2002)	-3.26	0	reject

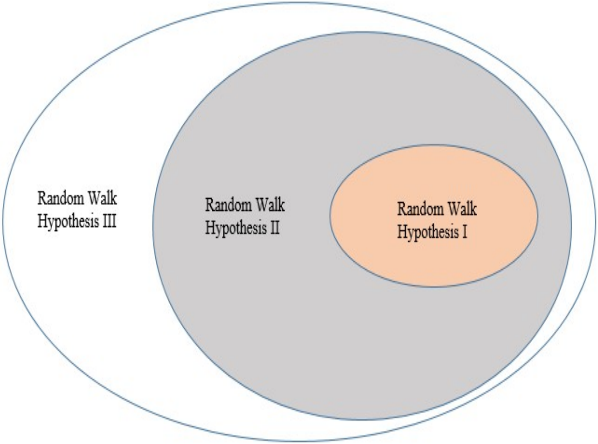
Random Walk Hypothesis II

- ไม่จำเป็นต้องมีการกระจายตัวเหมือนกันทุกช่วงเวลา แต่ยังคงเป็นอิสระต่อกันอยู่ (independently but not identically distributed:INID) $\mathrm{E}\left[\epsilon_{t+k}|\epsilon_t\right] = 0$
- สมมติฐานการเดินสุ่มแบบที่ 1 (Random Walk Hypothesis I) เป็นส่วนหนึ่ง ของสมมติฐานการเดินสุ่มแบบที่ 2 (Random Walk Hypothesis II)
- การทดสอบการเป็นอิสระต่อกัน แต่ไม่จำเป็นต้องมีการกระจายตัวที่เหมือนกัน ซึ่งการทดสอบนั้นค่อนข้างยุ่งยากจึงไม่เป็นที่นิยม

Random Walk Hypothesis III

- ถูกเรียกว่าเป็น "weakest test" เพราะ แค่ทดสอบว่าตัวแปรที่ทดสอบไม่มีความสัมพันธ์ข้ามเวลา (uncorrelated over time) การทดสอบแบบนี้ เป็นที่นิยมในปัจจุบัน
- ตัวอย่าง เช่น ข้อมูลที่มี $\mathrm{Cov}(\epsilon_t, \epsilon_{t+k}) = 0$ (serial uncorrelation) แต่ $\mathrm{Cov}(\epsilon_t^2, \epsilon_{t+k}^2) \neq 0$
- การทดสอบกลุ่มนี้ทำได้หลายแบบ เช่น
  - การทดสอบโดยใช้สัมประสิทธิ์สหสัมพันธ์(Autocorrelation Coefficient)
  - การทดสอบโดยสัดส่วนของความแปรปรวน(Variance Ratio)

Random Walk Hypothesis III



Autocorrelation Coefficient

- Autocorrelation Coefficient Test เป็นการทดสอบ การไม่มีความสัมพันธ์ข้ามเวลา (serial uncorrelation) อย่างง่าย โดยใช้ ค่าสัมประสิทธิ์สหสัมพันธ์ ระหว่างผลตอบแทนที่ต่างกัน k ช่วง

$$\hat{\rho}(k) = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} \tag{15}$$

เมื่อ

$$\hat{\gamma}(k) = \frac{1}{T}\sum_{t=1}^{T-k}(r_t - \bar{r}_T)(r_{t+k} - \bar{r}_T) \tag{16}$$

และ $\bar{r}_T = \frac{\sum_{t=1}^{T} r_t}{T}$

- ภายใต้ null hypothesis ค่าสัมประสิทธิ์สหสัมพันธ์ จะมีการกระจายตัวดังต่อไปนี้

$$\sqrt{T}\hat{\rho}(k) \sim N(0,1) \tag{17}$$

Variance Ratio

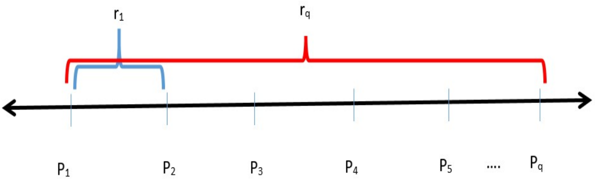
- Variance Ratio ประยุกต์มาจากการทดสอบ Autocorrelation Coefficient แต่มีข้อดีคือสามารถทำได้ภายใต้เงื่อนไขที่ ϵ มี heteroskedasticity
- หลักการที่สำคัญคือ การเปรียบเทียบ variance ของอัตราผลตอบแทนสุทธิของการถือครองหลักทรัพย์ในระยะเวลาที่ต่างกัน
- variance ของอัตราผลตอบแทนสุทธิของการถือครองหลักทรัพย์ 1 ช่วงเวลา เท่ากับ

