Investment in High-Frequency Trading Technology: A Real Options Approach*

Laura Delaney†

January 30, 2018

Abstract

This paper derives an optimal timing strategy for a regular slow trader considering investing in a high-frequency trading (HFT) technology. The market is fragmented, and slow traders compete with fast traders for trade execution. Given this optimal timing rule, I then characterise the equilibrium level of fast trading in the market as well as the welfare-maximising socially optimal level. I show that there is always a unique cost of investment such that the equilibrium level of fast trading and the socially optimal level coincide. Finally I discuss potential policy responses to addressing equilibrium and social optimality misalignment in HFT.

Keywords: Finance, High frequency trading, Fragmented markets, Real options.

JEL Classification Numbers: C61, G10, G20.

1 Introduction

Over the last decade, the state of financial markets has changed considerably. In the first instance, markets have become highly fragmented. There are now more than 50 trading venues for U.S. equities - 13 registered exchanges and 44 so called Alternative Trading Systems (see Biais et al. [4] and O'Hara and Ye [22]). Hence, traders must search across many markets for quotes and doing so can be costly as it may delay full execution of their orders.

In response to the increase in market fragmentation, so called high frequency trading (HFT) technologies have been developed to reduce the associated costs borne by traders. HFT is a type of algorithmic trading that uses sophisticated computer algorithms to implement vast amounts of trades in extremely small time intervals. For example, traders can buy colocation rights (the placement of their computers next to the exchange's servers) which gives them fast

*Thanks to Jacco Thijssen, Saqib Jafarey, Carol Alexander, two anonymous referees, and seminar participants at the Young Finance Scholars Conference 2016 and the Bachelier Finance Society Conference 2016 for their much appreciated comments and suggestions.

†Department of Economics, City University, Northampton Square, London, EC1V 0HB, UK. Email: `laura.delaney.1@city.ac.uk Tel.:+44-(0)20-7040-4129.`

access to the exchange's data feed, they can invest in smart routers which can instantaneously compare quotes across all trading venues and then allocate their orders accordingly, or they can invest in high-speed connections to the exchanges via fiber optic cables or microwave signals. Proprietary trading desks, hedge funds, and so called pure-play HFT outlets are investing large sums of money into such technologies in an effort to outpace the competition. Indeed, according to Hoffman [17], recent estimates suggest that HFTs are now responsible for more than 50% of trading in U.S. equities.

In a recent paper, O'Hara [21] details the many ways in which market microstructure has changed over the past decade and calls for a new approach to research in this area which "reflects the new realities of the high frequency world". Nevertheless, there has been a growth in the finance literature on HFT in recent years (see a survey by Foucault [11]). Much of the literature is empirical and, on the whole, the consensus has been that HFT improves liquidity through lower bid-ask spreads (Hendershott et al. [16] and Hasbrouck and Saar [14]); is highly profitable (Menkveld [19] and Baron et al. [2]); and facilitates price discovery (Hendershott and Riordan [15] and Brogaard et al. [5]). The theoretical literature has also been growing in recent years. For example, Hoffman [17] presents a stylised model of HFT in a limit order market where agents differ in their trading speed; Pagnotta and Philippon [23] propose a model in which trading venues invest in speed and compete for traders who choose where and how much to trade; and Biais et al. [4] develop a model of equilibrium investment in a HFT technology in a Glosten and Milgrom [13] type framework.

While the finance literature strengthens our understanding of the nature and implication of HFT in many dimensions, to the best of my knowledge, the decision to be fast is always taken to be exogenous. However, the decision to be fast or slow is a real investment like any other. In particular, there is uncertainty associated with the payoff generated from investing in the HFT technology and the investment involves an upfront investment cost which is sunk. Moreover, the slow trader can adopt the technology at any future point in time with no terminal date. This adds a dynamic aspect to the investment decision which is not accounted for in other models. Hence, it makes sense to endogenise the investment decision and determine how the optimality of investment timing has implications for HFT in the marketplace. To this end, the problem has a place in the operations research literature.

In this paper I use a real options approach to determine analytically the optimal time for financial market traders to invest in a HFT technology such that the market is fragmented and slow traders compete with HFTs for trade execution. Given this optimal timing strategy, I then compare the equilibrium level of fast trading in the market with the welfare-maximising socially optimal level. While optimal investment timing has been well-developed in the operations research literature through its application to many different types of problems (see, for example, Banerjee et al. [1]; Battauz et al. [3]; Munoz et al. [20]; Delaney and Thijssen [8]), it is the first application of the approach in a HFT environment.

There are a number of novel results generated by the model, all of which arise from the optimal timing policy derived, in particular, the inclusion of a value of waiting to invest into the slow trader's value function. The results are as follows. (i) It is optimal to wait longer to invest if the level of high frequency trading in the market increases, and early adoption is optimal if the slow trader's probability of finding a liquid venue decreases. (ii) It is also optimal to wait longer if the uncertainty of the profit process increases and, if the probability of finding a liquid venue is low, if the discount rate increases and/or the shortfall in the expected rate of return from holding the option to invest decreases. However, if the probability of finding a liquid venue is high, an increase in the discount rate and a decrease in the shortfall make early adoption optimal. These comparative static results with respect to the discount rate and the shortfall are novel from a real options perspective, which I discuss in a later section of the paper. (iii) There is always a unique equilibrium level of fast trading in the market. (iv) The equilibrium level of fast trading decreases in the probability of finding a liquid venue. (v) There can be either under-investment or over-investment in equilibrium relative to the trading industry welfare-maximising socially optimal level. Over-investment arises when the cost of investing in the technology is low, and under-investment arises when the cost is high. (vi) There is a unique cost of investment such that the socially optimal level of fast trading and the equilibrium level coincide. (vii) Increases in the discount rate and/or uncertainty over the profit process alleviate the extent of over-investment in equilibrium and exacerbate the extent of under-investment. However, over-investment is alleviated and under-investment exacerbated by decreases in the shortfall in the expected rate of return from holding the option to invest.

I discuss the efficacy of using Pigouvian taxes or subsidies to align the equilibrium level of HFT with the socially optimal level. In the case of under-investment, subsidising slow traders' investment cost by an amount equal to the difference between the marginal effect of the level of HFT on utilitarian welfare at the socially optimal level and the equilibrium level appears to be an appropriate policy response to aligning these levels, but in the case of over-investment, taxing HFTs by an amount equal to this difference may not be the most effective response because it will not alter the prevailing level of HFT activity in the market. Instead, when there is over-investment, subsidising HFTs to exit the market by refraining from more fast trading, by an amount equal to the size of this difference, may be a more effective response in that case.

Finally I present and discuss some empirical implications which are generated by the model.

The remainder of this paper is organised as follows. The set-up of the model is described in the next section. In Section 3 the solution to the optimal stopping problem is given, as well as a brief discussion on the comparative statics. Section 4 characterises and compares the market equilibrium and welfare-maximising socially optimal levels of fast trading, as well a providing some discussion on policy implications. Section 5 provides some empirical implications and Section 6 concludes. All proofs are placed in the appendix.

2 The Model

Consider a risk-neutral market trader contemplating investment in a HFT technology. Time is continuous, the horizon is infinite, and indexed by $t \in [0, \infty)$. The trader discounts the future at the risk-free rate $r > 0$. Investing in the technology incurs a sunk cost $I > 0$.

The objective is to determine the optimal time to invest in the technology so that the trader's discounted expected payoff from investing is maximised. To solve for this optimal stopping problem, we must determine the expected present value of trading profits for the trader as if he were (i) fast and (ii) slow. These value functions depend on the trading environment which I describe in following subsection, and then I derive the value functions in accordance with this environment.

2.1 The Trading Environment

The market is fragmented and, as such, liquidity conditions vary across venues so that traders must search for quotes which can lead to delayed execution of orders. At every instant a fraction of the trading venues are "liquid". In the context of this model, a trading venue is liquid if the trader's order, when sent to the exchange venue, is fully executed.

At any instant, the trader can only send an order to one trading venue and his choice of venue is random. HFTs have extremely fast connection speeds to the market and can observe all venues instantaneously. Thus, if the trader is fast, he always finds a liquid one immediately upon order submission. He is indifferent between liquid venues. However, if the trader is slow, he must search for liquid trading venues and finding one can take time. Thus, at each instant he executes his trade with probability λ. Otherwise, with complementary probability, he must continue to search for a liquid venue.[1]

Obtaining empirical estimates of this probability λ is possible using the Fragulator software supplied by http://fragmentation.fidessa.com. This software provides comprehensive data and statistics on market fragmentation in the U.S., Europe, Japan, China, and Australia. One such statistic is the Fidessa Fragmentation Index (FFI) which provides a measure of how different stocks are fragmenting across primary markets and alternative venues. The FFI can range from 1 to V_N, where V_N is the number of venues trading a given stock. An FFI = 1 indicates trading residing on one venue, whereas an FFI = V_N indicates trading for the stock is spread evenly across all venues. Therefore, FFI/V_N would provide an appropriate empirical estimate of λ.

