Learning (not) to trade: Lindy's law in retail traders

Teodor Godina

Serge Kassibrakis

Semyon Malamud

Alberto Teguia

Jiahua (Java) Xu

November 11, 2020

Two quick surveys

Survey I

Imagine you are a stock trader. You have performed very well during the last year: 10% of return. Now you wonder whether to keep trading or exit.

Other things equal, you'd be more likely to exit, if you:

- A have been trading for 10 years
- B have been trading for 2 years

Survey II

Imagine you are a stock trader. You have performed very poorly during the last year: -10% of return. Now you wonder whether to keep trading or exit.

Other things equal, you'd be more likely to exit, if you:

- A have been trading for 10 years
- B have been trading for 2 years

Research motivation and background

Exit procrastination—Lindy's law?

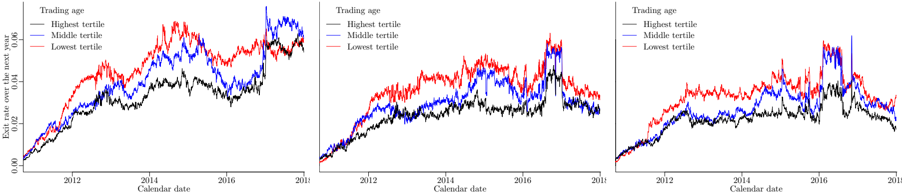


Figure 1: At any given return level (left to right: lowest to highest tertile), higher trading age is associated with lower exit rate

Purpose of the study

To model trading inertia, a.k.a Lindy's law, in retail traders.

Literature review

- [Seru et al., 2010]: two learning routes of investors—the first type improves over time by learning from their trading experience, while the second type realizes their ability is insufficient and ceases to trade.
- [Nicolosi et al., 2009]: investors benefit from trading experience and can learn to place more profitable trades.
- [Gervais and Odean, 2001]: traders develop assessments of their own trading ability through experience.
- [Schraeder, 2016]: young investors trade frequently but obtain lower returns.

Data and Methodology

Sample collection

- Source: one of the biggest stock trading platforms for retail traders in Switzerland
- $82,072$ traders' status (exit/remain) from 2001-05-04 to 2019-01-29
- $90,628$ traders' trading activity (buy/sell) from 2001-05-08 to 2018-12-31, with in total $6,766,393$ trader-action observations
- $64,037$ traders' daily account value from 2009-01-01 to 2018-12-28, with in total $107,332,300$ account-day observations

The model I

Assumptions

- Based on their trading skills, each trader falls in either of the two categories: (i) with trading skills, i.e. good type, denoted by G, or (ii) without trading skills, i.e. bad type, denoted by B.
- Traders do not know their types, but have a prior belief about their probability of being Type G equal to π_0.
- Type B traders never place successful trades with profits sufficient to cover the total trading cost.
- For type G traders, successful trades arrive at jump times of a time-inhomogeneous Poisson process, N_t, with an exogenously given intensity $f(t)$, namely, $\mathbb{P}[N_{t+dt} - N_t = 1] = f(t)dt$.
- Type G traders improve their trading skills with experience. Thus, $f(t)$ is a monotone increasing function of the trader's trade age, t.
- Traders learn about their types and continuously update their belief π_t from observing their trading history.

The model II

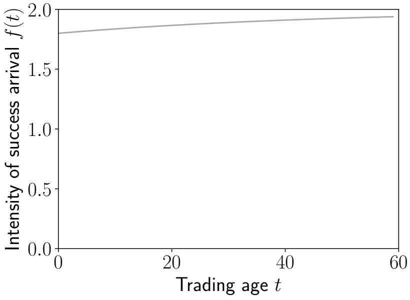
Assumptions

- Entry of new traders happens at a constant rate ρ_{entry}.
- Traders exit from trading for two reasons: (i) an exogenous "exit shock" that occurs at a Poisson intensity ρ_{exit}, or (ii) the trader's posterior probability of being Type G dropping below a threshold value, i.e. when the trader becomes sufficiently confident that he is Type B.
- Each trader has an idiosyncratic running fixed cost of trading φ

