อ่านกระดาน
Reading the Book

Part 7: Order Flow Imbalance & Price Impact — flow ดันราคายังไง

Order Flow & Microstructure สำหรับรายย่อย · เลนส์ execution ของซีรีส์

จากภาพนิ่ง (OBI) → วิดีโอการเปลี่ยน (OFI) → วัดแรงดันราคาด้วย Kyle's λ

Part 7 — OFI & Price Impact

🎯 **อ่านจบ Part นี้คุณจะ**

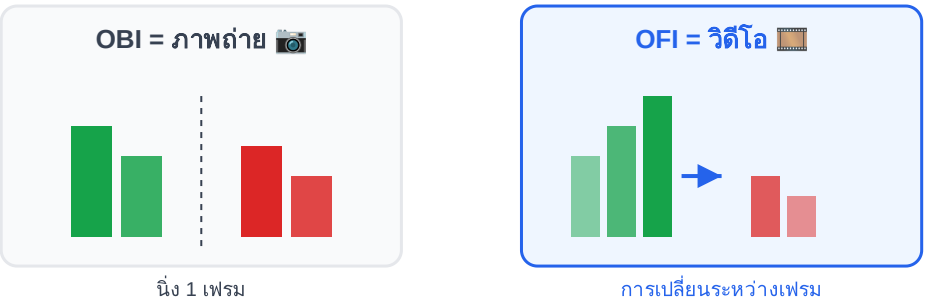
- แยก **OBI (state/ภาพนิ่ง)** ออกจาก **OFI (flow/การเปลี่ยน)** ได้ชัด (มีตารางเทียบ)
- คำนวณ OFI แบบ event-based ที่ best level ครบ 3 กรณี (ราคาขึ้น/เท่า/ลง)
- เข้าใจว่า **Kyle's λ = ความชันของเส้น regression** ΔP บน OFI — แรงดันราคาต่อหนึ่งหน่วย flow
- รู้จัก **multi-level OFI** (Cont 2023, รวมหลายชั้นด้วย PCA) และทำไมต้อง stationarize ก่อน regress

OBI = ภาพถ่าย / OFI = วิดีโอ

Part 5 สอน OBI ซึ่งเป็น "**ภาพถ่ายนิ่ง**" ของกระดาน ณ วินาทีหนึ่ง — บอกว่าตอนนี้ฝั่งไหนหนัก แต่ภาพนิ่งบอกไม่ได้ว่ามันกำลัง *เปลี่ยนไปทางไหน*

OFI (Order Flow Imbalance) คือ "**วิดีโอการเปลี่ยน**" — มันวัด *การเปลี่ยนแปลงสุทธิ* ของ depth ระหว่างสอง snapshot โดยรวมทั้ง 3 เหตุการณ์: limit เพิ่มเข้า, cancel ถอนออก, market กินออก ต่างจาก Δ/CVD (Part 6) ที่จับเฉพาะ market — OFI จับ **ทุกแรงที่ขยับกระดาน**

มิติ	OBI (Part 5)	OFI (Part 7)
วัดอะไร	**state** — ปริมาณที่ตั้งรอ ณ จุดเวลาหนึ่ง	**flow** — การเปลี่ยนสุทธิระหว่างสองช่วงเวลา
คำอุปมา	ภาพถ่ายนิ่ง	วิดีโอการเปลี่ยน
รวมเหตุการณ์	แค่ snapshot (ไม่รู้สาเหตุ)	limit + cancel + market รวมกัน
ความสัมพันธ์กับราคา	correlation (เดาทิศ)	**regression เชิงสาเหตุ** ΔP=λ·OFI
จับ "การเปลี่ยน" ไหม	ไม่ — เป็นภาพนิ่ง	ใช่ — เป็นแกนกลาง



ภาพ 7.1 · OBI (state) เทียบ OFI (flow)