2.2 Valuations

The trading activities of a slow trader yields a stream of profits X^S in perpetuity, and the activities of a fast trader yields a profit stream X^F, such that both processes depend on a stochastic process $(X_t)_{t\geq 0}$ which is a geometric Brownian motion of the form:

$$dX = \mu X dt + \sigma X dW, \tag{1}$$

for constants $0 < \mu < r$ and $\sigma > 0$, which represent the drift and volatility of the process respectively, and $W = (W_t)_{t\geq 0}$ is a standard Brownian motion.

I describe the different profit process specifications for a slow and a fast trader later in the section, but will determine a process for each type that depends positively on the stochastic component X. Hereafter, I refer to X simply as the "profit process".

[1]This notion of a fragmented market environment has the flavour of that in Biais et al. [4].

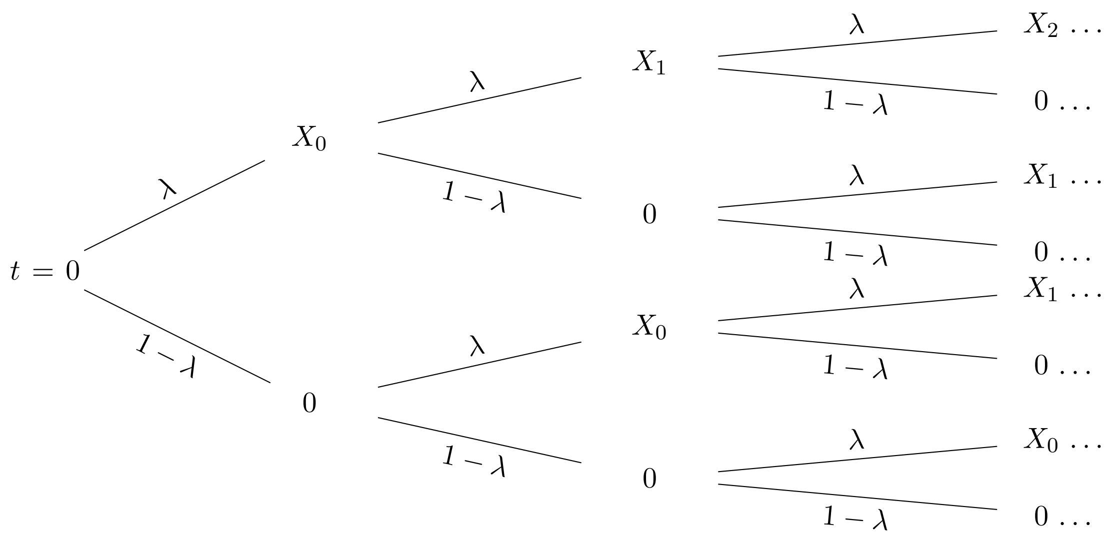

Figure 1: The Search Process

2.2.1 The Search Process

As discussed, if the trader is slow, he does not have access to the HFT technology and cannot observe all trading venues instantaneously. Thus, he may sometimes send orders to a trading venue which is not liquid. He can only send orders to one trading venue in any instant. If he sends an order to some venue at some instant t, it will be executed with probability λ. If it is not executed, I assume that the trader continues to search for a liquid venue for that order until it is executed. To describe the process, I let the dependence on time be explicit so that X_t denotes the value of X at time t. The process works as follows (a graphical depiction of which is given in Fig. 1 below):

- At $t = 0$: the trader submits an order which delivers a profit X_0 with probability (w.p.) λ, otherwise gets 0.
- At $t = 1$ the trader will do *one* of the following:
  1. If his order is executed at $t = 0$, he submits another order at $t = 1$ which will yield, if filled, X_1. It gets filled with probability λ.
  2. If, however, his $t = 0$ order is not filled, at $t = 1$ he is searching for a liquid venue for his $t = 0$ order and gets it filled w.p. λ. If filled, it yields the profit X_0.

  Thus,

$$\begin{aligned} E^0[X_1^S] =&\lambda^2 E^0[X_1] + (1-\lambda)\lambda X_0 \\ =&\lambda X_0 \left(\lambda e^{\mu} + (1-\lambda)\right) \end{aligned} \tag{2}$$

  since X_1 and X_0 are governed by (1).
- The process continues in this way as long as the trader is slow.

Therefore, for some arbitrary instant $t \geq 0$, the search process described implies that

$$\begin{aligned} E^0[X_t^S] =& \lambda(1-\lambda)^t X_0 \sum_{i=0}^{t} \binom{t}{i} \left(\frac{\lambda e^{\mu}}{1-\lambda}\right)^i \\ =& \lambda X_0((1-\lambda)+\lambda e^{\mu})^t \end{aligned} \tag{3}$$

by the Binomial theorem.

If, however, the trader is fast, he gets his order filled with certainty whenever he sends it to the market because he can observe which venues are liquid. Thus, he does not need to search for liquidity and at every instant t that he trades, he gets a profit of X_t. Thus, $X_t^F \equiv X_t$ and

$$E^0[X_t^F] = X_0 e^{\mu t}, \tag{4}$$

which is, in fact, Eq. (3) for $\lambda = 1$.

Moreover, since $\lambda < 1$, $E^0[X_t^F] > E^0[X_t^S]$ for all t. This says if the trader is fast, he has higher trading profits in expectation than if he were slow because his execution delay costs are zero. This is intuitive and necessary because if this condition did not hold, investment in the HFT technology would never be optimal.

2.2.2 Prevalence of HFT in Market

Much of the literature in this area documents that HFTs generate adverse selection costs for slow traders because their accelerated access to value-relevant information for the asset implies they profit more at the expense of slow traders (see, for example, empirical studies by Baron et al. [2] and Brogaard et al. [5]). Moreover, as documented in Biais et al. [4], anecdotal evidence suggests that the profitability of HFTs has declined in recent years which may be due to the fact that the prevalence of high frequency trading in the market place has increased. I capture these stylized facts by assuming that $X_0 := (1+\alpha)^{-1}x$, for some $x > 0$, and where I define $\alpha < 1$ as the fraction of HFT activity in the market at each instant. Therefore, since the profit processes both follow a geometric Brownian motion, the current (discounted) expected profit from trading for *each type* of trader, for any future time, decreases in the level of HFT activity.

I should point out here that since a higher level of α is simply equivalent to a lower initial level of the profit process, this parameter is of particular interest in Section 4, where I examine the equilibrium level of fast trading in the market, as well as the trading industry value maximising level.

3 The Optimal Stopping Problem

If the trader decides to adopt the technology at time τ, the present value of the entire investment opportunity, denoted by $V(x)$, is given by the expected discounted profit flow of a slow trader between now (i.e., $t = 0$) and time τ, at which time he pays the investment cost I and becomes a fast trader. From time τ onwards, the present value of the investment opportunity is the expected discounted profit flow of a fast trader[2]; i.e.;

$$V(x) = E^0 \left[\int_0^\tau e^{-rt} X_t^S dt + \int_\tau^\infty e^{-rt} X_t^F dt - e^{-r\tau} I \right] \tag{5}$$

where $E^{t'}$ denotes the expectation operator applied at time t'.

Using the strong Markov property of diffusions, we can re-write Eq. (5) as

$$\begin{aligned} V(x) =& E^0 \left[\int_0^\infty e^{-rt} X_t^S dt - \int_\tau^\infty e^{-rt} X_t^S dt + \int_\tau^\infty e^{-rt} X_t^F dt - e^{-r\tau} I \right] \\ =& E^0 \left[\int_0^\infty e^{-rt} X_t^S dt \right] + E^0 \left[e^{-r\tau} E^\tau \left(\int_0^\infty e^{-rt} \left(X_t^F - X_t^S \right) dt - I \right) \right]. \end{aligned} \tag{6}$$

The problem is to find a value function $V^*(x)$ and a stopping time $\tau^* > 0$ such that the following optimal stopping problem is solved:

$$V^*(x) = \sup_{\tau \in \mathcal{T}} E^0 [e^{-r\tau} F(X_\tau^S, X_\tau^F)], \tag{7}$$

for $\mathcal{T}$ the set of stopping times, and

$$F(X_\tau^S, X_\tau^F) = E^\tau \left[\int_0^\infty e^{-rt} \left(X_t^F - X_t^S \right) dt \right] - I \tag{8}$$

is the trader's payoff from adopting the technology. Hereafter I denote $F(X_t^S, X_t^F)$ by $F(X_t)$ since both X^S and X^F are functions of X.

By Fubini's theorem, and using Eqs. (3) and (4), Eq. (8) can be re-written as

$$\begin{aligned} F(X_\tau) =& X_\tau \left(\frac{1}{\delta} - \lambda \int_0^\infty e^{-rt} \left((1-\lambda) + \lambda e^{r-\delta} \right)^t dt \right) - I \\ =& \Omega(r, \delta, \lambda) X_\tau - I \end{aligned} \tag{9}$$

where $\delta := r - \mu > 0$ is the convenience yield (or shortfall)[3], and the scaling term

$$\Omega(r, \delta, \lambda) = \frac{1}{\delta} - \frac{\lambda}{(r - \ln((1-\lambda) + \lambda e^{r-\delta}))} > 0 \tag{10}$$

represents the relative advantage (in terms of profits) from being fast. Note that I have replaced μ by $r - \delta$ for the purpose of analysing the solution and comparing it with that obtained in a standard real options model with the payoff flow being obtained in perpetuity. Hereafter, δ will be referred to as the "shortfall".