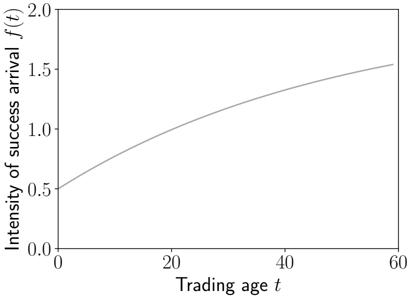
Time-inhomogeneous Poisson process I

The arrival of success becomes increasingly frequent for Type G.

$f(t) = -10e^{-0.02t} + 20$



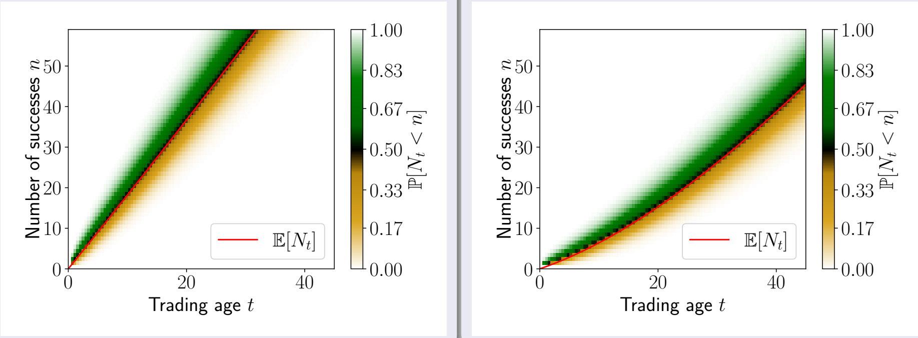
$f(t) = -1.5e^{-0.02t} + 2$



Time-inhomogeneous Poisson process II

Cum. prob. of no. of successes n achieved by age t (green-orange fan):
$\mathbb{P}[N_t < n] = \sum_{k=0}^{n-1} \frac{(F(t))^k}{e^{F(t)} k!}$
Expected no. of successes n by age t for Type G (red line):
$\mathbb{E}[N_t] = F(t) = \int_0^t f(t)dt$

$f(t) = -10e^{-0.02t} + 20$ | $f(t) = -1.5e^{-0.02t} + 2$



Time-inhomogeneous Poisson process III

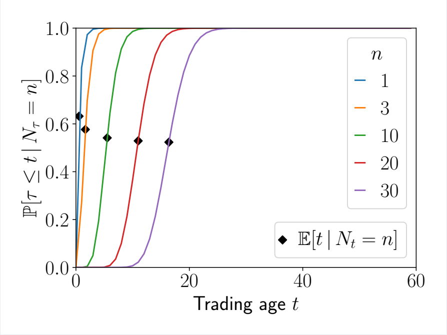
Probability of experiencing n successes by age t for Type G (color lines):

$\mathbb{P}[\tau \leq t \mid N_\tau = n] = 1 - \sum_{k=0}^{n-1} \frac{(F(t))^k}{e^{F(t)} k!}$

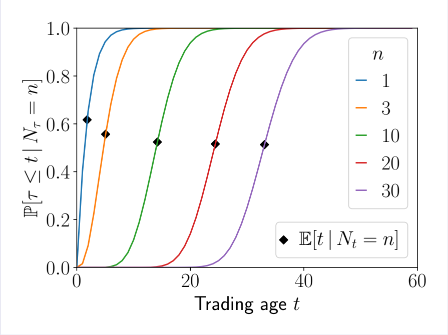
Expected time needed to achieve n successes (black diamonds):

$\mathbb{E}[t \mid N_t = n] = \int_0^\infty \frac{t f(t)(F(t))^{n-1}}{e^{F(t)}(n-1)!} dt$

