Bank of England

Price formation in markets with trading delays

Staff Working Paper No. 1,023
April 2023

Gabor Pinter and Semih Üslü

Bank of England

Staff Working Paper No. 1,023

Price formation in markets with trading delays

Gabor Pinter[1] and Semih Üslü[2]

Abstract

We develop a parsimonious price formation model to study information aggregation and information acquisition in the presence of trading delays. If delays apply uniformly to uninformed and informed traders, the level of delays does not affect information aggregation. Traders' information acquisition incentives are, however, weaker in a market with longer delays. Therefore, the equilibrium fraction of informed traders is lower if delays are longer, establishing an inverse relationship between trading delays and price informativeness. We also show that risk premia and price dispersion tend to be non-monotonic functions of the level of delays when information acquisition is endogenous. We document novel empirical evidence from the UK corporate bond market, which largely corroborates the implications of our theory.

Key words: Trading frictions, trading delays, price informativeness, information aggregation, information acquisition, liquidity.

JEL classification: D49, D53, D82, D83, G11, G12, G14.

---

(1) Bank of England. Email: gabor.pinter@bankofengland.co.uk
(2) Johns Hopkins Carey. Email: semihuslu@jhu.edu

The views expressed in this paper are those of the authors, and not necessarily those of the Bank of England or its committees. An earlier version of this article was circulated under the title 'Informed trading in markets with trading delays'. We would like to thank, for helpful comments and suggestions Bruno Biais as well as Daniel Andrei, Simon Board, Will Cong, Adrien d'Avernas, Nicolae Gearleanu, Ben Lester, Wei Li, David McAdams, Tomasz Sadzik, Zhaogang Song, Aleh Tsyvinski, Laura Veldkamp, Guner Velioglu, Pierre-Olivier Weill, and the participants at the UCLA Theory Proseminar.

The Bank's working paper series can be found at www.bankofengland.co.uk/working-paper/staff-working-papers

Bank of England, Threadneedle Street, London, EC2R 8AH
Email: enquiries@bankofengland.co.uk

ISSN 1749-9135 (on-line)

1 Introduction

Execution quality is important to traders in financial markets. A key dimension of execution quality is the speed at which their orders are executed. This speed varies widely across markets. Pagnotta and Philippon (2018) state that more than half of the US equity orders take five seconds to ten minutes to fully execute. On the other hand, the same number was identified to be 1-3 days in the corporate bond market (Duffie, Gârleanu, and Pedersen, 2007, Pagnotta and Philippon, 2018, and Kargar, Lester, Plante, and Weill, 2022) and five days in the municipal bond market (Hugonnier, Lester, and Weill, 2020). How do these trading delays, small or drastic, affect price discovery? This is the main question we analyze in our paper.

We develop a competitive asset trading model that nests Grossman and Stiglitz (1980) as a special case. By trading in a market without trading delays, traders in Grossman and Stiglitz (1980) fully benefit from any private information they possess, while traders in our model trade in a market with trading delays, in which private information progressively becomes public. In two special cases, the trading environment of our model coincides with the Grossman and Stiglitz (1980) environment: if private information is infinitely-lived (i.e., the noise of public information revelation goes to infinity) or if trading delays disappear (i.e., if trading speed goes to infinity). Hence, apart from these two special cases, trading delays and the lifespan of private information are expected to affect all the equilibrium outcomes such as risk premia, price dispersion, and measures of price informativeness.

To capture trading delays in a meaningful way, we set up a dynamic trading game. Two types of traders, informed and uninformed, gain one-time access to a competitive market at i.i.d. stochastic times. Importantly, we assume that trading delays apply uniformly to uninformed and informed traders. Conditional on market access, a trader trades to obtain her optimal position given all the information available to her at the time of her market access. Interestingly, while the market price becomes a better forecast of the asset value as time passes because private information progressively becomes public, we show that the measure of unique information the market price contains, which we term *revelatory price efficiency*, is stationary and coincides with what obtains in the equilibrium of Grossman and Stiglitz (1980) model. This leads to a conditional irrelevance result: By taking as given the fraction of informed traders, revelatory price efficiency (RPE) is independent of the level of trading delays.