[2]This is essentially an exchange of one profit flow for another by paying a fixed cost; see for example Dixit and Pindyck [9] Chapter 9, Section 3. However, in that example, there are two investors in a particular project and the leader's profit flow is exchanged for a different (lesser) profit flow once the follower invests.

[3]In the context of real options models of this type, the convenience yield is interpreted as the shortfall in the expected rate of return from holding the option to invest rather than the HFT technology, and hence represents an opportunity cost of waiting rather than investing now (Dixit and Pindyck [9]). This ought not to confused with the market price of risk, which assumes a similar notation in other financial equilibrium models such as the CAPM, for example.

I present the derivation of Eq. (10) in Appendix A, where I also prove that $\Omega(r, \delta, \lambda) > 0$. The relative advantage of being fast is induced by the elimination of delayed execution arising from the need to search for quotes in fragmented markets.

Theorem 1. *Let $\beta_1 > 1$ be the positive root of the quadratic equation*

$$\mathcal{Q}(\beta) = \frac{1}{2}\sigma^2\beta(\beta - 1) + (r - \delta)\beta - r = 0. \tag{11}$$

Investment takes place at the first passage time $\tau^* = \inf\{t : X_t \geq X^*\}$, for some constant

$$X^* = \frac{\beta_1}{\beta_1 - 1} \left(\Omega(r, \delta, \lambda)\right)^{-1} I. \tag{12}$$

Moreover, the optimal stopping problem (7) is solved by

$$V^*(x) = \begin{cases} \frac{\lambda(1+\alpha)^{-1}x}{\left(r - \ln\left((1-\lambda)+\lambda e^{r-\delta}\right)\right)} + \left(\frac{(1+\alpha)^{-1}x}{X^*}\right)^{\beta_1} F(X^*) & \text{if } x < (1+\alpha)X^* \\ \frac{(1+\alpha)^{-1}x}{\delta} - I & \text{if } x \geq (1+\alpha)X^*, \end{cases} \tag{13}$$

where $F(X)$ is given by Eq. (9).

Proof. See Appendix B. ■

3.1 Comparative Statics

In this subsection I discuss the economics underlying the optimal investment strategy.

First I discuss the effects of the two HFT related parameters, α and λ, on the optimal investment threshold X^*.

Proposition 1. *An increase in the level of fast trading activity, α, and an increase in the probability of finding a liquid venue, λ, makes it optimal for a slow trader to wait longer before investing.*

The level of HFT activity in the market does not directly affect the profit threshold X^* above which it is optimal to become fast. However, an increase in the level of α reduces a slow trader's current profit from trading since $X_0 := (1 + \alpha)^{-1}x$. Therefore, the slow trader will have to wait longer before it is optimal to invest since an increase in α increases the distance between the current profit level and X^*.

The probability of finding a liquid venue impacts the optimal threshold via its effect on the relative advantage of being fast; in particular, through its effect on $\Omega(r, \delta, \lambda)$. The relative advantage of being fast is high when the probability of finding a liquid venue is low and thus, it is optimal to invest sooner if the probability of finding a liquid venue decreases. This is because it becomes harder for the slow traders to find quotes quickly and owing to the search process, the more they lag behind the HFTs in terms of trading profits, i.e.; $E^0[X^F] - E^0[X^S]$ decreases in the probability of finding a liquid venue.

The following proposition states the effect of the other parameters which are standard in all real option models, but not particular to a HFT environment, on the optimal investment threshold. The results, are interesting in their own right because they differ somewhat from their effects on the investment threshold in the more standard real options investment model (see for example, Dixit and Pindyck [9], Chapter 6) as I discuss below, but they are also useful in understanding the driving forces underlying the extent of over- and under-investment in the HFT technology in equilibrium, relative to the trading industry welfare maximising level of investment, which I discuss in Section 4.1

Proposition 2. *The optimal investment threshold X^* increases in the uncertainty in the profit process σ. Moreover, X^* increases in the discount rate r and decreases in the shortfall δ when the probability of finding a liquid venue is low, and vice versa when this probability is high.*

Proof. See Appendix C. ■

The effect of uncertainty on X^* is as expected from other standard models of investment under uncertainty (see for example, Dixit and Pindyck [9] and McDonald and Siegal [18]). It impacts the threshold via the value of waiting. In particular, the more volatile are the profits from trading, the greater is the opportunity cost of investing in the technology since the cost of doing so is high and the investment, once made, is irreversible.

The discount rate and the shortfall impact the optimal threshold X^* via their effect on the value of waiting (the option effect) and their effect on the present value of the difference in perpetuity factors of fast and slow traders (i.e., its effect on $\Omega(r, \delta, \lambda)$)

The equilibrium relationship $\delta = r - \mu$ must be maintained. Therefore, if r increases, so too does the rate of growth of profit from being fast μ and, hence, the expected appreciation in the value of the option to invest in the HFT technology increases making it costlier to invest immediately rather than to wait. However, if δ increases while keeping the discount rate fixed, μ will decrease implying that it is costlier to wait rather than to invest immediately because the expected appreciation in the value of the option to invest decreases.[4]

However, an increase in r also increases the present value of the relative advantage of being fast when δ is fixed (cf. Eq. (C.4) in Appendix C). This is because the expected present value of a fast trader's profit flow does not change when the shortfall is constant, but an increase in r reduces the expected present value of a slow trader's profit flow. Hence, the relative advantage of being fast increases implying earlier investment is optimal. On the other hand, an increase in δ, while keeping r fixed, reduces the present value of the relative advantage of being fast (cf. Eq. (C.6) in Appendix C) because it reduces the expected present value of a fast trader's profit flow more than it increases the expected present value of a slow trader's profit flow. This is

[4] We can verify both of these results technically since

$$\frac{\partial}{\partial r}\left(\frac{\beta_1}{\beta_1 - 1}\right) = -\frac{1}{(\beta_1 - 1)^2}\frac{\partial \beta_1}{\partial r} > 0$$

and

$$\frac{\partial}{\partial \delta}\left(\frac{\beta_1}{\beta_1 - 1}\right) = -\frac{1}{(\beta_1 - 1)^2}\frac{\partial \beta_1}{\partial \delta} < 0$$

since $\partial\beta_1/\partial r < 0$ when δ fixed and $\partial\beta_1/\partial\delta > 0$ when r fixed (cf. Appendix C).

because an increase in δ is synonymous with a decrease in the rate of profit growth from being fast. Thus, an increase in δ implies waiting longer is optimal.

So long as the probability of a slow trader finding a liquid venue λ is very low, then the relative advantage of being fast will always remain high so that any change in either r or δ will not significantly impact $\Omega(r, \delta, \lambda)$. Therefore, the effect of any shift in either of the parameters on the value of waiting dominates its effect on the relative advantage of being fast. This implies that an increase in r makes waiting longer optimal, but an increase in δ makes earlier investment optimal when λ is low owing to the option effect.

On the other hand, when the probability of a slow trader finding a liquid venue is high, then value of waiting will always be high and this will not be impacted significantly by any shift in r or δ. Hence, the effect of a change in r or δ on the relative value of being fast dominates their effects on the value of waiting implying that an increase in r will make earlier investment optimal and an increase in δ will make waiting longer optimal in this case, both owing to the present value effect.

The results with respect to r and δ are particularly interesting from the real options perspective because they differ from more standard result on the (positive) effects of the discount rate (when δ is fixed) and the shortfall (when r is fixed) on the optimal investment threshold when the post-investment profit flow is received in perpetuity rather than as a lump sum payout (see, for example, Dixit and Pindyck [9], Chapter 6).

In their model, the effect of r on the threshold is always via its effect on the value of waiting when the shortfall δ is kept constant. This is because the pre- and post- investment profit flows are linear functions of the underlying geometric Brownian motion, implying that the present value of the investment is unaffected by changes in r when δ is fixed. Hence, all the impact of r on the threshold comes from its effect on the value of waiting. However, in my model, the post-investment profit flow is a linear function of the geometric Brownian motion describing the profit flow, but the pre-investment profit flow is not. This is because the slow trader must search for quotes in order to find a liquid venue. Hence, the present value of investing is not unaffected by a change in r in this model and, therefore, the non-linearity of the pre-investment profit flows, as a result of the search procedure, implies that the threshold may increase or decrease in r owing to the option or present value effects, respectively.

In the case of δ, there are two opposing effects in the Dixit and Pindyck [9] model also, but the present value effect always dominates the option effect so that an increase in δ always implies waiting longer is optimal. This corresponds with my result for a high λ, but an additional effect whereby early investment is optimal will be generated in my model when the probability of finding a liquid venue is low. This is because I include in the model the search procedure for slow traders when finding liquid venues is not certain, and this produces the additional effect.