$f(t) = -10e^{-0.02t} + 20$



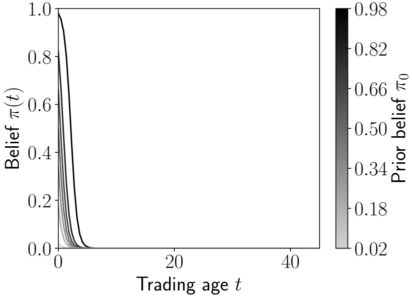
$f(t) = -1.5e^{-0.02t} + 2$



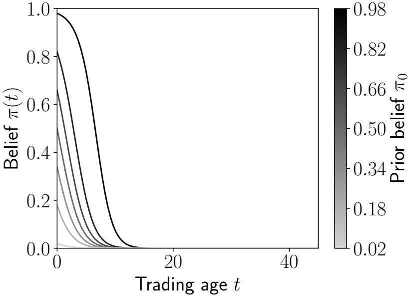
Time-inhomogeneous Poisson process IV

The belief of being Type G decreases before encountering the first success.
$\pi(t) = \frac{1}{(\frac{1}{\pi_0}-1)e^{F(t)}+1}, \;\; t < \tau_N$

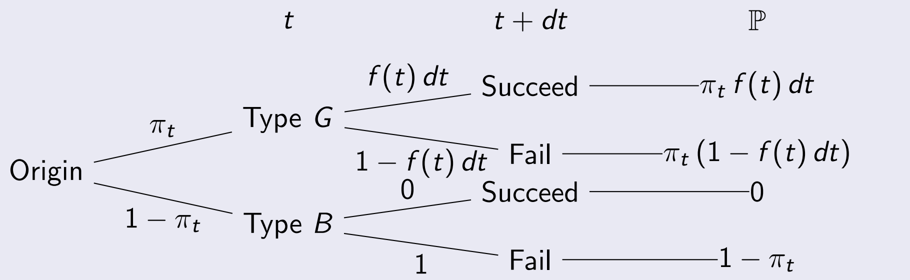
$f(t) = -10e^{-0.02t} + 20$



$f(t) = -1.5e^{-0.02t} + 2$



Evolution of belief I



$$d\pi_t = \pi_{t+dt} - \pi_t = \frac{\pi_t\big(1 - f(t)dt\big)}{\pi\big(1 - f(t)dt\big) + (1 - \pi_t)} - \pi_t = -f(t)\pi_t(1 - \pi_t)dt$$

for $t < \tau_N$, while $\pi_t = 1$ for $t \geq \tau_N$.

Evolution of belief II

Log transformation for mathematical convenience:

$$\ell_t \; = \; \ln \frac{\pi_t}{1 - \pi_t}$$

By direct calculation, ℓ_t follows the dynamics:

$$d\ell_t \; = \; -f(t)dt$$

for $t < \tau_N$ and $\ell_t = \infty$ for $t \geq \tau_N$.

