Deutsche Bank

Internalisation by electronic FX Spot Dealers

Jim Gatheral's 60th Birthday Conference ♫♩♩♪

NYU Courant, New York, October 14, 2017

Roel Oomen
(joint work with Max Butz)

Motivation

- FX is traded OTC and makes up the world's largest financial market
  ... but often poorly understood ☹
- Traders obtain liquidity on a bi-lateral and bespoke basis from dealers
- The product is often standardised (e.g. spot) but the OTC nature lies in the delivery of liquidity
- There are a number of interesting topics here that have received limited academic interest thus far, e.g.

  1. traders aggregate dealer liquidity: how to do this?
  2. regulators are interested in "last look" : what is it (needed for)?
  3. dealers act as principal and risk manage in different ways : how and why?

Deutsche Bank

Risk management approach of principal dealers

Externalisation

Virtu Financial Inc (2014, p2) : "*Our strategies are also designed to lock in returns through precise and nearly instantaneous hedging, as we seek to eliminate the price risk in any positions held.*"

Internalisation

Bank of England, H.M. Treasury, and Financial Conduct Authority (2014, p. 59): "*Market participants have indicated that some dealers with large enough market share can now internalise up to 90% of their client orders in major currency pairs*"

Question to a specialist audience of FX traders at major industry event

Question *how long does it take a tier-1 LP to internalise a EURUSD ticket over the most active period of the day?*

(a) seconds

(b) tens of seconds

(c) minutes

(d) tens of minutes

Fast markets . . .

. . . by "fast markets" I mean those markets where these trends have gone furthest: most obviously major equity, foreign exchange (FX) and futures markets. In these markets, there is less need for intermediaries to warehouse risk, due to the inherent liquidity characteristics that attract a wide range of participants, making it easier to find a near-instant match between buyers and sellers.

Chris Salmon, "Keeping up with fast markets," Speech at 13th Annual Central Bank Conference on the Microstructure of Financial Markets, London 6 October 2017, `www.bankofengland.co.uk/speeches`

Contribution

Why is it important to understand internalisation?

- internalisation it is the process with which the majority (about 66%) of FX liquidity is “generated”
- there is a polarisation amongst dealers into internalisers and externalisers - how should traders evaluate this?
- the speed of internalisation (or externalisation) impacts transaction costs and (should) influence a trader’s execution strategy
- there is virtually no data available on it (the BIS 2016 triennial survey now includes a question on it)

Contribution of this paper is to provide a framework to analyse and understand the process

The model setup

- Let X_t, $t \geq 0$ denote the position of a liquidity provider (LP) at time$-t$, defined as the accumulation of completed buy- and sell-transactions of unit size:

$$X_t = N_t^+ - N_t^-$$

  where N_t^+ and N_t^- count the number of buy and sell transactions up to time$-t$

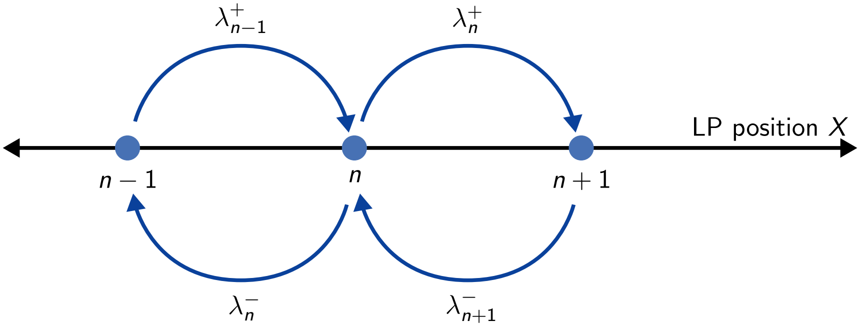
- X follows a compound Poisson process with position dependent arrival rates of buys and sells, i.e. when $X_t = n$, then $E(dN_t^+) = \lambda_n^+ dt$ and $E(dN_t^-) = \lambda_n^- dt$

- The LP can control the arrival rate of buys and sells by skewing its prices and does so in a manner that encourages risk reducing trades once the position exceeds some specified threshold $n^* \geq 0$, i.e.

$$\lambda_n^+ < \lambda_n^- \quad \text{when} \quad n > n^* \qquad \text{and} \qquad \lambda_n^+ > \lambda_n^- \quad \text{when} \quad n < -n^*.$$