4 Levels of Fast Trading

In this section, I characterise the equilibrium and socially optimal levels of fast trading conditional on the optimal timing strategy derived above.

I assume that the initial realised profit levels are i.i.d across traders and continuously distributed on $[0, (1+\alpha)^{-1}\overline{x}]$, where $\bar{x} > (1+\alpha)X^*$. In other words, x^i is i.i.d. across traders i and continuously distributed on $[0, \overline{x}]$ with cumulative distribution function $G(\cdot)$ and density $g(\cdot)$, where $x^i := (1+\alpha)X_0^i$ from Section 2.2.2.[5] From Theorem 1 it is optimal for all traders i whose $x^i < (1+\alpha)X^*$ to refrain from investing, and it is optimal for all traders i whose $x^i \geq (1+\alpha)X^*$ to invest immediately. Hence, given their initial endowments, the fraction of traders in the market who find it optimal to be slow is $G((1+\alpha)X^*)$ and the remaining $1-G((1+\alpha)X^*)$ find it optimal to be fast. Therefore, the equilibrium level of fast trading must satisfy the following condition:

$$\alpha^* = 1 - G((1+\alpha^*)X^*). \tag{14}$$

This condition characterises a level of equilibrium that is always unique and the fact that there is always uniqueness of equilibrium is novel. The uniqueness arises from the optimal stopping rule in the following way. A high level of α implies that the initial profit level of a trader, irrespective of whether he is fast or slow, is low. However, this high value of α implies that the range of initial profit levels over which it is optimal to be slow ($[0, (1+\alpha)X^*)$) is also high. In other words, $G((1+\alpha)X^*)$ increases in α. Correspondingly, a high level of α implies the range of profit levels over which it is optimal to be fast ($[(1+\alpha)X^*, \bar{x}]$) is low; i.e., $1 - G((1+\alpha)X^*)$ decreases in α. The monotonic increase in $G(1+\alpha)X^*$ implies that the condition (14) yields an equilibrium level of HFT that is unique.

Multiple equilibria can arise in other financial market models because of virtuous or vicious circles. For example, in Glosten and Milgrom [13] and Dow [10], if traders anticipate the market will be liquid, they will submit many orders, and hence, the market is liquid, and vice versa. In a similar vein, in Biais et al. [4], one equilibrium can arise in which no trader invests because each expects others not to invest, and one in which all traders should become fast because all expect everyone else to be fast. This situation of multiple equilibria arising from self-fulfilling cycles could only arise in the current model if $G((1+\alpha)X^*)$ were to increase in the level of HFT for some values of α and decrease in α for other values. If $G((1+\alpha)X^*)$ were to decrease in α, then we would have a situation in which for low values of α, it will be optimal for more traders to be slow than to be fast, and vice versa. Therefore, multiplicity of equilibria is ruled out in my model because the optimal stopping rule given in Theorem 1 ensures $G((1+\alpha)X^*)$ increases everywhere in α.

Proposition 3. *The equilibrium level of fast trading decreases in the probability of finding a liquid venue.*

Proof. See Appendix D. ■

The result is intuitive and the reasoning is as follows. A decrease in the probability of finding a liquid venue implies earlier investment is optimal (cf. Proposition 1). But earlier investment

[5] Differences in X_0^i (via differences in x^i) reflect the fact that trading activity in the market is heterogeneous in the sense that traders, irrespective of whether they are fast or slow, trade different assets, different volumes, etc. at any instant. Therefore, initial profit levels vary across traders in the market so that at any given instant, it will be optimal for some to be fast and others to be slow according to the optimal policy given in Theorem 1.

being optimal implies that the range of values for which it is optimal to be fast increases; i.e., the fraction of traders who find it optimal to be fast increases and, thus α^*, increases.

4.1 Socially Optimal versus Equilibrium Fast Trading

The socially optimal level of fast trading is the level of α that maximises utilitarian welfare. Say the level of HFT in the market at time t is α. From Theorem 1, the average profit from all the slow trading in the market, denoted by $V^s(\alpha)$ is given by

$$V^s(\alpha) = \int_0^{(1+\alpha)X^*} \left[V^*(x^i)\Big|x^i < (1+\alpha)X^*\right] g^s(x^i)dx^i \tag{15}$$

and the average profit from all fast trading is given by

$$V^f(\alpha) = \int_{(1+\alpha)X^*}^{\overline{x}} \left[V^*(x^i) + I\Big|x^i \geq (1+\alpha)X^*\right] g^f(x^i)dx^i \tag{16}$$

where $V^*(x^i)$ is given in Eq. (13), and $g^f(\cdot)$ and $g^s(\cdot)$ denote the particular value of the density function which pertains to fast and slow trading, respectively. Thus

$$g^s(x^i) = \frac{g(x^i)}{G(x^i)}$$

and

$$g^f(x^i) = \frac{g(x^i)}{1 - G(x^i)}$$

where $G(\cdot)$ and $g(\cdot)$ are as previously defined.

I denote the relative advantage of fast trading in the market by $\Delta(\alpha)$, where

$$\Delta(\alpha) := V^f(\alpha) - V^s(\alpha).$$

Hence, the utilitarian welfare is given by

$$\begin{aligned} W(\alpha) =& \alpha\left(V^f(\alpha) - I\right) + (1-\alpha)V^s(\alpha) \\ =& \alpha\left(\Delta(\alpha) - I\right) + V^s(\alpha). \end{aligned} \tag{17}$$

Denoting the socially optimal level of fast trading in the market by α^{SO}, then α^{SO} solves

$$W'(\alpha^{SO}) = \Delta(\alpha^{SO}) - I - \left(-\alpha^{SO}\frac{\partial V^f(\alpha)}{\partial \alpha}|_{\alpha=\alpha^{SO}} - (1-\alpha^{SO})\frac{\partial V^s(\alpha)}{\partial \alpha}|_{\alpha=\alpha^{SO}}\right) = 0, \tag{18}$$

where $W'(\alpha^{SO}) = \frac{\partial W(\alpha)}{\partial \alpha}|_{\alpha=\alpha^{SO}}$.

The socially optimal level of HFT is not necessarily zero. If the socially optimal level were to be zero, then the following condition would have to hold: $W'(0) = \Delta(0) - I + \frac{\partial V^s(\alpha)}{\partial \alpha}\Big|_{\alpha=0} = 0$. However, I show in Appendix E that $\partial V^s(\alpha)/\partial\alpha > 0$, for all α, implying that the condition cannot hold if $\Delta(0) - I > 0$; in other words, it cannot hold if the relative value of being fast is high when everyone else is slow. Since $G((1+\alpha)X^*)$ increases in α, if $\alpha^{SO}=0$, then $G((1+\alpha^{SO})X^*)=G(X^*)$ is low; i.e., if $\alpha^{SO}=0$, the fraction of traders for which it is optimal to be fast is relatively high. Hence, in the model, the relative value of being fast is high when everyone else is slow. Therefore, the socially optimal level of fast trading will be non-zero. This implies that it is the relative gain from fast trading which gives HFT its social value in the model. This relative gain is owing to the reduction in delay costs which trading in fragmented markets can inflict on those without access to the technology.

On the other hand, it would be socially optimal for all traders in the market to be fast if $W'(1)=\Delta(1)-I+\frac{\partial V^f(\alpha)}{\partial\alpha}\Big|_{\alpha=1}=0$ holds. However, as shown in Appendix E, $\partial V^f(\alpha)/\partial\alpha<0$ for all α, the condition will not be satisfied if $\Delta(1)-I<0$; i.e., if the expected relative value of fast trading in the market is low when everyone if fast, which must hold since $G((1+\alpha)X^*)$ increases in α. Thus, it is not socially optimal for all traders to be fast when the expected relative value of fast trading in the market is low if everyone in the market is fast.

Proposition 4. *When the cost of investing in the HFT technology is low, the equilibrium level of fast trading is never lower than the socially optimal level. However, for sufficiently high levels of I, the equilibrium level will be lower than the socially optimal level. Moreover, there is always a cost level such that the equilibrium level and the socially optimal level coincide.*

The reasoning is as follows. If α^{SO}, the solution to Eq. (18), is less than α^*, there is over-investment in the fast trading technology in equilibrium relative to what is socially optimal. I show in Appendix D that α^* decreases in X^*. Thus, over-investment in equilibrium will arise when it is optimal to invest early and this is the case when the cost of investing I is low because the low cost implies a low value of waiting.