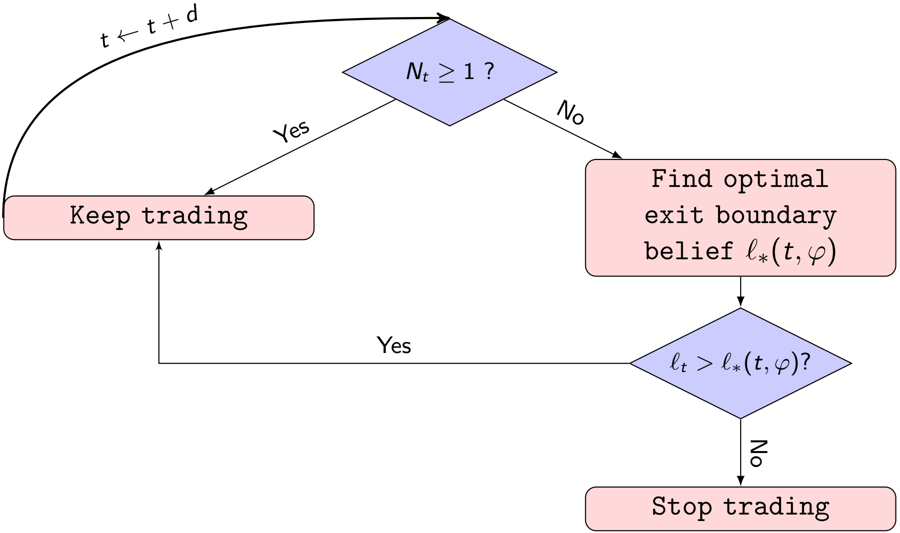


Figure 2: A trader's decision tree (absent exogenous shock)

Value functions

- Type G:
  $V_G(t,\varphi) \; = \; \int_t^{\infty} e^{-r(s-t)}\big(f(s) - (\varphi + 0.5c^{-1}\eta^2)\big)\,ds$
- Type unknown:
  $V(\ell,t,\varphi) \; = \; \max_{\tau_E} \mathbb{E}\Big[- \int_t^{\min\{\tau_E,\tau_N\}} e^{-rs}(\varphi + 0.5c^{-1}\eta^2)ds \; + \; e^{-r\tau_N} 1_{\tau_N<\tau_E}\big(1 \; + \; V_G(\tau_N,\varphi)\big)\Big]$

Theorems

- exit boundary $\ell_*(t,\varphi)$
  $\frac{e^{\ell_*(t,\varphi)}}{e^{\ell_*(t,\varphi)}+1} \; = \; \frac{\varphi}{f(t)(1+V_G(t,\varphi))}$

Variables and parameters I

Observable

- 1_{exit}: account status—closed/open
- R individual trading performance: one-year excess log return
- t trading age: time since account open
- $m(t)$ total trading mass by age: number of open accounts

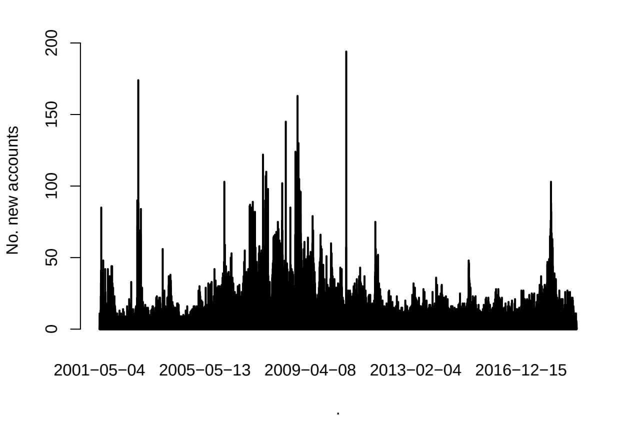
Unobservable

- Success
- Type (G or B) and its distribution
- ℓ belief of being Type G and its (conditional) distribution (given type)
- φ running cost and its distribution
- Conditional distribution (given type and belief) of R
- ρ_{exit} exit rate due to exogenous shock

Variables and parameters II

Other

- c cost factor per trading effort: arbitrarily fixed
- ρ_{entry} entry rate: assumed time invariant



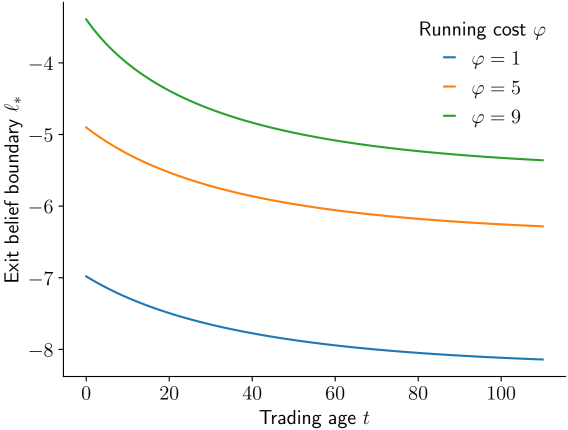
Numerical simulation I

- Success arrival intensity: $f(t) = \alpha\, e^{\beta} + \gamma$, where $\alpha = -10, \beta = -0.02, \gamma = 20$
- Discount rate: $r = 0.1$
- Constant entry rate: $\rho_{entry} = 1$
- Exit rate due to exogenous shock: $\rho_{exit} = 0.4$