Deutsche Bank

The model setup



The global balance equation of a Markov chain ($\lambda^+_n \phi_n = \lambda^-_{n+1}\phi_{n+1}$) gives the position distribution ϕ_n

$$\phi_n = \phi_0 \prod_{k=1}^{n} \frac{\lambda^+_{k-1}}{\lambda^-_k} \quad \text{for} \quad n > 0 \qquad \text{and} \qquad \phi_n = \phi_0 \prod_{k=1}^{n} \frac{\lambda^-_{1-k}}{\lambda^+_{-k}} \quad \text{for} \quad n < 0.$$

Example - binary position skew

- Let $\lambda_n^+ + \lambda_n^- = 2\lambda_0$ with

$$\lambda_n^+ = \begin{cases} \lambda_0 & |n| \leq R \\ \lambda_0(1-\alpha) & n > R \\ \lambda_0(1+\alpha) & n < -R \end{cases}$$

  for some fixed threshold $R \geq 0$ and $0 < \alpha \leq 1$.

- The distribution of the LP's position is stationary and given as:

$$\phi_n = \phi_0 \quad \text{for} \quad |n| \leq R \qquad \text{and} \qquad \phi_n = \frac{\phi_0}{1+\alpha}\left(\frac{1-\alpha}{1+\alpha}\right)^{|n|-1-R} \qquad \text{otherwise}$$

  where $\phi_0 = (2R + 1 + \alpha^{-1})^{-1}$.

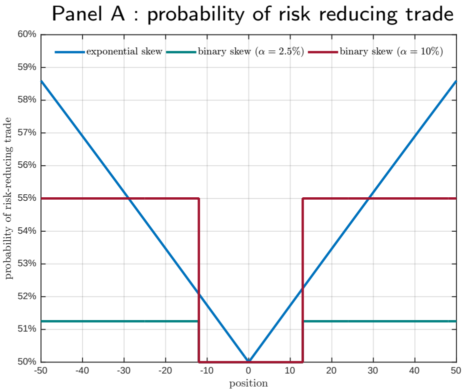
Deutsche Bank

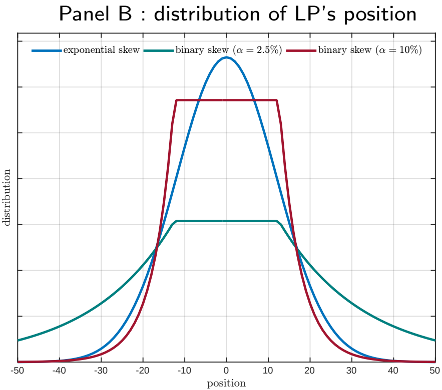
Example - exponential position skew

- Let $\lambda_n^+ = \lambda_{-n}^- = \lambda_0 e^{-\frac{1}{2} n / R^2}$ for $R > 0$.
- The distribution of the LP's position is stationary and given as:

$$\phi_n = \frac{1}{\sqrt{2\pi R^2}} e^{-\frac{1}{2} n^2 / R^2}.$$

Position skewing and distribution



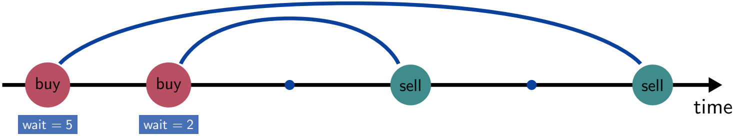


Internalisation horizon – definition

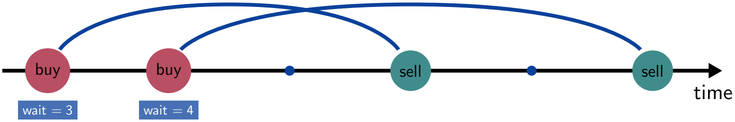
We define the internalisation horizon as the length of time a given trade forms part of the LP's risk position before it is fully offset by another trade in the opposite direction.