The over-investment in speed in equilibrium relative to the socially optimal level result has been found in other papers, but for different reasons to the one in this paper. See, for example, Budish et al. [6] who develop a model in which traders invest in speed so that they can quickly react to the arrival of public information but, in contrast to this model, gains from trade are not modelled. This absence in trading gains leads to a result that slow trading is always socially optimal. In Biais et al. [4], each investor also chooses his speed at which it operates in a given market, but in that paper, over-investment in speed always arises and that is because fast traders benefit from the adverse selection cost they inflict on slow traders, but they do not internalise this cost in equilibrium. This increases the level of HFT in equilibrium, but the socially optimal level does not increase in the benefits from adverse selection. Hence, over-investment arises. Moreover, Glode et al. [12] view investment in HFT as an arms race, and there is an equilibrium in which every investor chooses to invest in the technology in the fear others will also do so, which leads to over-investment.

On the other hand, if $\alpha^{SO}>\alpha^*$, there is under-investment in equilibrium relative to the socially optimal level. This arises when X^* is high and, thus, when the cost of investing is high making the value of waiting relatively valuable.

There is little evidence in the literature of under-investment in the HFT technology. An extension of the model in Biais et al. [4], which includes markets where fast trading is not permitted, can lead to under-investment in equilibrium. This is because slow traders opt to trade only in slow markets which, in turn, reduces the expected profits of HFTs. In their extended model, only two types of equilibria can arise: all traders trade in slow markets or all are fast. By contrast, under-investment can occur in my model without considering the possibility of slow only markets. This is because I view investment in the HFT technology as an optimal stopping problem and, owing to the inclusion of the value of waiting, under-investment can occur in equilibrium (and not necessarily either an all fast ($\alpha^* = 1$) or all slow ($\alpha^* = 0$) equilibrium) when the cost of investing is very high. In my case, when under-investment arises, it does so because it is not optimal for enough traders to avail of the reduction in delay cost which investment would provide because the benefit of this reduction is not sufficiently high to warrant paying such a high investment cost owing to the value of the option to wait.

Therefore, owing to the optimal timing policy, it is the case that for some levels of investment cost $\alpha^{SO} < \alpha^*$, and for other levels $\alpha^{SO} > \alpha^*$, implying there is a unique I such that the condition

$$\alpha^{SO} = 1 - G((1 + \alpha^{SO})X^*) \tag{19}$$

is satisfied, and this is the level of I such that the equilibrium level and the socially optimal level of HFT coincide.

Another interesting question is how the regions of over and under-investment in equilibrium are related to changes in the standard real option parameters r, σ, and δ (via a change in μ). In essence, we want to understand how shifts in these parameter values impact the extent of over- and under-investment. The following proposition summarises the impact they have.

Proposition 5. *The extent of over-investment decreases in r and σ and increases in δ. The extent of under-investment increases in r and σ and decreases in δ.*

The extent of over-investment in equilibrium is alleviated when α^* decreases. Therefore, any parameter shift that causes a ceteris paribus increase in X^* will reduce the extent of over-investment in equilibrium (cf. Eq. (D.2) in Appendix D). Conversely, any parameter shift that causes a decrease in X^* will alleviate the extent of under-investment.

We know from Proposition 2 that X^* increases in σ and in r and decreases in δ when the probability of finding a liquid venue is low. However, when λ is low, we know from Proposition 3 that the equilibrium level of investment is high and hence, we can infer that the over-investment scenario arises in this case. Thus, when we have over-investment in equilibrium relative to the socially optimal level, X^* increases in r and σ and decreases on δ implying that increases in σ and/or r and/or decreases in δ alleviates the extent of over-investment.

By the same reasoning, an increase in σ will exacerbate the extent of under-investment in equilibrium. However, in the case of under-investment, the probability of finding a liquid venue will be high and, therefore, X^* will decrease in r and increase in δ (by Proposition 2). So, a decrease in r and/or an increase in δ will lead to an increase in X^* and, consequently, an exacerbation in the extent of under-investment in equilibrium.

4.2 Implications for Policy

The analysis so far leads to the inference that one appropriate policy response would be to tax or subsidise HFT activity in a manner akin to imposing a Pigouvian tax or subsidy. This tax (or subsidy), denoted by T^{**}, would align the equilibrium level of HFT with the socially optimal level when it satisfies the following equation

$$W'(\alpha^{SO}) = W'(\alpha^*) + T^{**}, \tag{20}$$

where $W'(\alpha)$ is defined in Eq. (18).

When we have over-investment in equilibrium relative to the socially optimal level, so T^{**} would take the form of a tax, but in the case of under-investment it would be a subsidy. The result from Proposition 4 indicates that the tax will be higher when the cost of investment is low because this encourages early investment in the HFT technology, and overall leads to over-investment in equilibrium relative to the socially optimal level. On the other hand, for sufficiently high levels of investment cost, the cost should actually be subsidised to encourage slow traders to invest sooner so that they avail of the reduction in delay cost that HFT provides in fragmented markets.

A caveat with this approach is that when there is over-investment, imposing a tax on HFTs when the cost is low will deter slow traders from adopting the technology as it essentially increases the cost of investing making waiting relatively more valuable, but it will not reduce the level of HFT that is already prevalent because there is no incentive with such a policy for HFTs to stop "trading fast". Such a caveat does not arise in the under-investment case because reducing the cost of investment by a subsidy will make investing relatively more attractive to slow traders.

One way of circumventing the issue in the over-investment case would be to subsidise HFTs to stop trading fast. This subsidy, S^{**} would need to be large enough so that the HFTs are at least indifferent between being fast or slow; i.e., S^{**} is such that

$$E[V^*(x)\Big| x < (1+\alpha)X^*] + S^{**} \geq E[V^*(x)\Big| x \geq (1+\alpha)X^*], \tag{21}$$

where $V^*(x)$ is defined in Eq. (13).

If $|S^{**}| = |T^{**}|$, then the level of HFT in equilibrium will also be socially optimal.

5 Empirical Implications

Most of the existing theoretical and empirical literature related to HFT is concerned with two main aspects (cf. Hoffman [17]) (i) the impact of HFT on market quality (see, for example, Hendershott et al. [16] and Hasbrouck and Saar [14]) and (ii) the identification of trading activity by HFTs in order to study the behaviour of different traders' types (see for example, Menkveld [19] and Brogaard et al. [5]). The novel contribution of this paper is to view HFT as a corporate investment opportunity and, to this end, my model generates a number of empirical implications relevant for HFT, but from a different perspective to those earlier research efforts. In this sense, the current model, and future extensions to it, have the potential to be of strong appeal to financial practitioners and regulators.

Implication 1. *The effect of market fragmentation on the prevalence of HFT in the market is concave over time.*

When the market is highly fragmented, the probability of finding a liquid venue λ is low. In my model, this makes early investment in the HFT technology optimal which, in turn, increases the prevalence of HFT activity in equilibrium. This implication corresponds well with the notion that market fragmentation stimulated the growth in fast trading technologies. Chiyachantana and Jain [7] find that in a sample of institutional investors, delays in order execution owing to market fragmentation are attributable to approximately one-third of total costs. Such delay costs can be reduced by investing in HFT technologies. However, as this prevalence grows, the initial profit levels of traders entering the market declines in my model (modelled in response to the empirical and anecdotal evidence that the profitability of HFT has declined in recent years (see Section 2.2.2) above) implying, from Proposition 1, that it becomes optimal for traders to wait longer before adopting the technology. This will eventually lead to a decline in the level of HFT. This latter feature is not considered in the current model, as it is beyond the scope of the present paper to consider the sequential nature of technology adoption, but one would expect that empirical studies investigating the effects of market fragmentation on the prevalence of HFT is concave over time.

Implication 2. *The relative value of HFT declines in the prevalence of fast trading.*

This implication arises from the optimal policy defined in Theorem 1 and the ensuing Proposition 1 that it is optimal to wait longer before investing when α is high. It is consistent with the anecdotal evidence, discussed in Section 2.2.2 that the profitability of HFT has declined in recent years which is likely to be attributable to the growth in the number of fast institutions.

To illustrate this effect, I plot the relative value of HFT against α in Fig. 2 below for a uniform distribution function and the following parameter values: $I = 1$, $\overline{x} = 3$, $\delta = 0.03$, $\sigma = 0.2$ and $\lambda = 0.2$.

Implication 3. *The extent of trading activity that takes place on over-the-counter and dark markets decreases in the cost of the HFT technology.*

HFT takes place on centralised electronic limit order book markets which are faster than over-the-counter and dark markets. When the cost of the technology is low, it is optimal to invest in the HFT technology early and, thus, the level of HFT in equilibrium will be high (see Proposition 4). Those traders for whom investing in the technology will never be optimal will seek out ways of insulating themselves from the predatory behaviour of HFTs as its prevalence grows and, as such, are likely to migrate to slower markets on which HFT does not take place. According to Biais et al. [4], an article in the New York Times in 2013[6] notes that "*Investors*

[6] See "As markets heat up, trading slips into the shadows", New York Times, March 31, 2013.