When traders are allowed to make *ex ante* investment in private information by anticipating the outcome of the dynamic trading game, however, the irrelevance breaks down. More specifically, traders' information acquisition incentives are weaker if trading delays are longer (or if private information is shorter-lived). Therefore, the equilibrium fraction of informed traders is lower if trading delays are longer, establishing an inverse relationship between trading delays and RPE. This inverse relationship between trading delays and price informativeness is the first implication of our theoretical model that we test in the empirical part of the paper.

Second, our theory has novel implications for price dispersion, which refers to the second moment of transaction prices that arise during the trading session. In our baseline model without endogenous information acquisition, we show that price dispersion is non-monotonic in trading delays. The key determinant behind this result is how trading delays affect the distribution of traders' trading times. If trading delays are small, both type of traders, uninformed and informed, can trade very early on. Like in the original Grossman and Stiglitz (1980) model, the distribution of information that traders possess is close to bimodal. As trading delays get more severe, traders access the market later and in more dispersed times, which means the distribution of information that (uninformed) traders possess is more dispersed. This, in turn, leads to larger price dispersion. As trading delays get even more severe, however, the market access of this uninformed crowd of traders gets extremely delayed, making their valuations closer together due to a high level of information revelation by then. This reduces price dispersion. Overall, price dispersion turns out to be a hump-shaped function of trading delays.

Using parametric examples, we also confirm that this non-monotonicity is preserved when information acquisition is endogenous. More specifically, we find that price dispersion and average risk premium tend to reach their maximum when trading delays are just severe enough to preclude all information acquisition in equilibrium. At this level of trading delays, delays are severe enough to keep all traders uninformed, but also moderate enough to allow these uninformed traders to trade before a meaningful revelation of public information. Thus, the transaction prices resulting from these uninformed trades exhibit a large risk premia. Similarly, prices exhibit high dispersion because of the high flow of uninformed orders getting executed. Shorter trading delays would reduce the fraction of uninformed traders, and so, reduce the sensitivity of the market price to the fluctuations in the public information process, thereby reducing price dispersion. Longer trading delays would delay the market access of this entirely uninformed crowd of traders, making their valuations closer together when they access the market, and in turn, reducing price dispersion.

Third, we discuss how to view the stylized facts regarding price informativeness in the existing empirical literature through the lens of our model. To offer theoretical insights into the growth of "big data" financial technology over the last decade, we ask how RPE changes in response to traders' easier access to better data technologies. We contrast two possible scenarios: a decline in the information acquisition cost vs. an increase in the rate of information depreciation due to the technology allowing traders to extract others' information. Our model predicts that while RPE increases in the former scenario, it decreases in the latter. Since both effects are, in principle, present for any asset albeit with different strengths, our model implies that RPE will increase for some assets and decrease for others. Bai, Philippon, and Savov (2016) find a secular decline in stock price informativeness for the universe of all firms, while they find a secular increase for the S&P500 firms. Our model rationalizes their findings if the improvement in RPE from the decline in the information acquisition cost dominates (resp. is dominated by) the decline in RPE from the increase in the rate of information depreciation for the S&P500 firms (resp. the universe of all firms).[1]

In the last part of the paper, we provide novel empirical evidence for some of the theoretical implications unique to our model. To this end, we use a transaction-level dataset which covers close to the universe of all secondary market trades in the UK corporate bond market. With its over-the-counter (OTC) structure, the corporate bond market is a textbook example of markets with trading delays.[2] We proxy (the inverse of) a bond's trading delay with its trade frequency measured in the data. We, then, utilize the cross-sectional variation in bonds' trade frequency to test two empirical hypotheses generated by our theoretical model: (i) *ceteris paribus*, price efficiency is decreasing in trading delays and (ii) *ceteris paribus*, the dispersion of transaction prices is non-monotonic in trading delays. Overall, our results from empirical tests are consistent with the implications of the theoretical model. We interpret this as pointing to the importance of taking trading delays into account when analyzing the informational inefficiency and liquidity costs in asset trading.