Internalisation horizon – illustration

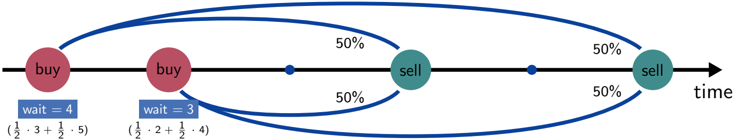
Last-in-first-out allocation



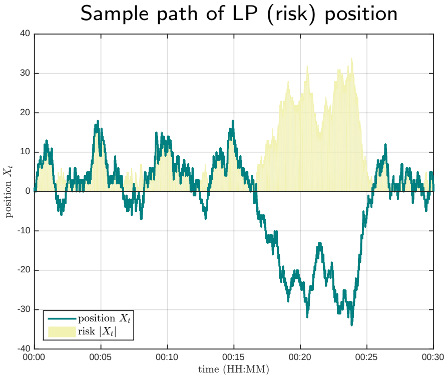
First-in-first-out allocation

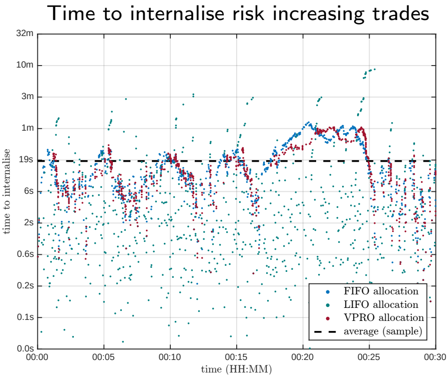


Pro-rated-by-volume allocation



Internalisation horizon – illustration





Distribution depends on allocation methodology → intractable?

Queuing theory

Little's Law

The average number of customers in a queuing system, denoted L, equals the average arrival rate of customers to the system, λ, multiplied by the average waiting time of a customer in the system, W, or $L = \lambda W$ (Little, 1961)

Internalisation as a queuing theory problem

- Reformulate buy/sell trades $\rightarrow$ risk increasing/decreasing trades
- Reformulate the dealer's long/short position $\rightarrow$ (absolute) risk position
- The time it takes for a risk increasing trade (the "customer") to be internalised by a risk decreasing trade ("serviced") once it has entered the dealer's risk position (the "queue") is what defines the internalisation horizon ("queuing time").
- Dealer uses position skewing to encourage risk-reducing flow ("deploying additional staff when restaurant is busy") to control the build-up of risk

Internalisation as a queuing theory problem

- Arrival rate of risk increasing trades ("customers"):

$$\lambda = \sum_{n \geq 0} \phi_n \lambda_n^+ + \sum_{n \leq 0} \phi_n \lambda_n^- .$$

- Dealer's risk position ("length of the queue")

$$L = \sum_{n \in \mathbb{N}} \phi_n |n|$$

Deutsche Bank

Internalisation horizon

With a binary position skew as specified in Example 1, the LP's average internalisation horizon of a trade is:

$$W = \frac{R}{\lambda_0}\frac{c_R}{4} \tag{1}$$

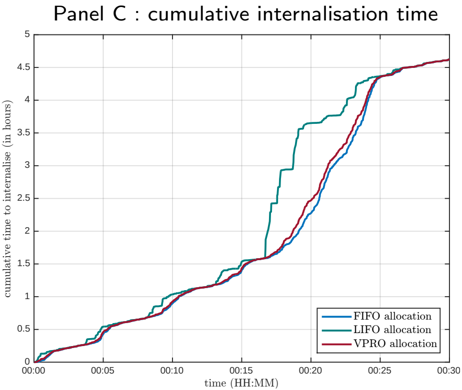
where $c_R = \frac{2\alpha^2 R + 2\alpha(1+\alpha) + (1+\alpha)/R}{2\alpha^2 R + \alpha(1+\alpha)} \to_R 1$.

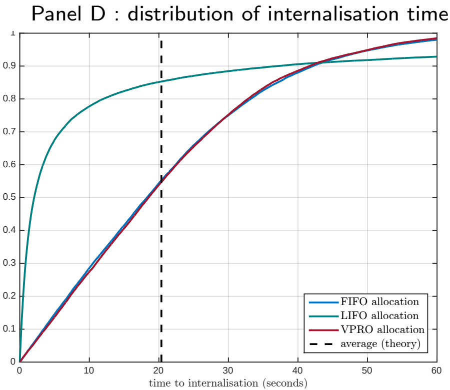
With an exponential position skew as specified in Example 2, the LP's average internalisation horizon of a trade is:

$$W \approx \frac{R}{\lambda_0}\frac{c_R}{\sqrt{2\pi}} \tag{2}$$

where $c_R = 2(1 - \Phi(\frac{1}{2}R^{-1}))e^{1/(8R^2)} + 1/\sqrt{2\pi R^2} \to_R 1$.