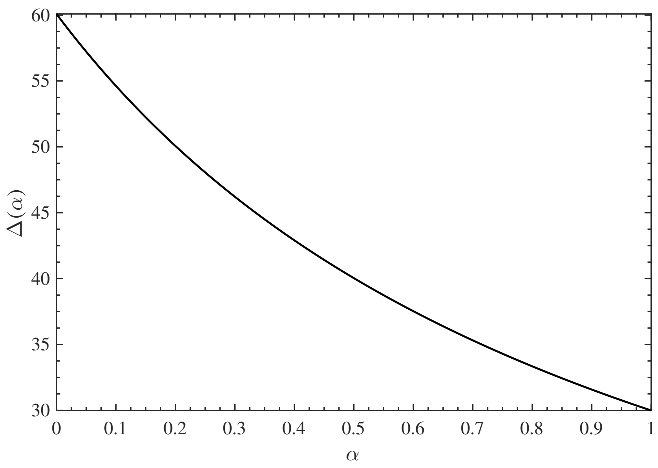

Figure 2: The effect of α on $\Delta(\alpha)$.

have said that they have moved more of their trading into the dark because they have grown more distrustful of the big exchanges like the NYSE and the Nasdaq. These exchanges have been hit by technological mishaps and become dominated by so-called high-frequency traders."

6 Concluding Remarks

In recent years, the state of market microstructure has changed considerably. There are many ways in which these changes have come about, but one of the biggest changes is that markets have become highly fragmented. When markets are fragmented, traders must search across many markets for venues which will execute their orders at their specified prices. This can result in delayed or partial execution which is costly. In response to the increase in market fragmentation, there has been a demand for speed by traders, and various types of expensive technologies have been developed. Such technologies enable traders to compare all trading venues instantaneously or obtain a glimpse of the true state of the market before everyone else.

In this paper I derive a dynamic model, using techniques from real options analysis, which provides an optimal timing strategy for slow traders to invest in a high frequency trading technology. The model prescribes waiting longer to invest if the level of high frequency trading in the market increases, and it prescribes earlier adoption if the probability of finding a liquid venue decreases. It also prescribes waiting longer if the uncertainty of the profit process increases and, if the probability of finding a liquid venue is low, if the discount rate increases and/or the shortfall decreases. However, if the probability of finding a liquid venue is high, an increase in the discount rate and a decrease in the shortfall make early adoption optimal.

Based on this optimal timing strategy, I then characterise the equilibrium level of fast trading in the market as well as the welfare-maximising socially optimal level. From this analysis, the following results emerge. There is always a unique equilibrium level of fast trading in the market, and this level decreases in the probability of finding a liquid venue. There is also a socially optimal level of fast trading such that, when the cost of adopting the technology is relatively cheap, there is over-investment in equilibrium relative to what is socially optimal, and when the cost is relatively high, there is under-investment in equilibrium. This implies that there is a unique cost of investment such that the equilibrium level and the socially optimal level of fast trading coincide. Moreover, increases in the discount rate and/or uncertainty over the profit process alleviate the extent of over-investment in equilibrium and exacerbate the extent of under-investment. However, over-investment is alleviated and under-investment exacerbated by by decreases in the shortfall. All of these results are driven by the optimal timing policy.

I discuss why a Pigouvian subsidy given to slow traders to encourage investment when the cost of investment is high is a reasonable policy response to aligning the equilibrium level of HFT with the socially optimal level in the case of under-investment, but that a Pigouvian tax imposed on fast traders when the cost is low, may not be the most effective means of obtaining alignment in the case of over-investment. However, subsidising HFTs to refrain from trading fast (i.e., trading using their fast technology) may be effective in that instance.

It is clear that further research on this issue is warranted and one possibility could be to extend the model by assuming regulators impose a stochastic per period tax on HFTs which is positively correlated with the extent to which the equilibrium level and the socially optimal level are misaligned. If the fast trader had then the option to “exit the market”, or in other words, revert to being slow if paying this tax became too costly relative to the benefit from having no delay costs, then this may be a more effective way of aligning market equilibrium with social optimality (than paying the HFTs to “exit” when there is over-investment) because this could then simultaneously change the *current* level of HFT as well as changing the optimal adoption threshold for slow traders. This would require a high level of surveillance to ensure that those fast traders who no longer pay the “fast tax” also no longer use co-location services, smart routers etc. The approach would also work in the under-investment case because if the subsidy is stochastic and positively correlated with the extent of under-investment, then if the under-investment problem is large, the subsidy would be large enough to encourage a high level of investment, but if the problem is small, it would be small enough to only encourage a small increase. Also, a stochastic subsidy would be unlikely to alter the level of HFT already prevalent because it would not discourage HFTs to refrain from trading fast. Certainly, investigating this type of policy response, along with the potential means of enforcement, is an interesting topic for future research.

Another interesting extension to this paper would be to incorporate the sequential nature of HFT adoption into the real options model. In the current version of the model, the optimal stopping problem is solved such that decisions are made at date $t = 0$ for a given mass of HFT α. However, if this level of α were to change over time as more traders adopt the HFT technology, then the threshold of investment would be impacted. For example, while it may be optimal to invest immediately if the current level of HFT were to prevail indefinitely, if, however, the level of HFT were to be stochastic over time, then it may be optimal for the trader to postpone investing in the HFT now because it would not be justifiable in the long term if α were to change significantly. Such an extension will be considered in future research.

Appendix

A Derivation of $\Omega(r, \mu, \lambda)$ and proof that $\Omega(r, \mu, \lambda) > 0$

A.1 Derivation

From Eq. (9),

$$\Omega(r, \mu, \lambda) := \left(\frac{1}{\delta} - \lambda \int_0^{\infty} e^{-rt} \left((1-\lambda) + \lambda e^{\mu} \right)^t dt \right)$$

To prove my result in Eq. (10), it is sufficient to simply present my derivation of $\int_0^{\infty} e^{-rt} \left((1-\lambda) + \lambda e^{\mu} \right)^t dt$ as follows:

$$\begin{aligned} \int_0^{\infty} e^{-rt} \left((1-\lambda) + \lambda e^{\mu} \right)^t dt &= \int_0^{+\infty} e^{-t(r - \ln((1-\lambda)+\lambda e^{\mu}))} dt \\ &= \left[-\frac{1}{r - \ln\left((1-\lambda) + \lambda e^{\mu} \right)} e^{-t(r - \ln((1-\lambda)+\lambda e^{\mu}))} \right]_0^{+\infty}. \end{aligned}$$

Now

$$\lim_{t \to +\infty} e^{-t(r - \ln((1-\lambda)+\lambda e^{\mu}))} = 0$$

if and only if $r - \ln\left((1-\lambda) + \lambda e^{\mu} \right) > 0$. Hence, we must prove that $r - \ln\left((1-\lambda) + \lambda e^{\mu} \right) > 0$.

If $\lambda = 1$, then the latter expression becomes $r - \mu$ $(= \delta)$ which is positive by assumption. Therefore, if the expression decreases in λ, then it is positive everywhere. Indeed,

$$\frac{\partial}{\partial \lambda} \left(r - \ln\left((1-\lambda) + \lambda e^{\mu} \right) \right) = \frac{1 - e^{\mu}}{(1-\lambda) + \lambda e^{\mu}} < 0$$

implying $r - \ln\left((1-\lambda) + \lambda e^{\mu} \right) > 0$.

Therefore,

$$\int_0^\infty e^{-rt}\left((1-\lambda)+\lambda e^{\mu}\right)^t dt = \frac{1}{r-\ln\left((1-\lambda)+\lambda e^{\mu}\right)}.$$

A.2 Proof that $\Omega(r, \mu, \lambda) > 0$

$$\Omega(r, \mu, \lambda) = \frac{1}{\delta} - \frac{\lambda}{\left(r-\ln\left((1-\lambda)+\lambda e^{\mu}\right)\right)} > 0$$

If $\lambda = 1$, then $\Omega(\cdot) = 0$. Hence, if $\Omega(\cdot)$ decreases in λ, then $\Omega(\cdot)$ must be positive.

$$\frac{\partial\Omega(\cdot)}{\partial\lambda} < 0 \Longleftrightarrow r-\ln\left((1-\lambda)+\lambda e^{\mu}\right) + \frac{\lambda\left(e^{\mu}-1\right)}{(1-\lambda)+\lambda e^{\mu}} > 0.$$

This is true since $\mu > 0$ by assumption, and since $r-\ln\left((1-\lambda)+\lambda e^{\mu}\right) > 0$, which is proven above.

B Proof of Theorem 1

The derivation of the optimal threshold uses well-developed standard techniques from real options theory (see, for example, Dixit and Pindyck [9]).

Once the trader invests in the technology, he obtains a flow of profits of $X_t^F \equiv X_t$ in perpetuity. Thus, the net present value from investing, denoted by $V^A(X_0)$, is given by

$$V^A(X_0) = E^0\left[\int_0^\infty e^{-rt}X_t dt\right] - I = \frac{X_0}{\delta} - I. \tag{B.1}$$

Prior to investing, the trader obtains a flow of profits X^S, as well as having the option to invest. Letting $V^B(X_0)$ denote the current value to the trader before investing, and using standard dynamic programming arguments from Dixit and Pindyck [9], $V^B(X_0)$ solves the following Bellman equation:

$$\frac{1}{2}\sigma^2 X_0^2 (V^B)''(X_0) + (r-\delta)X_0(V^B)'(X_0) - rV^B(X_0) + X_0^S = 0, \tag{B.2}$$

where $(V^B)''(X_0) := \partial^2 V^B(\cdot)/\partial X_0^2$ and $(V^B)'(X_0) := \partial V^B(\cdot)/\partial X_0$.