$$\mathrm{Var}(p_t - p_{t-1}) = \mathrm{Var}(r_1) = \sigma^2 \tag{18}$$

- variance ของอัตราผลตอบแทนสุทธิของการถือครองหลักทรัพย์ q ช่วงเวลา เท่ากับ

$$\begin{aligned}\mathrm{Var}(r_q) &= \mathrm{Var}([p_t - p_{t-1}] + [p_{t-1} - p_{t-2}] + \cdots + [p_{t-(q-1)} - p_{t-q}]) \\ &= q\mathrm{Var}(r_1) = q\sigma^2 \end{aligned} \tag{19}$$

Variance Ratio



Variance Ratio

- ภายใต้ null hypothesis H_0:

$$VR(q) = \frac{Var(r_q)}{qVar(r_1)} = 1 \tag{20}$$

- ค่าสถิติเพื่อทดสอบสมมุติฐาน ในกรณีที่มี homoskedasticity คือ

$$z(q) = \frac{VR(q) - 1}{\sqrt{(\sigma_{ho})}} \sim N(0, 1) \tag{21}$$

เมื่อ $\sigma_{ho} = \frac{2(2q-1)(q-1)}{3qN}$ และ N จำนวนข้อมูลของอัตราผลตอบแทนสุทธิ

- ค่าสถิติเพื่อทดสอบสมมุติฐาน ในกรณีที่มี heterskedasticity คือ

$$z(q) = \frac{VR(q) - 1}{\sqrt{(\sigma_{he})}} \sim N(0, 1) \tag{22}$$

เมื่อ $\sigma_{he} = \frac{4\sum_{\tau=1}^{q-1}(q-\tau)^2 b_\tau}{q^2 N}$ และ $b_\tau = \frac{N\sum_{i=1}^{N-\tau}(r_i-\bar{r})^2(r_{i+\tau}-\bar{r})^2}{(\sum_{i=1}^{N}(r_i-\bar{r})^2)^2}$

ตัวอย่างการทดสอบ โดยใช้ run test และ variance ratio

FEDR	runtest			variance ratio test		
	$\hat{z}$	P-value	test-hypothesis	$\hat{z}$	P-value	test-hypothesis
Weekly(2002-2014)	-0.45	0.65	fail to reject	-0.18	0.85	fail to reject
Monthly(1975-2002)	-3.26	0	reject	1.09	0.28	fail to reject

- การทดสอบ Random Walk Hypothesis ของ ข้อมูลรายสัปดาห์ (Weekly) ``ยอมรับสมมติฐาน'' ของ Random Walk Hypothesis II และ``ยอมรับสมมติฐาน'' ของ Random Walk Hypothesis III
- การทดสอบ Random Walk Hypothesis ของข้อมูลรายเดือน (Monthly) ``ยอมรับสมมติฐาน'' ของ Random Walk Hypothesis III เท่านั้น
- ดังนั้น ถ้าเราทดสอบแค่ variance ratio จะยอมรับเพียง Random Walk Hypothesis III เท่านั้น

Regression Predictability

การทดสอบ Random Walk Hypothesis ของอัตราผลตอบแทนนั้น เป็นการทดสอบที่เกี่ยวกับประสิทธิภาพของตลาด (efficient market hypothesis) โดยดูรูปแบบของราคาในอดีต แต่อย่างไรก็ตามก็ยังมีตัวแปรอื่นๆ ที่อาจจะสามารถพยากรณ์ราคาหรือผลตอบแทนได้ เช่น

$$r_{t+1} = a + bx_t + \epsilon_{t+1} \tag{23}$$

- r_{t+1} คือ อัตราผลตอบแทนสิทธิในช่วงเวลา t+1
- x_t คือ ตัวแปรการพยากรณ์ที่เราทราบค่า(obsevred variable)
- ϵ_{t+1} คือ ตัวแปรสุ่มในช่วงเวลา t+1

ชัดเจนว่า ถ้าอัตราผลตอบแทนสิทธิสามารถคาดเดาได้ หมายความว่า

- สัมประสิทธิ์ b มีค่าไม่เท่ากับ 0 “อย่างมีนัยสำคัญ ”คือ x_t สามารถพยากรณ์ผลตอบแทนได้
- แบบจำลองสามารถพยากรณ์ได้ดี เมื่อ R^2 มีค่าสูง