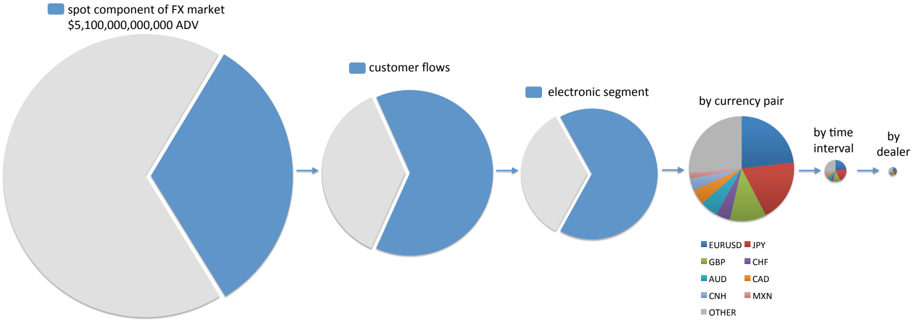
Internalisation horizon





Deutsche Bank

The size of the FX market



- FX is the world's largest financial market, ADV of $5,067bn (BIS, 2016) but ...
- $1,652bn of this is spot
- $1,047bn of this is customer flow
- $692bn of this is executed electronically
- further break-down by currency, time-period, and then individual dealer

Illustration of FX market flow rates in e-spot

- The EURUSD accounts for 23.1% of trading volumes which translates into

$$23.1\% \times \$692\text{bn/day} = \$159,852\text{mn}/1440\text{min} = \$111\text{mn/min}.$$

- Individual tier-1 dealer captures a fraction of this. Euromoney (2016) estimates that top 10 dealers hold 66% marketshare ...

$$\frac{1}{2} \times \$111\text{mn/min} \times 6.6\% = 3.66\text{mn/min} \text{ of each type.}$$

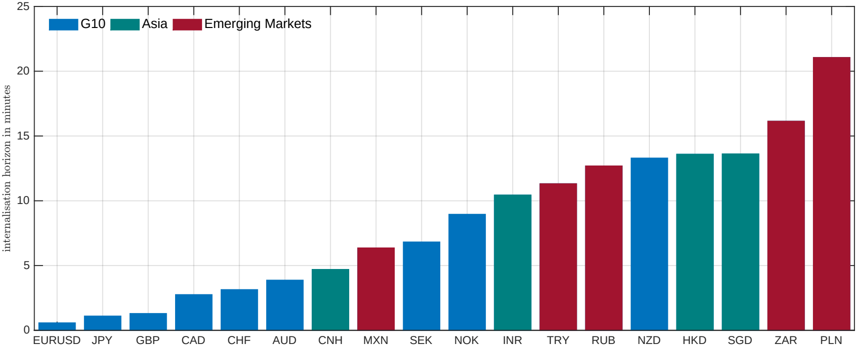
Illustration of internalisation horizons in practice

- Assume dealer adopts exponential price skewing and sets her risk limit so that the absolute risk position is within a \$25mn corridor 95% of the time, i.e. $R = \$25\text{mn}/\Phi^{-1}(97.5\%) = \$12.75\text{mn}$.
- This yields a top-down measurement of the average internalisation horizon of a tier-1 dealer in EURUSD as:

$$\frac{\$25\text{mn}/1.96}{\frac{1}{2}\times\$111\text{mn/min}\times 6.6\%}\times\frac{1.00}{\sqrt{2\pi}} = \frac{\$12.75\text{mn}}{3.66\$\text{mn/min}}\times\frac{1.00}{2.51} = 1.39 \text{ minutes.}$$

Deutsche Bank

Illustration of internalisation horizons in practice



Deutsche Bank

Illustration of internalisation horizons in practice

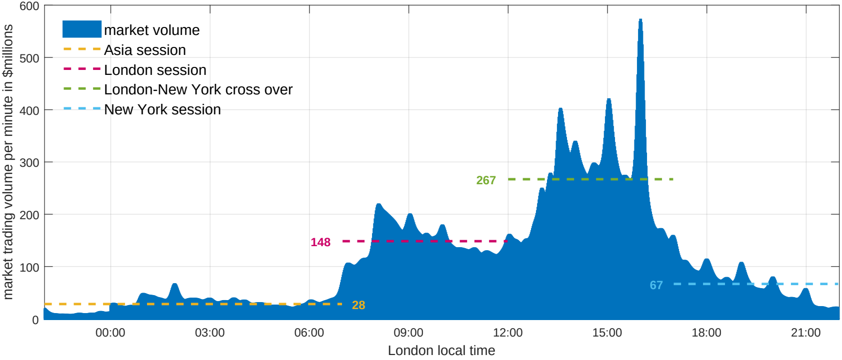
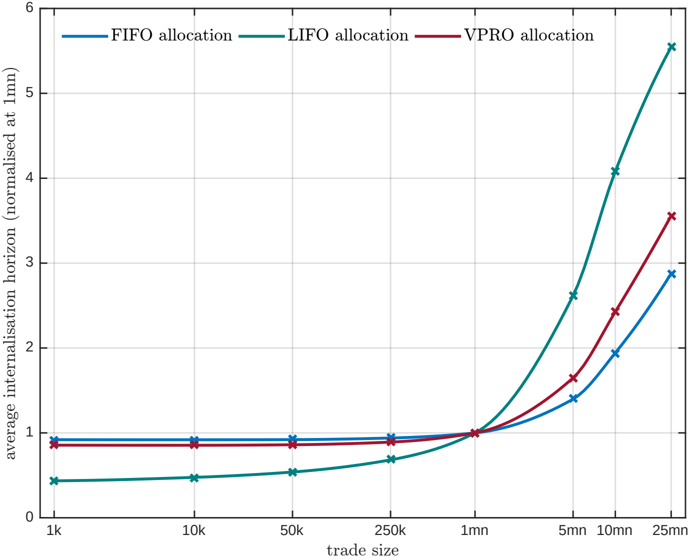