The solution to this equation takes the following general form

$$V^B(X_0) = A_1 X_0^{\beta_1} + A_2 X_0^{\beta_2} + \frac{X_0^S}{\delta},$$

where A_1 and A_2 are constants to be determined, and β_1 and β_2 are the two (real) roots of the quadratic equation:

$$\frac{1}{2}\sigma^2\beta(\beta-1) + (r-\delta)\beta - r = 0.$$

Intuitively, the present value of trading for the slow trader ought to be comprised of the expected present value of his profit flow as if he were never to invest in the HFT technology, but adjusted for the fact that he has the option to invest in the technology should the expected payoff from investing become sufficiently large to warrant the cost of doing so (Dixit and Pindyck [9]). Hence, the general form of $V^B(X_0)$ should comply with this intuition. To this end the first two terms are the value of his option to invest, and the term X_0^S/δ represents the expected present value of his profit flow as if he were never to invest. Hence, it must be the case that

$$\begin{aligned}\frac{X_0^S}{\delta} \equiv& E^0\left[\int_0^\infty e^{-rt}X_t^S dt\right]\\ =&\lambda X_0 \int_0^\infty e^{-rt}\left((1-\lambda)+\lambda e^{r-\delta}\right)^t dt\\ =&\frac{\lambda X_0}{r-\ln\left((1-\lambda)+\lambda e^{r-\delta}\right)}\end{aligned} \tag{B.3}$$

(cf. Eq. (3) and Appendix A.1 for technical details.)

If the slow trader's current profit X_0^S is zero, it will stay at zero forever because X^S must follow a geometric Brownian motion, which has an absorbing barrier at zero, owing to its dependence on X (see Eq. (B.3)). Thus $X_0^S = 0$ iff $X_0 = 0$. However, since $X_0 \equiv X_0^F$, this implies that investing and becoming fast will never yield any profit either and, hence, that the value of the option to invest has no value. Therefore, the following condition must be satisfied at the boundary:

$$V^B(0) = 0.$$

But since $\beta_2 < 0$, this condition is satisfied iff $A_2 = 0$.

Therefore, a solution to Eq. (B.2) which satisfies to the boundary condition is given by

$$V^B(X_0) = A_1 X_0^{\beta_1} + \frac{\lambda X_0}{r-\ln\left((1-\lambda)+\lambda e^{r-\delta}\right)}. \tag{B.4}$$

Finally, there is a value of X at some time τ^*, which I denote by X^*, at which the trader is indifferent between investing and waiting. Moreover, at this threshold, the value functions before and and after investing should meet tangentially. In other words, the following conditions must be satisfied

$$V^B(X^*) = V^A(X^*)$$

and

$$(V^B)'(X^*) = (V^A)'(X^*),$$

where $(V^i)'(X^*) = \frac{\partial V^i(X_0)}{\partial X_0}\Big|_{X_0=X^*}$ (for $i = \{A, B\}$).

Together these these conditions give

$$X^* = \frac{\beta_1}{\beta_1 - 1}(\Omega(r,\delta,\lambda))^{-1} I, \tag{B.5}$$

and

$$A_1 = \left(\Omega(r,\delta,\lambda)X^* - I\right)(X^*)^{-\beta_1} \equiv (X^*)^{-\beta_1}F(X^*) \tag{B.6}$$

where

$$\Omega(r,\delta,\lambda) := \frac{1}{\delta} - \frac{\lambda}{r - \ln\left((1-\lambda) + \lambda e^{r-\delta}\right)}$$

and $F(X^*)$ is given by Eq. (9). Finally, since $X_0 := (1+\alpha)^{-1}x$, the result for the value function which solves the optimal stopping problem can be verified.[7] ■

C Proof of Proposition 2

$$X^* = \frac{\beta_1}{\beta_1 - 1}\left(\Omega(r,\delta,\lambda)\right)^{-1} I$$

where

$$\Omega(r,\delta,\lambda) = \frac{1}{\delta} - \frac{\lambda}{r - \ln\left((1-\lambda) + \lambda e^{r-\delta}\right)}.$$

(Hereafter I just write Ω in the interest of preserving space.)

Let

$$\mathcal{Q}(\beta_1) := \frac{1}{2}\sigma^2\beta_1(\beta_1 - 1) + (r-\delta)\beta_1 - r = 0.$$

Then for $\zeta \in \{r, \delta, \sigma\}$:

$$\frac{\partial X^*}{\partial \zeta} = -\frac{\beta_1}{\beta_1 - 1}\frac{1}{\Omega^2}\frac{\partial \Omega}{\partial \zeta} I - \frac{1}{(\beta_1 - 1)^2}\frac{\partial \beta_1}{\partial \zeta}\frac{I}{\Omega}. \tag{C.1}$$

The first term on the right hand side measures the sensitivity of the threshold with respect to the returns on the investment and is referred to as the present value effect. The second term measures the sensitivity of the threshold with respect to β_1 and, thus, measures the option effect. This decomposition is standard in real options models.

Now

$$\frac{\partial \beta_1}{\partial \zeta} = -\frac{\partial \mathcal{Q}(\beta_1)/\partial \zeta}{\partial \mathcal{Q}(\beta_1)/\partial \beta_1} = -\frac{\partial \mathcal{Q}(\beta_1)}{\partial \zeta}\frac{1}{\sigma^2\left(\beta_1 - \frac{1}{2}\right) + r - \delta}$$

(cf. Dixit and Pindyck [9] pp. 144) so that we can re-write Eq. (C.1) as follows:

$$\frac{\partial X^*}{\partial \zeta}\frac{\chi}{X^*} = -\chi\frac{\partial \Omega}{\partial \zeta} + \frac{\Omega^2}{\beta_1}\frac{\partial \mathcal{Q}(\beta_1)}{\partial \zeta}, \tag{C.2}$$

where

$$\chi := (\beta_1 - 1)\left(\sigma^2\left(\beta_1 - \frac{1}{2}\right) + r - \delta\right)\Omega^2.$$

From Eq. (C.2), since $\partial\Omega/\partial\sigma = 0$ and $\partial\mathcal{Q}(\beta_1)/\partial\sigma = \sigma\beta_1(\beta_1 - 1) > 0$,

$$\frac{\partial X^*}{\partial \sigma} > 0$$

owing solely to the option effect.

[7] Indeed, using the result in Dixit and Pindyck [9] pp. 315-316, we know that $E^0[e^{-r\tau}] = \left(\frac{X_0}{X^*}\right)^{\beta_1}$, where τ is the time of investment. Therefore, $(X^*)^{-\beta_1} = (X_0)^{-\beta_1}E^0[e^{-r\tau}] = ((1+\alpha)^{-1}x)^{-\beta_1}E^0[e^{-r\tau}]$. Thus, we can write A_1 as $A_1 = E^0[e^{-r\tau}]\left((1+\alpha)^{-1}x\right)^{-\beta_1}F(X^*)$. In other words, the value of the option to invest is just the expected discounted value of the net present value at the time investment takes place, which is intuitive.

Also, from Eq. (C.2), we see that

$$\frac{\partial X^*}{\partial r} > 0 \Longleftrightarrow \frac{\Omega^2}{\beta_1}(\beta_1 - 1) > \chi \frac{\partial \Omega}{\partial r} \tag{C.3}$$

where the right hand side of the second inequality in (C.3) gives the present value effect and the left hand side the option effect. Moreover,

$$\frac{\partial \Omega}{\partial r} = \frac{\lambda(1-\lambda)}{\left(r - \ln\left((1-\lambda) + \lambda e^{r-\delta}\right)\right)^2 \left(1 - \lambda + \lambda e^{r-\delta}\right)} > 0. \tag{C.4}$$

If Ω is large, then the option effect will dominate and therefore drive the effect of r on X^*. In this case, X^* will increase in r.

However, if λ is high, and Ω is therefore small, then we have that X^* will be impacted by r owing to the PV effect because Ω increases in r. In this case, X^* will decrease in r.

$$\frac{\partial X^*}{\partial \delta} > 0 \Longleftrightarrow \chi \frac{\partial \Omega}{\partial \delta} + \Omega^2 < 0 \tag{C.5}$$

where

$$\frac{\partial \Omega}{\partial \delta} = -\frac{1}{\delta^2} + \frac{\lambda^2 e^{r-\delta}}{\left(r - \ln\left((1-\lambda) + \lambda e^{r-\delta}\right)\right)^2 \left(1 - \lambda + \lambda e^{r-\delta}\right)} < 0 \tag{C.6}$$

(since $\partial\Omega/\partial\delta$ is continuous in λ and $\partial\Omega/\partial\delta = 0$ for $\lambda = 1$ and $\partial\Omega/\partial\delta = -1/\delta^2 < 0$ for $\lambda = 0$).