Regression Predictability

- เราสามารถหาความเชื่อมโยงระหว่างสมการพยากรณ์ (23) และแบบจำลองที่ใช้ทดสอบจากสมการ (7) หรือ (8) ดังต่อไปนี้
- หากเราสมมติให้ x_t ไม่สามารถพยากรณ์ผลตอบแทนได้ เราหา conditional expectation ได้เป็น

$$\begin{aligned} r_{t+1} &= a + \epsilon_{t+1} && (24) \\ E_t(r_{t+1}) &= a && (25) \end{aligned}$$

- แต่หากเราสมมติให้ x_t สามารถพยากรณ์ผลตอบแทนได้ เราหา conditional expectation ได้เป็น

$$\begin{aligned} r_{t+1} &= a + bx_t + \epsilon_{t+1} && (26) \\ E_t(r_{t+1}) &= a + bx_t && (27) \end{aligned}$$

Regression Predictability

- ตัวอย่างที่ง่ายที่สุดของตัวอย่างนี้ คือ

$$r_{t+1} = a + br_t + \epsilon_{t+1} \tag{28}$$

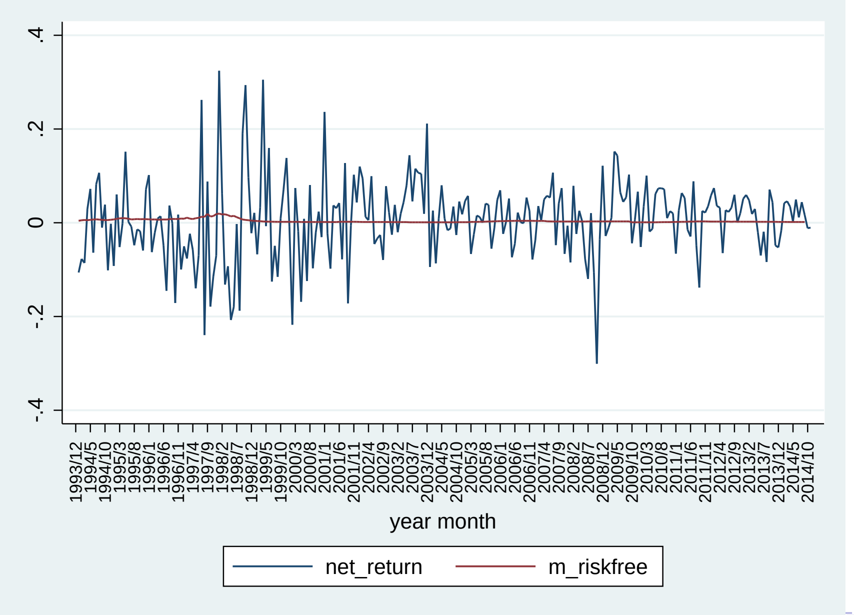
- เรามักจะใช้วิธีกำลังสองน้อยสุด (Ordinary Least Square: OLS) ในการประมาณค่าพารามิเตอร์ของสมการ Regression Predictability โดยผลลัพธ์ที่เราสนใจคือ ค่าสัมประสิทธิ์ b
  - หาก $b > 0$ อย่างมีนัยสำคัญ เราก็จะสรุปว่า หลักทรัพย์ที่ผลตอบแทนดีในอดีตจะเป็นไปได้สูงที่จะให้ผลตอบแทนที่ดีในช่วงเวลาต่อมา ซึ่งเราเรียกกว่า "momentum"
  - หาก $b < 0$ อย่างมีนัยสำคัญ เราก็จะสรุปว่า หลักทรัพย์ที่ผลตอบแทนดีในอดีตจะเป็นไปได้สูงที่จะให้ผลตอบแทนที่ไม่ดีในช่วงเวลาต่อมา ซึ่งเราเรียกกว่า "mean reversion"