Illustration of internalisation horizons in practice

currency	market share	per minute trade flow rate (in $mn) avg	by trading session APAC	LON	NYLON	NY	average internalisation horizon (in mins) min	avg	by trading session APAC	LON	NYLON	NY
Panel A: G10 currencies (73.3% market share)												
EURUSD	23.1%	111	28	148	267	67	1	1	5	1	1	2
JPY	19.3%	93	86	103	140	47	1	2	2	1	1	3
GBP	11.2%	54	13	84	119	32	1	3	12	2	1	5
CHF	4.4%	21	4	31	49	14	3	7	40	5	3	11
AUD	5.5%	26	23	32	40	13	4	6	7	5	4	12
CAD	4.6%	22	6	17	56	23	3	7	27	9	3	7
NZD	1.5%	7	6	9	12	4	13	21	26	18	13	40
SEK	2.0%	10	1	19	23	3	7	16	60^+	8	7	48
NOK	1.5%	7	1	14	17	3	9	21	60^+	11	9	60^+
Panel B: Asian currencies (8.1% market share)												
CNH	3.8%	18	20	33	14	5	5	8	8	5	11	34
SGD	1.6%	8	7	11	10	3	14	20	21	14	16	59
HKD	1.5%	7	6	11	9	3	14	21	26	14	17	46
INR	1.1%	5	5	15	1	0	10	29	29	10	60^+	60^+
Panel C: Emerging markets currencies (5.6% market share)												
MXN	1.8%	9	1	5	24	11	6	18	60^+	34	6	14
TRY	1.3%	6	1	13	14	2	11	25	60^+	12	11	60^+
RUB	1.1%	5	0	8	12	5	13	29	60^+	18	13	34
ZAR	0.8%	4	0	6	10	2	16	40	60^+	24	16	60^+
PLN	0.6%	3	0	6	7	1	21	53	60^+	28	21	60^+

Internalisation horizon for non-unit trade sizes . . .



Deutsche Bank

What are the costs of position skewing / internalisation?

Let's start with an informal 3-step approach to build intuition ...

1. Assume linear approximation to exponential skewing, i.e. when $X_t = n$

$$\text{Probability of risk reducing trade } P_n \approx \frac{1}{2} + \frac{|n|}{4R^2}$$

2. Assume simplest Oomen (2017) model with two LPs competing for uninformed trader's flow

$$\text{Required price skew } \theta_n \propto -\frac{n}{R^2}$$

3. Cost of skewing is then the sum of ...
   3.1 cost when skewing "in"

$$\mathbb{C}_S^- \propto \sum_n |\theta_n| P_n \phi_n \approx \frac{\kappa_-}{2R} + \frac{\sqrt{2\pi}\kappa_-}{8R^2}$$

   3.2 minus revenues when skewing "out"

$$\mathbb{C}_S^+ \propto \sum_n |\theta_n|(1 - P_n)\phi_n \approx \frac{\kappa_-}{2R} - \frac{\sqrt{2\pi}\kappa_-}{8R^2}.$$

Deutsche Bank

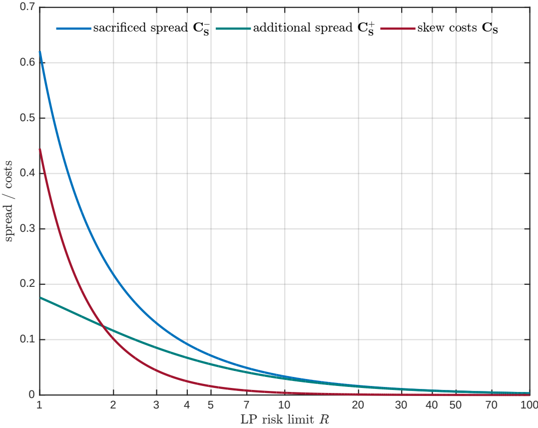
What are the costs of position skewing / internalisation?