As in the case of r, there are two effects at play. Specifically, if Ω is large, the option effect will dominate the PV effect of δ on X^*. In this is the case, X^* will decrease in δ. On the other hand, if Ω is small, the present value effect will dominate and X^* will increase in δ.

D Proof of Proposition 3

$$\frac{\partial \alpha^*}{\partial \lambda} = \frac{\partial \alpha^*}{\partial X^*} \frac{\partial X^*}{\partial \lambda}.$$

We know from Proposition 1 that $\partial X^*/\partial \lambda > 0$. Therefore

$$\frac{\partial \alpha^*}{\partial \lambda} > 0 \Longleftrightarrow \frac{\partial \alpha^*}{\partial X^*} > 0.$$

From the equilibrium condition given by Eq. (14)

$$\begin{aligned} \frac{\partial \alpha^*}{\partial X^*} &= -\frac{\partial[(1+\alpha^*)X^*]}{\partial X^*} G'((1+\alpha^*)X^*) \\ &= -\left[(1+\alpha^*) + X^* \frac{\partial \alpha^*}{\partial X^*}\right] G'((1+\alpha^*)X^*) \end{aligned} \tag{D.1}$$

Therefore

$$\frac{\partial \alpha^*}{\partial X^*} = -\frac{(1+\alpha^*)G'((1+\alpha^*)X^*)}{1 + X^* G'((1+\alpha^*)X^*)} < 0 \tag{D.2}$$

since $G'((1+\alpha^*)X^*) > 0$. Therefore, α^* decreases in λ.

E The Effect of α on $V^f(\alpha)$, $V^s(\alpha)$, and $\Delta(\alpha)$

First I use Leibniz rule for differentiating integrals to show that $\Delta(\alpha)$ decreases in α.

$$\begin{aligned}
\frac{\partial \Delta(\alpha)}{\partial \alpha} =& \frac{\partial V^f(\alpha)}{\partial \alpha} - \frac{\partial V^s(\alpha)}{\partial \alpha} \\
=& -\frac{1}{\delta(1+\alpha)^2}\int_{(1+\alpha)X^*}^{\bar{x}} x^i g^f(x^i)dx^i - \frac{(X^*)^2}{\delta} g^f((1+\alpha)X^*) \\
& - \Big[-\frac{\lambda}{(1+\alpha)^2}\frac{1}{(r-\ln((1-\lambda)+\lambda e^{r-\delta}))}\int_0^{(1+\alpha)X^*} x^i g^s(x^i)dx^i \\
& + \frac{\lambda(X^*)^2}{(r-\ln((1-\lambda)+\lambda e^{r-\delta}))} g^s((1+\alpha)X^*) \\
& - \beta_1(1+\alpha)^{-\beta_1-1}(X^*)^{-\beta_1}F(X^*)\int_0^{(1+\alpha)X^*}(x^i)^{\beta_1}g^s(x^i)dx^i \\
& + X^*F(X^*)g^s((1+\alpha)X^*)\Big],
\end{aligned} \tag{E.1}$$

where the term in the square brackets is $\partial V^s(\alpha)/\partial\alpha$.

It is clear from the equation that $\partial V^f(\alpha)/\partial\alpha < 0$. Then it must be the case that $\partial\Delta(\alpha)/\partial\alpha < 0$ if $\partial V^s(\alpha)/\partial\alpha > 0$.

$\partial V^s(\alpha)/\partial\alpha > 0$ if

$$(X^*)^2 g^s((1+\alpha)X^*) > \frac{1}{(1+\alpha)^2}\int_0^{(1+\alpha)X^*} x^i g^s(x^i)dx^i \tag{E.2}$$

and if

$$X^* g^s((1+\alpha)X^*) > \beta_1(1+\alpha)^{-\beta_1-1}(X^*)^{-\beta_1}\int_0^{(1+\alpha)X^*}(x^i)^{\beta_1}g^s(x^i)dx^i. \tag{E.3}$$

Using the integration by parts technique to evaluate the integrals in the latter two equations, it is sufficient to approximate the integrals using only the "uv" term in the standard formula to show the conditions hold if we assume that $\frac{\partial g^s(x^i)}{\partial x^i} \geq 0$. This is because the "$\int vdu$" term is positive in that case and, hence, $\int udv < uv$. Indeed, there are many specifications of $g^s(\cdot)$ such that the derivative is nonnegative. For example, assuming the x^i's follow a uniform distribution, $g^s(x^i) = \frac{1}{(1+\alpha)X^*}$, the derivative of which is zero. Letting $u = g^s(x^i)$ and $dv = x^i dx^i$ or $dv = (x^i)^{\beta_1}dx^i$, then $[uv]_0^{(1+\alpha)X^*} = \frac{1}{2}(1+\alpha)^2(X^*)^2 g^s((1+\alpha)X^*)$ for Eq. (E.2), and $[uv]_0^{(1+\alpha)X^*} = \frac{1}{\beta_1+1}(1+\alpha)^{\beta_1+1}(X^*)^{\beta_1+1}g^s((1+\alpha)X^*)$ for (E.3).

Eq. (E.2) becomes

$$(X^*)^2 g^s((1+\alpha)X^*) > \frac{1}{2}(X^*)^2 g^s((1+\alpha)X^*), \tag{E.4}$$

which clearly holds, and Eq. (E.3) becomes

$$X^* g^s((1+\alpha)X^*) > \frac{\beta_1}{\beta_1+1}X^* g^s((1+\alpha)X^*), \tag{E.5}$$

which also clearly holds.

Therefore, $\partial V^s(\alpha)/\partial\alpha > 0$ and $\Delta(\alpha)$ decreases in α everywhere. ■

References

[1] S. Banerjee, U. Gucbilmez, and G. Pawlina. Optimal exercise of jointly held real options: A Nash bargaining approach with value diversion. *European Journal of Operational Research*, 239:565–578, 2014.

[2] M.D. Baron, J. Brogaard, B. Hagströmer, and A. Kirilenko. Risk and return in high frequency trading. Available at SSRN: `http://ssrn.com/abstract_id=1859265`, 2016.

[3] A. Battauz, M. De Donno, and A. Sbuelz. Real options and American derivatives: The double continuation region. *Management Science*, 61(5):1094–1107, 2015.

[4] B. Biais, T. Foucault, and S. Moinas. Equilibrium fast trading. *Journal of Financial Economics*, 116(2):292–313, 2015.

[5] J. Brogaard, T. Hendershott, and R. Riordan. High frequency trading and price discovery. *Review of Financial Studies*, 27(8):2267–2306, 2014.

[6] E. Budish, P. Cramton, and J. Shim. The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics*, 130(4): 1547–1621, 2015.

[7] C. Chiyachantana and P. Jain. Institutional trading frictions. Unpublished working paper, University of Memphis, 2009.

[8] L. Delaney and J.J.J. Thijssen. The impact of voluntary disclosure on a firm's investment policy. *European Journal of Operational Research*, 242:232–242, 2015.

[9] A.K. Dixit and R.S. Pindyck. *Investment under uncertainty.* Princeton University Press, Princeton, NJ, 1994.

[10] J. Dow. Is liquidity self-fulfilling? *Journal of Business*, 77:895–908, 2005.

[11] T. Foucault. *Market Microstructure: Confronting many viewpoints.* John Wiley & Sons, Chichester, UK, 2012.

[12] V. Glode, R. Green, and R. Lowery. Financial expertise as an arms race. *Journal of Finance*, 67:1723–1759, 2012.

[13] L. Glosten and P. Milgrom. Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1):71–100, 1985.

[14] J. Hasbrouck and G. Saar. Low latency trading. *Journal of Financial Markets*, 16(4): 646–679, 2013.

[15] T. Hendershott and R. Riordan. Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis*, 48(4):1001–1024, 2013.

[16] T. Hendershott, C. Jones, and A. Menkveld. Does algorithmic trading improve liquidity? *Journal of Finance*, 66(1):1–33, 2011.

[17] P. Hoffman. A dynamic limit order market with fast and slow traders. *Journal of Financial Economics*, 113:156–169, 2014.

[18] R. McDonald and D.R. Siegal. The value of waiting to invest. *Quarterly Journal of Economics*, 101:707–728, 1986.

[19] A. Menkveld. High frequency trading and the new market makers. *Journal of Financial Markets*, 16(4):712–740, 2013.

[20] J.I. Munoz, J. Contreras, J. Caamano, and P.F. Correia. A decision making tool for project investments based on real options: The case of wind power generation. *Annals of Operations Research*, 186:465–490, 2011.

[21] M. O'Hara. High frequency market microstructure. *Journal of Financial Economics*, 116(2):257–270, 2015.

[22] M. O'Hara and M. Ye. Is market fragmentation harming market quality? *Journal of Financial Economics*, 100:459–474, 2011.

[23] E. Pagnotta and T. Philippon. Competing on speed. Unpublished working paper, Singapore Management University, University of Memphis, 2015.