Regression Predictability

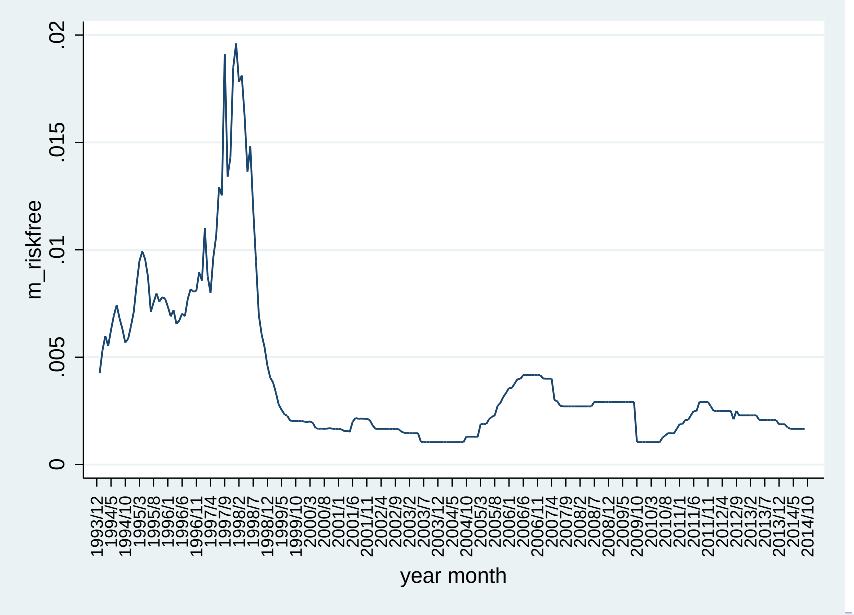
Table: ตารางเปรียบเทียบสัมประสิทธิ์อัตราผลตอบตอบแทนรายเดือนจากสมการ (28)

	b	t(b)	p-value	standard error of b	R^2
stock(1975-2013)	0.019999	0.12	0.905	0.1666939	0.0004
พันธบัตรซื้อคืน 1 เดือน (1994-2013)	0.688727	4.11	0.001	0.1675811	0.4984

Regression Predictability



Regression Predictability



Long horizon Regression

- ที่ผ่านมา ดูเหมือนว่า เราจะไม่สามารถพยากรณ์ผลตอบแทนระยะสั้นได้
- คำถามต่อไปก็คือ
  - แล้วเราสามารถสามารถพยากรณ์ผลตอบแทนระยะยาว (ผลตอบแทนจากการถือหลักทรัพย์ระยะยาว เช่น 1 ปี เมื่อเทียบกับ 1 เดือน) ได้หรือไม่?
  - ถ้าได้ ตัวแปรอะไรที่ช่วยเราพยากรณ์ได้ดี?
- ยกตัวอย่างเช่น

$$r^e_{t\rightarrow t+k} = a + b\frac{D_t}{P_t} + \epsilon_{t+k} \tag{29}$$

เมื่อ excess return (r^e_t) ได้มาจาก

$$r^e_t = (R_t - 1 - r_{ft}) \tag{30}$$

และ

$$r^e_{t\rightarrow t+k} = \sum_{i=1}^{k} r^e_{t+i} \tag{31}$$

Long horizon Regression

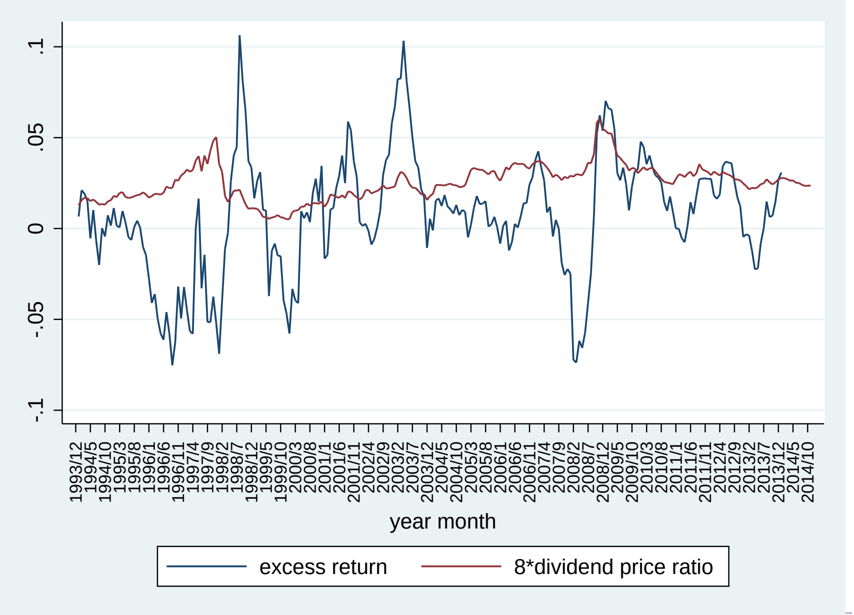
Table: long horizon regression of excess return

k(months)	b	t(b)	R^2
1	0.0026	0.46	0.0014
5	0.0456	2.99	0.0579

จากข้อมูลรายเดือนตั้งแต่ 01-1994 จนถึง 10-2014

ผลการประมาณการณ์ชี้ให้เห็นอย่างชัดเจนว่า เราสามารถพยากรณ์อัตราผลตอบแทนในช่วง 5 เดือน ได้ดีกว่าช่วง 1 เดือน