Heuristic approach shows:

(a) costs increase with reduction in LP's willingness to hold risk

(b) skew components $O(R^{-1})$

(c) net skew costs $O(R^{-2})$



What are the costs of position skewing / internalisation?

There are a number of deficiencies with this approach

- adverse selection (set of deals won) is affected by the act of skewing
- trader's actions may depend on price (skew) shown
- dealer may skew prices "asymmetrically"

What are the costs of position skewing / internalisation?

Starting point model by Oomen (2017) where LPs compete on best price for a trader's flow

- True price process is unobserved

$$p_t^* = p_{t-1}^* + \varepsilon_t$$

- LPs make independent assessment of true price:

$$p_t^{(i)} = p_t^* + d_t^{(i)} \quad \text{where} \quad d_t^{(i)} \sim \text{ i.i.d. } \mathcal{N}(0, \omega^2)$$

  for $i \in \{1, 2\}$ and $\text{corr}(d_t^{(1)}, d_t^{(2)}) = \rho_d$

- LPs charge a nominal spread around this price:

$$b_t^{(i)} = p_t^{(i)} - \frac{1}{2}s \quad \text{and} \quad a_t^{(i)} = p_t^{(i)} + \frac{1}{2}s$$

Deutsche Bank

Two extensions

1. Explicit decomposition of "true price deviation" into $d_t = m_t + \theta_t$ where
   - m_t is the measurement error, random i.i.d. $(0,\kappa^2)$
   - θ_t is the LP's position skew, known and deterministic to the LP, but random to others.
2. The trader is "informed" to the extent that her order placement depends on the LPs prices as follows:
   - trader gets a signal on true price that is i.i.d. (p_t^*, ω_T^2)
   - they sells to LP–i when:

     LP–i shows best bid $b_{t_j}^{(i)} > b_{t_j}^{(\neq i)}$

     Selling is more attractive than buying $p_{t_j}^{(0)} - b_{t_j}^{(i)} < \min(a_{t_j}^{(i)} - p_{t_j}^{(0)}, a_{t_j}^{(\neq i)} - p_{t_j}^{(0)})$
   - trader fully informed when $\omega_T = 0$, noise trader when $\omega_T = \infty$.

Spread metrics

Assume two identical LPs and absence of position skewing (i.e. $\theta = 0$).

The expected observed spread is defined as $S = E(\min_i a_t^{(i)} - \max_i b_t^{(i)})$ and equal to:

$$S = s - \kappa_- \sqrt{2/\pi}, \tag{3}$$

where $\kappa_\pm^2 = 2\kappa^2(1 \pm \rho_m)$.

The expected effective spread is defined as $\mathbb{S} = 2E(|x_{t_j} - p_{t_j}^*|)$ and equal to:

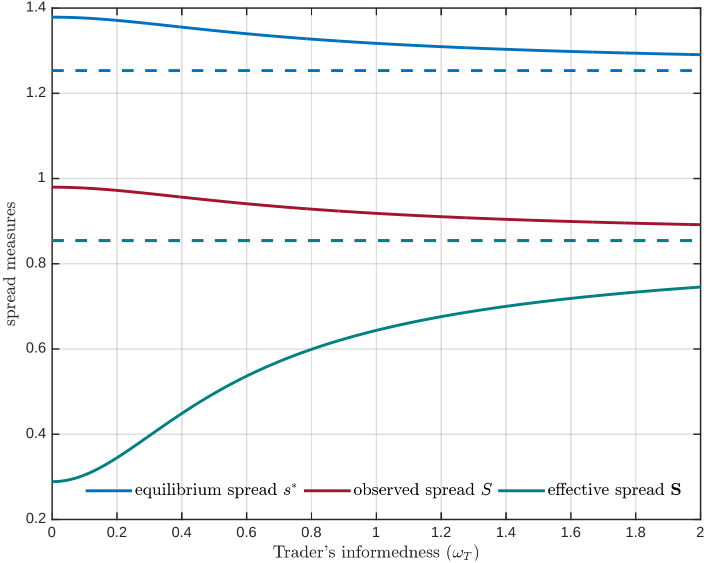
$$\mathbb{S} = S - \xi^{-1}\kappa_+^2 \sqrt{2/\pi}, \tag{4}$$

where $\xi^2 = \kappa_+^2 + 4\omega_T^2$.