Next, we briefly review the related literature and our paper's contribution. Section 2 lays out the main model environment, while Section 3 defines and characterizes a noisy rational expectations equilibrium for this environment. Section 4 endogenizes the traders' information acquisition decisions and the informativeness of the market price. Section 5 analyzes the extent to which our theoretical findings are consistent with the corporate bond trading patterns in practice. Section 6 concludes.

[1] Bai, Philippon, and Savov (2016) find that the decline in market efficiency for the universe of all firms is highest during the very last period (2010-2014) of their dataset, when the big data arguably plays the most prominent role.

[2] Pintér and Üslü (2022) propose a structural estimation of a model of OTC asset markets by using the same dataset. Their estimation implies that the median client faces a trading delay of around three quarters of a trading day in the most liquid segment of the UK corporate bond market with 57 bonds.

1.1 Related literature

The first strand of literature to which our paper contributes is the voluminous literature on information acquisition with papers by Grossman and Stiglitz (1980), Verrecchia (1982), Admati and Pfleiderer (1986), Veldkamp (2006a), Veldkamp (2006b), Cespa (2008), Lee (2013), and Dugast and Foucault (2018), among others. See Veldkamp (2011) for a survey of this literature. These papers abstract away from trading frictions and trading delays and focus on the details of assets' payoff and information structures. We keep the payoff and information structure simple while allowing for rich dynamics in the equilibrium price process, and model explicitly the trading delays that are characteristic of many real-world financial markets, especially the illiquid ones like the markets for corporate and municipal bonds.

The second strand of literature to which our paper contributes is the fast-growing literature on trading delays with papers by Duffie, Gârleanu, and Pedersen (2005), Weill (2007), Gârleanu (2009), Lagos and Rocheteau (2009), Afonso and Lagos (2015), Farboodi, Jarosch, and Shimer (2015), Pagnotta and Philippon (2018), Üslü (2019), and Hugonnier, Lester, and Weill (2020), among others. See Weill (2020) for a recent survey. These papers abstract away from private information and focus exclusively on trading frictions and delays. We borrow the details of trading delays from this literature. In particular, traders in our model gain access to a market with Walrasian pricing mechanism at idiosyncratic random times as in Gârleanu (2009) and Pagnotta and Philippon (2018). The same mechanism obtains in Lagos and Rocheteau (2009) as well, when dealers' bargaining power is set to be zero. Models in this literature typically obtain that illiquidity premium and price dispersion are monotone increasing in trading delays. The novelty of our model is to introduce private information about common asset value to these papers' trading environment with delays, and we show that illiquidity premium and price dispersion can be non-monotonic functions of the delay.

The combination of trading delays and private information about common asset value is also present in a relatively limited set of papers by Wolinsky (1990), Blouin and Serrano (2001), Duffie, Malamud, and Manso (2014), Golosov, Lorenzoni, and Tsyvinski (2014), Dang and Morath (2015), Lauermann and Wolinsky (2016), Lauermann, Merzyn, and Virág (2017), and Lester, Shourideh, Venkateswaran, and Zetlin-Jones (2018). With the exception of Duffie, Malamud, and Manso (2014) and Dang and Morath (2015), these papers do not study information acquisition. What distinguishes our model from Duffie, Malamud, and Manso (2014) and Dang and Morath (2015) is mainly the trading protocol (and implicitly the price setting mechanism). While prices that aggregate information are privately negotiated prices in bilateral meetings in

Duffie, Malamud, and Manso (2014) and Dang and Morath (2015), the Walrasian price in our model acts as a public signal as in the extant literature on information acquisition following Grossman and Stiglitz (1980).

Traders' desire to take advantage of short-lived private information in our model is in common with the recent literature that analyzes high-frequency trading. Examples include Aït-Sahalia and Sağlam (2013), Budish, Cramton, and Shim (2015), Biais, Foucault, and Moinas (2015), Foucault, Hombert, and Roşu (2016), and Li (2018). While our paper's main focus is on trading delays in illiquid markets, these papers model high-speed trading in liquid markets. Accordingly, they end up using very different sets of modeling assumptions than our model's assumptions. Information aggregation or revelation by market prices is also analyzed together with different types of frictions, e.g., with technological trading costs by Subrahmanyam (1988), Dow and Rahi (2000), Vives (2017), and Dávila and Parlatore (2020), with sticky relationships by Babus and Kondor (2018) and Li and Song (2018), and with long intermediation chains by Glode and Opp (2016) and Glode, Opp, and Zhang (2019).