Numerical simulation II

- Distribution of types at entry: Type G 10%, Type B 90%
- Distribution of φ at entry: uniform distribution on $[0, \varphi^*]$, where $\varphi^* = \alpha + \gamma - 0.5\,(r + \frac{\alpha\,\beta}{\alpha+\gamma})$
- Distribution of belief ℓ_0 at entry: inverse exponential distribution with $\lambda = 60$ for Type G and $\lambda = 50$ for Type B, on $[\ell(0, \varphi), +\infty)$
- Distribution of return based on true type and belief:
    - Being G believing G: $\mathbb{N}(2, 0.8^2)$
    - Being G believing B: $\mathbb{N}(1, 1^2) \leftarrow$ *underconfidence*
    - Being B believing G: $\mathbb{N}(-2, 1^2) \leftarrow$ *overconfidence*
    - Being B believing B: $\mathbb{N}(-1, 1.2^2)$

Belief lower boundary $\ell_*(t, \varphi)$



$\ell_*(t, \varphi)$ decreases with t

Exit likelihood

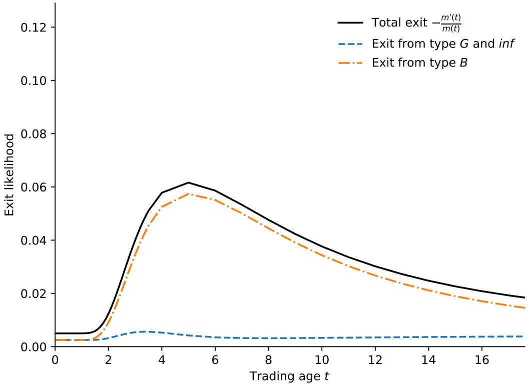


Figure 3: Exit likelihood by trading age t

Exit likelihood increases and then decreases

Performance by fixed cohort

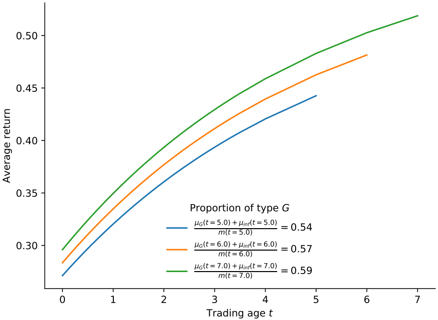


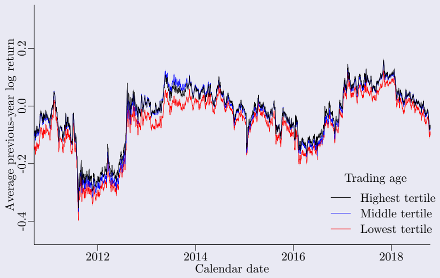
Figure 4: Average return for a specific cohort

Empirical results

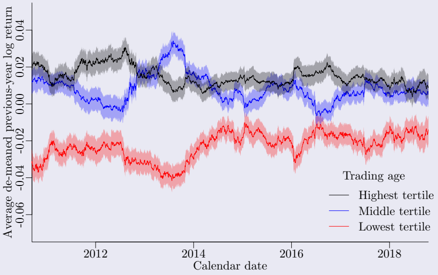
Evolution of clients' performance

Higher trading age associated with higher performance: more seasoned clients perform better.

Original



Demeaned



Learning effect: more seasoned clients perform better

$$R_{i,t} = \alpha_T + \beta^t\, t + \boldsymbol{X}_i \boldsymbol{\beta}^{\boldsymbol{X}} + \epsilon_{i,t} \qquad (1)$$

Table 1: Panel regressions on performance with respect to trading age and other control variables.