The equilibrium spread at which neither LP can profit by unilaterally making a change to their own spread is equal to:

$$s^* = \frac{4\kappa^2 + \pi\xi\kappa_-}{\xi + \kappa_-}\sqrt{2/\pi}. \tag{5}$$

Spread metrics



Deutsche Bank

LP's position skewing rule

Symmetric price skewing $\theta_t^{(i)} = -\gamma X_t^{(i)}$

Asymmetric generalisation

$$\begin{aligned} b_t^{(i)} &= m_t^{(i)} - \frac{1}{2}s - \gamma_{\mathrm{in}} \min(X_t^{(i)}, 0) - \gamma_{\mathrm{out}} \max(X_t^{(i)}, 0), \\ a_t^{(i)} &= m_t^{(i)} + \frac{1}{2}s - \gamma_{\mathrm{in}} \max(X_t^{(i)}, 0) - \gamma_{\mathrm{out}} \min(X_t^{(i)}, 0), \end{aligned}$$

Cost of skewing

Assume LP–2 does not control position via skewing.

LP–1 compares revenues with ($\mathbb{R}_\gamma$) and without ($\mathbb{R}_0$) position skewing $\rightarrow$ difference is cost of skewing where ...

$$\begin{aligned} \mathbb{R}_c \quad = \quad & E(a_{t_j}^{(1)} - p_{t_j}^* \mid \gamma = c, \text{ trader buys from LP-1}) \cdot \Pr(\text{trader buys from LP-1} \mid \gamma = c) \\ & + E(p_{t_j}^* - b_{t_j}^{(1)} \mid \gamma = c, \text{ trader sells to LP-1}) \cdot \Pr(\text{trader sells to LP-1} \mid \gamma = c) \end{aligned}$$

Note that $\mathbb{R}_0 = \mathbb{S}$.

Internalisation metrics (first order)

Costs of position skewing for LP–1 is:

$$\mathbb{C}_\gamma = \vartheta \left(s^* - s\right) \frac{\gamma_{\mathrm{in}} - \gamma_{\mathrm{out}}}{4\pi} R + \mathcal{O}_C \left(R^{-2}\right).$$

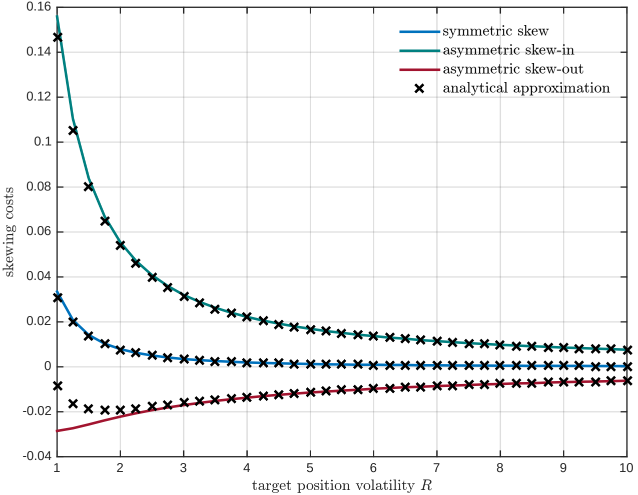
The market share of LP–1 is:

$$\mathbb{M}_\gamma = \frac{1}{2} + \vartheta \frac{\gamma_{\mathrm{in}} - \gamma_{\mathrm{out}}}{2\pi} R + \mathcal{O}_M \left(R^{-2}\right).$$

The average spread shown by LP–1, i.e. $E(a_t^{(1)} - b_t^{(1)})$*, is:*

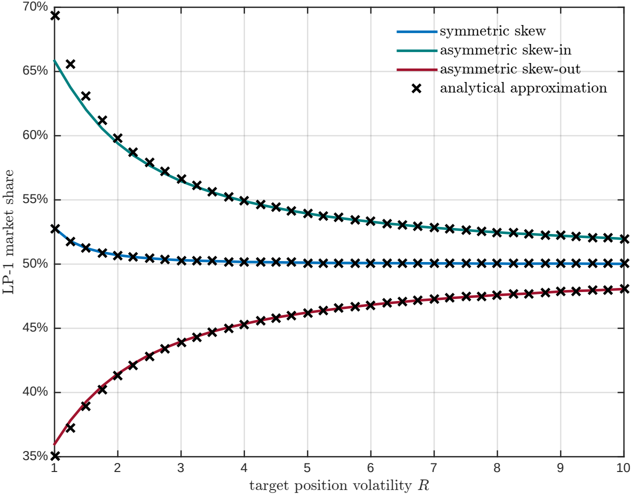
$$\overline{s}_\gamma = s - \frac{\gamma_{\mathrm{in}} - \gamma_{\mathrm{out}}}{\sqrt{\pi/2}} R + \mathcal{O}_{\overline{s}} \left(R^{-2}\right).$$

Internalisation metrics – illustration of skewing costs



Note: parameters set as $\omega_T = \frac{1}{2}\kappa$, and $s = \frac{1}{2}s^*$, i.e. a scenario with a highly informed trader and LPs that charge a tight nominal spread.