นิยาม OFI ที่ best level — 3 กรณี

OFI event-based (Cont-Kukanov-Stoikov) วัดการเปลี่ยน depth ที่ best bid/ask ระหว่าง snapshot ก่อน (n−1) กับตอนนี้ (n) โดยขึ้นกับว่า **ราคา best ขยับหรือไม่**:

```
ฝั่ง bid (e^b_n) — เทียบ P^b กับ P^b ก่อนหน้า:
  ถ้า P^b_n > P^b_{n−1} → e^b = V^b_n (ราคา bid ขึ้น = แรงซื้อเข้า)
  ถ้า P^b_n = P^b_{n−1} → e^b = V^b_n − V^b_{n−1} (ราคาเท่า = ดูปริมาณ
เปลี่ยน)
  ถ้า P^b_n < P^b_{n−1} → e^b = −V^b_{n−1} (ราคา bid ลง = แรงขายกิน)

ฝั่ง ask (e^a_n) — สมมาตรกลับด้าน:
  ถ้า P^a_n < P^a_{n−1} → e^a = V^a_n
  ถ้า P^a_n = P^a_{n−1} → e^a = V^a_n − V^a_{n−1}
  ถ้า P^a_n > P^a_{n−1} → e^a = −V^a_{n−1}

OFI_n = e^b_n − e^a_n (บวก = แรงดันขึ้น, ลบ = แรงดันลง)
```

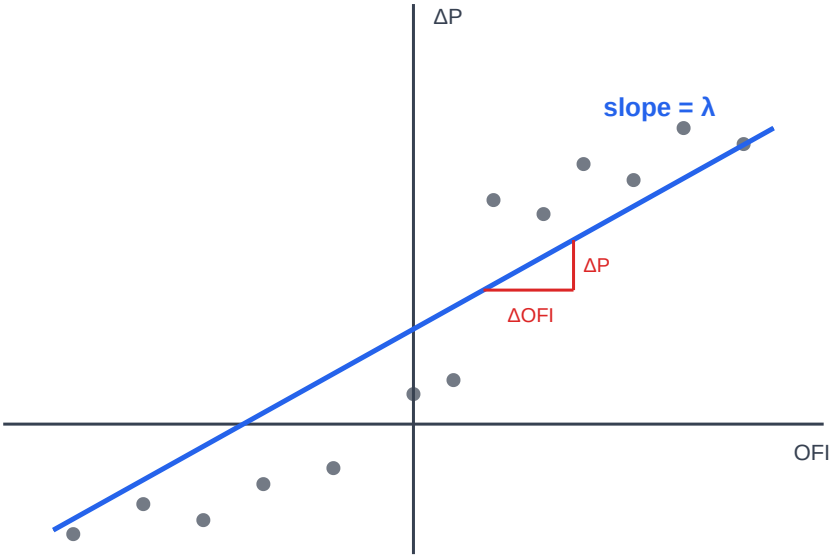
Kyle's λ — แรงดันราคาต่อหนึ่งหน่วย flow

คำถามสำคัญ: OFI หนึ่งหน่วยดันราคาเท่าไร? ตอบด้วย **regression เชิงเส้น** ของการเปลี่ยน mid (ΔP) บน OFI ความชัน (slope) ที่ได้คือ **Kyle's λ** — "ค่าความลึก/แพงของตลาด" (price impact coefficient)

```
ΔP = λ · OFI + ε ← สมการ price impact (linear)

λ = ความชันเส้น regression (OLS slope):
  λ = Cov(ΔP, OFI) / Var(OFI) ← สูตร OLS ที่เล่ม math สอน
```

```
λ สูง = ตลาดบาง (flow นิดเดียวดันราคาแรง)
λ ต่ำ = ตลาดลึก (flow มากแต่ราคาขยับน้อย — liquid)
```



แต่ละจุด = หนึ่งช่วงเวลา · เส้นผ่านจุดกำเนิด (two-sided) · ความชัน = price impact

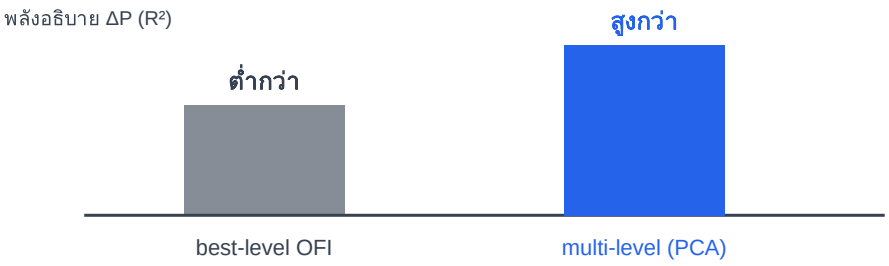
ภาพ 7.2 · (ภาพสำคัญ) scatter ΔP บน OFI — เส้น regression ที่ความชัน = Kyle's λ