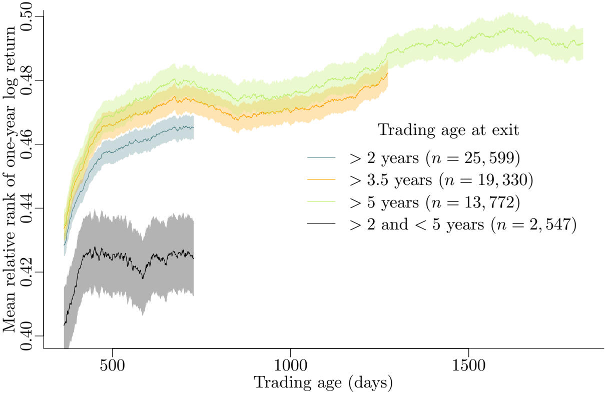
Regressand: return $R_{i,t}$	(1)	(2)	(3)	(4)	(5)
Trading age t (days)	$1.0 \times 10^{-5***}$	$1.1 \times 10^{-5***}$	$1.1 \times 10^{-5***}$	$1.0 \times 10^{-5***}$	$1.0 \times 10^{-5***}$
	(0.000)	(0.000)	(0.000)		(0.000)
Gender (Female = 1)		0.0325^{***}			0.0293^{***}
		(0.000)			(0.000)
Biological age at entry (years)			0.0008^{***}		0.0007^{***}
			(0.000)		(0.000)
Trades options (Yes = 1)				-0.0667^{***}	-0.0658^{***}
				(0.000)	(0.000)
Time effect (dates)	Yes	Yes	Yes	Yes	Yes
No. obs.	106,564,611	106,564,611	106,398,965	106,564,611	106,564,611
R^2	0.001	0.001	0.002	0.003	0.003

Significantly positive β^t

Trading age positively correlated with performance.

Learning effect: surviving clients learn faster

Figure 5: Evolution of performance rank as trading age increases for fixed cohorts



Learning effect: surviving clients learn faster

$$R_{i,t} = \alpha_T + \beta^t \, t + \boldsymbol{X}_i \beta^{\boldsymbol{X}} + \epsilon_{i,t}, \tag{2}$$

Table 2: Effect of trading age on performance for lifetime-classified cohorts at different stages of trading life.

Age range	1–1.5 years		1.5–2 years		2–3.5 years		3.5–5 years		
Cohort	$\overline{\alpha}$	β	$\overline{\alpha}$	β	$\overline{\alpha}$	β	$\overline{\alpha}$	β	n
2-year	39.8*** (0.000)	0.0129*** (0.000)	45.0*** (0.000)	0.0025*** (0.000)					25,599
3.5-year	40.4*** (0.000)	0.0138*** (0.000)	46.1*** (0.000)	0.0024*** (0.000)	46.6*** (0.000)	0.0009*** (0.000)			19,330
5-year	40.1*** (0.000)	0.0152*** (0.000)	46.1*** (0.000)	0.0035*** (0.000)	47.4*** (0.000)	0.0007*** (0.000)	49.0*** (0.000)	0.0001** (0.013)	13,772
2to5	39.6*** (0.000)	0.0060*** (0.000)	41.1*** (0.000)	0.0020*** (0.000)					2,547

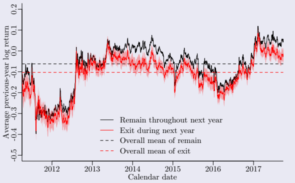
Higher β^t for longer surviving cohorts

Longer surviving clients learn faster.