Internalisation metrics – illustration of market share



Note: parameters set as $\omega_T = \frac{1}{2}\kappa$, and $s = \frac{1}{2}s^*$, i.e. a scenario with a highly informed trader and LPs that charge a tight nominal spread.

Deutsche Bank

Cost of internalisation

Trader informedness	Nominal spread s = equilibrium spread s^* perfect ($\omega_T = 0$)	high ($\omega_T = \kappa$)	medium ($\omega_T = 2\kappa$)	none ($\omega_T = \infty$)
Panel A: spread metrics in absence of position skewing ($\times 100$)				
Nominal half-spread ($\frac{1}{2}s$)	68.9	67.4	65.9	62.7
Observed half-spread ($\frac{1}{2}S$)	49.0	47.4	45.9	42.7
Effective half-spread ($\frac{1}{2}\mathbb{S}$)	14.4	24.8	32.2	42.7
Panel B: cost of symmetric position skewing ($\times 100$)				
Low risk ($R = 1$)	2.0	3.1	4.1	6.1
Medium risk ($R = 5$)	0.1	0.1	0.2	0.3
High risk ($R = 10$)	0.0	0.0	0.0	0.1
Panel C: cost of asymmetric-in position skewing ($\times 100$)				
Low risk ($R = 1$)	9.1	12.1	15.2	30.0
Medium risk ($R = 5$)	0.3	0.4	0.5	0.6
High risk ($R = 10$)	0.1	0.1	0.1	0.1
Panel D: cost of asymmetric-out position skewing ($\times 100$)				
Low risk ($R = 1$)	1.5	2.6	3.5	5.0
Medium risk ($R = 5$)	0.1	0.2	0.3	0.4
High risk ($R = 10$)	0.0	0.1	0.1	0.1

Cost of internalisation

Trader informedness	Nominal spread $s = \frac{1}{2}$ equilibrium spread s^*			
	perfect ($\omega_T = 0$)	high ($\omega_T = \kappa$)	medium ($\omega_T = 2\kappa$)	none ($\omega_T = \infty$)
Panel A: spread metrics in absence of position skewing ($\times 100$)				
Nominal half-spread ($\frac{1}{2}s$)	34.5	33.7	32.9	31.3
Observed half-spread ($\frac{1}{2}S$)	14.5	13.7	13.0	11.4
Effective half-spread ($\frac{1}{2}\mathbb{S}$)	−20.0	−8.9	−0.7	11.4
Panel B: cost of symmetric position skewing ($\times 100$)				
Low risk ($R = 1$)	3.0	3.9	4.7	6.0
Medium risk ($R = 5$)	0.1	0.1	0.2	0.2
High risk ($R = 10$)	0.0	0.0	0.0	0.1
Panel C: cost of asymmetric-in position skewing ($\times 100$)				
Low risk ($R = 1$)	14.6	17.4	20.4	34.8
Medium risk ($R = 5$)	1.7	1.7	1.8	1.8
High risk ($R = 10$)	0.8	0.8	0.7	0.8
Panel D: cost of asymmetric-out position skewing ($\times 100$)				
Low risk ($R = 1$)	−3.3	−2.2	−1.2	0.4
Medium risk ($R = 5$)	−1.2	−1.1	−1.0	−0.8
High risk ($R = 10$)	−0.6	−0.6	−0.6	−0.5

Concluding remarks

- *In the world's largest financial market of FX, where the majority of spot transactions are conducted electronically at a pace approaching the speed of light, the popular perception is that dealers tend to hold risk positions for only a matter of seconds. We show this is not the case.*

  Liquidity is finite and takes time to "produce" . . .
  (Jim – of course – already knew that long time ago!)

- *Efficient execution requires the execution objectives of the trader and the hedging approach of the LPs to be aligned*
  - more LPs not always better
  - execution style of the trader impacts the liquidity they can access
  - polarisation between internalisers and externalisers overly simplistic
  - direct trade-off between urgency and costs

    (note: big difference between efficient execution and "best execution")

Deutsche Bank

Disclaimer

Deutsche Bank