Long horizon Regression and Persistent Variable



Long horizon Regression and Persistent Variable

- คำถามก็คือ ทำไม $\frac{D_t}{P_t}$ สามารถช่วยพยากรณ์อัตราผลตอบแทนได้
- เราสามารถเขียนอัตราผลตอบแทนได้เป็น

$$
\begin{aligned}
R_{t+1} &= \frac{P_{t+1} + D_{t+1}}{P_t} && (32) \\
&= \frac{D_t}{P_t}\left(\frac{D_{t+1}}{D_t}\right) + \frac{P_{t+1}}{P_t} && (33)
\end{aligned}
$$

ซึ่งจะเห็นได้ว่า R_{t+1} มีความสัมพันธ์เชิงคณิตศาสตร์กับ $\frac{D_t}{P_t}$

Long horizon Regression and Persistent Variable

- คำตอบที่ถูกต้องกว่าก็คือ ตัวแปรที่จะช่วยพยากรณ์ได้ดีคือ ตัวแปรที่มีพฤติกรรมสอดคล้องกับสมการต่อไปนี้

$$x_{t+1} = \mu + \rho x_t + \delta_{t+1} \tag{34}$$

  โดยที่ ρ ต้องมีค่าเข้าใกล้ 1 มากๆ

- เราเรียกตัวแปรนี้ว่า “persistent variable”
- ยิ่งไปกว่านั้น เราสามารถเขียน Long horizon Regression ของอัตราผลตอบแทน ในรูปแบบทั่วไป ได้ดังต่อไปนี้

$$\begin{aligned} r_{t+1} &= a + bx_t + \epsilon_{t+1} && (35) \\ x_{t+1} &= \mu + \rho x_t + \delta_{t+1} && (36) \end{aligned}$$

R^2 of Long Horizon Regression Increases with n

- ถ้าเรากำหนดค่าคงที่ในทั้ง 2 สมการเป็น 0 และสมมติว่าเราลงทุนไป 2 ช่วงเวลา

$$
\begin{aligned}
r_{t\to t+2} = r_{t+1} + r_{t+2} &= \left(bx_{t+1} + \epsilon_{t+2}\right) + \left(bx_t + \epsilon_{t+1}\right) && (37)\\
&= b\rho x_t + b\delta_{t+1} + bx_t + \epsilon_{t+2} + \epsilon_{t+1} && (38)\\
&= b(1+\rho)x_t + \left(b\delta_{t+1} + \epsilon_{t+2} + \epsilon_{t+1}\right) && (39)
\end{aligned}
$$

- ถ้าเราถืออัตราผลตอบแทน T ปี จะได้ว่า

$$
\sum_{i=1}^{T} r_{t+i} = b\left(\sum_{i=1}^{T}\rho^{i-1}\right)x_t + b\sum_{i=1}^{T-1}\left(\sum_{j=1}^{T-i}\rho^{T-i-j}\right)\delta_{t+i} + \sum_{i=1}^{T}\epsilon_{t+i} \quad (40)
$$

R^2 of Long Horizon Regression Increases with n

ถ้าเรากำหนด $\rho = 1$ ดังนั้น

$$\sum_{i=1}^{T} r_{t+i} = Tbx_t + b\sum_{i=1}^{T-1}(T-i)\delta_{t+i} + \sum_{i=1}^{T}\epsilon_{t+i} \tag{41}$$

เห็นได้ชัดว่า

- สัมประสิทธิ์ของ x_t มีขนาดมากขึ้นเป็น T เท่า คือ Tb
- ขนาดของ R^2 มีค่ามากขึ้นเป็น T เท่า จาก

$$R^2 = \frac{Var(Tbx_t)}{Var(\sum_{i=1}^{T} r_{t+i})} = T^2\frac{Var(bx_t)}{TVar(r_{t+1})} = T\frac{Var(bx_t)}{Var(r_{t+1})}$$

โดยสรุป จะเห็นว่า ความสามารถพยากรณ์ของ persistent variable (ในแง่ของ R^2) จะเพิ่มขึ้นตามระยะเวลาที่ถือครองทรัพย์สิน