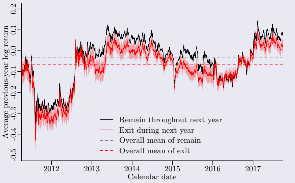
Performance: exiting vs remaining

Original

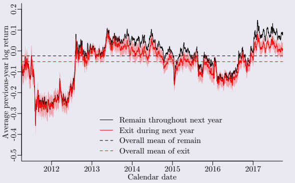
Lowest trading age tertile



Middle trading age tertile

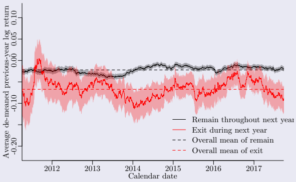


Highest trading age tertile

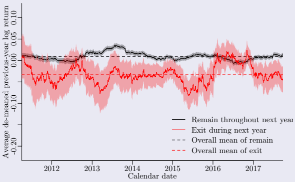


Demeaned

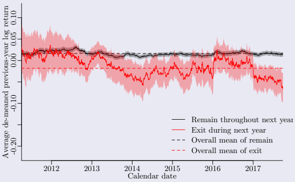
Lowest trading age tertile



Middle trading age tertile



Highest trading age tertile



Clients exit due to poor performance

$$\ln h_i(t \mid R_{i,t}, \boldsymbol{X}_i(t)) = \ln h_0(t) + \ln \alpha_i(t) + \beta^R R_{i,t} + \boldsymbol{X}_i(t)\beta^{\boldsymbol{X}} \qquad (3)$$

Table 3: Cox proportional hazard regressions with respect to trading performance and other control variables

Regressand: log hazard $\ln h_i(t)$	(1)	(2)
One-year log return $R_{i,t}$	-0.7696^{***}	-0.7833^{***}
	(0.000)	(0.000)
Gender (Female = 1)		0.2563^{***}
		(0.000)
Biological age at entry (years)		0.0019^{**}
		(0.022)
Trades Options (Yes = 1)		-0.0133
		(0.746)
Individual effect	Yes	Yes
Concordance	0.563	0.567
No. obs.	106,564,611	106,398,965

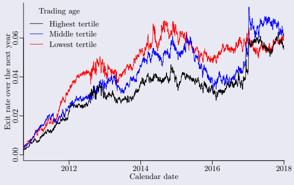
Significantly negative β^R

Poor performance increases exit likelihood.

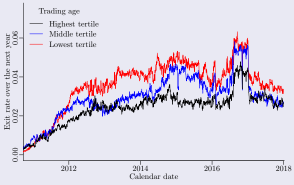
Trading age vs exit likelihood

Higher trading age associated with lower exit likelihood with performance level controlled for: exit procrastination (Lindy's Law)?

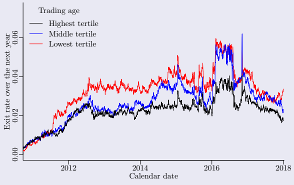
Lowest return tertile



Middle return tertile



Highest return tertile



Trading age vs exit likelihood

$$E_{i,t} = \alpha_T + \beta^t\, t + \beta^R R_{i,t} + \boldsymbol{X}_i \beta^{\boldsymbol{X}} + \epsilon_{i,t} \tag{4}$$

$E_{i,t}$: at time t client i exit = 1, remain = 0.

Table 4: Logistic regressions on exit with regard to trading age and other control variables

	Full sample		Subsample: age < 3 yr		Subsample: age ≥ 3 yr	
	(1)	(2)	(3)	(4)	(5)	(6)
Trading age (days) t	$-10.6 \times 10^{-5***}$	$-10.6 \times 10^{-5***}$	$1.0 \times 10^{-5***}$	$1.0 \times 10^{-5***}$	$-14.9 \times 10^{-5***}$	$-14.9 \times 10^{-5***}$
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
One-year log return $R_{i,t}$	-0.2747^{***}	-0.2877^{***}	-0.2971^{***}	-0.3020^{***}	-0.2709^{***}	-0.2834^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Gender (Female = 1)		0.1921^{***}		0.1358^{***}		0.2056^{***}
		(0.000)		(0.000)		(0.000)
Biological age at entry (years)		0.0030^{***}		-0.0035^{***}		0.0052^{***}
		(0.000)		(0.000)		(0.000)
Trades options (Yes = 1)		-0.0515^{***}		-0.1382^{***}		-0.0168
		(0.000)		(0.000)		(0.585)
Time effect (dates)	Yes	Yes	Yes	Yes	Yes	Yes
No. obs.	98,533,540	98,372,371	22,914,154	22,893,456	75,619,386	75,478,915

Significantly negative β^t after initial trading period

Lower exit likelihood with the increase of trading age.