Multi-level OFI + การ stationarize

OFI ที่ best level เดียวอาจไม่พอ **Cont, Cucuringu & Zhang (2023)** รวม OFI หลายชั้น (multi-level) ซึ่งมักมี correlation สูงระหว่างกัน จึงใช้ **PCA** ยุบเหลือองค์ประกอบหลัก (integrated/PC1 OFI) ก่อนใส่ regression — เพิ่มพลังอธิบายได้

```
multi-level: OFI^(k) สำหรับชั้น k = 1..L → integrated OFI =
PCA(OFI^(1)..OFI^(L)) เอา PC1

stationarize ก่อน regress (กันผลลวง/spurious):
  OFI_norm = OFI / DepthBar (หารด้วย depth เฉลี่ยช่วงนั้น → สเกลคงที่)
```



ภาพ 7.3 · multi-level OFI มัก explain ΔP ได้ดีกว่า best-level (เชิงสัมพัทธ์ — ค่าจริงขึ้นกับตลาด)

📑 งานวิจัยรองรับ — และข้อควรระวังเรื่องตัวเลข

Cont, Kukanov & Stoikov (2014) วาง OFI event-based และความสัมพันธ์เชิงเส้น ΔP=λ·OFI · **Cont, Cucuringu & Zhang (2023)** ขยายเป็น multi-level + cross-impact ด้วย PCA ตัวเลขที่มักอ้างว่า OFI อธิบาย ΔP ได้ **R^2~65%** (สูงกว่า OBI/trade imbalance) นั้น **ขึ้นกับสินทรัพย์ / timeframe / จำนวนชั้น / วิธี stationarize** อย่ายึดเป็นค่าตายตัว — ให้ผู้อ่าน calibrate กับ market ตัวเองเสมอ

⚠️ กับดักมือใหม่

- **regress OFI ดิบโดยไม่ stationarize** → ได้ R^2 หลอก (spurious) เพราะ depth สเกลเปลี่ยนตามเวลา ต้อง normalize ด้วย DepthBar ก่อน
- สับสน OBI กับ OFI — **OBI=ภาพนิ่ง, OFI=การเปลี่ยน** ใส่ OBI ลงสมการ price impact ผิดที่
- เชื่อว่า λ คงที่ — λ เปลี่ยนตามสภาพ liquidity (เวลาข่าว/ตลาดบางก็พุ่ง)

🔗 เชื่อมโยงเล่มอื่น (เน้น math)

Part นี้คือจุดที่เล่ม **math** เชื่อมเต็มตัว: **Kyle's λ = OLS slope** = Cov(ΔP,OFI)/Var(OFI) ตรงกับบทถดถอยเชิงเส้น · **R^2** = สัดส่วนความแปรปรวนที่อธิบายได้ · **multi-level OFI = PCA** ยุบตัวแปร correlated ตรงกับบท PCA ทั้งหมดนี้นำไปต่อ **Part 13 (stat-arb ด้วย log-OFI)** และเป็นรากของ **Part 8 (VPIN)** ที่เปลี่ยนจาก "แรงดันราคา" เป็น "ความเป็นพิษของ flow"

✍️ แบบฝึก (ทำได้ด้วยข้อมูลคริปโตฟรี)

ดึง Binance `@depth` ของ BTC/USDT มาทำ snapshot ต่อเนื่อง คำนวณ OFI ที่ best level ต่อช่วง 1 วินาที และ ΔP (การเปลี่ยน mid) แล้ว **รัน regression สองชุด**: (1) ΔP ~ OFI, (2) ΔP ~ OBI เทียบ **R^2** ของทั้งสอง — คุณจะเห็นด้วยมือว่าทำไม "การเปลี่ยน (OFI)" อธิบายราคาได้ดีกว่า "ภาพนิ่ง (OBI)" ลองเพิ่มการ stationarize (หาร DepthBar) แล้วดูว่า R^2 เปลี่ยนไหม

ต่อ Part 8: VPIN — เมื่อ flow ไม่ใช่แค่ดันราคา แต่ "เป็นพิษ" จนตลาดพัง →