Testing correlation between trading age at exit and performance at exit

$$\text{Performance}_{i,t_i^E} = \alpha + \beta^{t^E}\, t_i^E + \boldsymbol{X}_i \beta^{\boldsymbol{X}} + \epsilon_{i,t}$$

Rank as performance proxy: relative rank compared to peer investors

Table 5: Panel A: Performance proxy controlled for calendar time effects.

$\text{Performance}_{i,t_i^E} = Rank_{i,t_i^E}$	(1)	(2)	(3)	(4)	(5)
Intercept	0.4175***	0.4124***	0.3982***	0.4194***	0.3988***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Trading age at exit (days) t_i^E	0.9×10^{-5}***	1.0×10^{-5}***	0.8×10^{-5}***	0.9×10^{-5}***	0.8×10^{-5}***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Gender (Female = 1)		0.0282***			0.0256***
		(0.001)			(0.002)
Biological age at entry (years)			0.0005***		0.0004**
			(0.009)		(0.031)
TradesOptions (Yes = 1)				−0.0252**	−0.0249**
				(0.040)	(0.043)
R^2	0.0015	0.0028	0.0022	0.0020	0.0036
No. obs.	9,665	9,665	9,608	9,665	9,608

Rank of rank as performance proxy: relative rank of relative rank history

$$Rank_{i,t}^{S} = \frac{O_{i,t}^{S}}{t_i^E}$$

Table 6: Panel B: Performance proxy controlled for both calendar time and individual effects.

$\text{Performance}_{i,t_i^E} = Rank^S_{i,t_i^E}$	(1)	(2)	(3)	(4)	(5)
Intercept	0.4906^{***}	0.4932^{***}	0.5080^{***}	0.4891^{***}	0.5072^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Trading age at exit (days) t_i^E	$-1.1 \times 10^{-5***}$	$-1.1 \times 10^{-5***}$	$-1.0 \times 10^{-5***}$	$-1.0 \times 10^{-5***}$	$-1.0 \times 10^{-5***}$
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Gender (Female = 1)		-0.0148^{*}			-0.0130
		(0.097)			(0.150)
Biological age at entry (years)			-0.0004^{**}		-0.0004^{*}
			(0.043)		(0.070)
Trades Options (Yes = 1)				0.0201	0.0198
				(0.135)	(0.142)
R^2	0.0016	0.0019	0.0020	0.0018	0.0025
No. obs.	9,665	9,665	9,608	9,665	9,608

The higher the trading age, the lower the performance at exit: traders become "easier" on themselves, more tolerant of poor performance as the trading age goes.

Summary

We build a model which demonstrates, in accordance with empirical findings, that:

- Experienced traders perform on average better than inexperienced traders.
- Nevertheless, survivorship bias does not explain everything: at a given performance level, experienced traders are less likely to exit than inexperienced traders.
  - There exists a downward moving exit boundary.

References I

Gervais, S. and Odean, T. (2001).
Learning to be overconfident.
Review of Financial Studies, 14(1):1–27.

Nicolosi, G., Peng, L., and Zhu, N. (2009).
Do individual investors learn from their trading experience?
Journal of Financial Markets, 12(2):317–336.

Schraeder, S. (2016).
Information processing and non-Bayesian learning in financial markets.
Review of Finance, 20(2):823–853.

Seru, A., Shumway, T., and Stoffman, N. (2010).
Learning by Trading.
Review of Financial Studies, 23(2):705–739.

The End