Our empirical analysis in Section 5 relates to two mains strands of the empirical literature. First, we draw on previous studies on the determinants of various trade cost measures in OTC markets including price dispersion. See Garbade and Silber (1976), Edwards, Harris, and Piwowar (2007), Jankowitsch, Nashikkar, and Subrahmanyam (2011), Di Maggio, Kermani, and Song (2017), O'Hara, Wang, and Zhou (2018), Li and Schürhoff (2019), Üslü and Velioğlu (2019), Dick-Nielsen, Poulsen, and Rehman (2021), and Pintér and Üslü (2022), among others. To our knowledge, we are the first to show that price dispersion is non-monotonic in trading delays in the empirical literature as well as in the theoretical literature. Second, we draw on previous studies on the determinants of price efficiency. See Hotchkiss and Ronen (2002), Hendershott and Jones (2005), Hou and Moskowitz (2005), Bai, Philippon, and Savov (2016), Lee, Naranjo, and Velioğlu (2018), and Indriawan, Pascual, and Shkilko (2022), among others. To our knowledge, we are the first to analyze trading delays as a determinant of price efficiency in the cross section of corporate bonds.

2 The model economy

We build a continuous-time model economy with an infinite time horizon: $t \in [0, \infty)$. We fix a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \Pr)$, on which is defined a standard Brownian motion, B. The economy is populated by a unit-mass continuum of traders, who have constant-absolute-risk-aversion (CARA) preferences over horizon wealth W_∞ with a common risk aversion coefficient of $\gamma > 0$. Traders are infinitely patient (i.e., their discount rate is zero) and are endowed with a technology that instantly produces cash at unit marginal cost in order to make side payments.[3] We normalize the traders' initial wealth to zero, $W_0 = 0$, because their initial wealth will not affect decision making, provided that they can produce cash at unit marginal cost.

A fraction $\alpha \in (0, 1)$ of the traders are informed and the remaining $1 - \alpha$ uninformed. There is one risky asset with payoff realized at date $t = \infty$. The payoff of this asset, denoted by v, consists of two parts,

$$v = \theta + \varepsilon,$$

where θ is observable to informed traders but ε is not observable to anyone. Random variable θ is characterized by a normal distribution with mean $\hat{\theta}_0$ and variance $1/S_\theta$. Random variable ε also follows a normal distribution with mean 0 and variance $1/S_\varepsilon$. These random variables, θ and ε, are independent.

In addition to being informed or uninformed, traders are heterogeneous with respect to their initial endowment of the asset. Immediately prior to time $t = 0$, half of traders are each endowed with a_1 shares of the risky asset, and the other half a_2, where a_1 and a_2 are identically and normally distributed with mean 0 and variance $4/S_A$. All these random variables, θ, ε, a_1, and a_2, are pairwise independent.

We denote with $N_t^j(\cdot)$ the cdf of traders' asset positions at time t conditional on their information type $j \in \mathcal{J} \equiv \{\text{inf}, \text{uninf}\}$. Thus, at $t = 0$:

$$\int_{\mathbb{R}} a \, dN_0^j(a) = \frac{a_1 + a_2}{2}$$

for any $j \in \mathcal{J}$.

At date 0, informed traders obtain a private signal which reveals the true value of θ. For $t > 0$, there is a "flow" of public signals Z_t, progressively reducing the informational edge of informed traders:

$$dZ_t = \theta dt + \sigma dB_t, \tag{1}$$

[3] Postponing all payoffs and consumption to an infinitely far horizon and assuming that traders are infinitely patient are modeling tricks to obtain simple formulas by eliminating the pure time cost of trading delays. These are not essential for our analysis because we want to study the effect of trading delays on price discovery in a model with common valuations. The fast-growing search literature, instead, studies the effect of the time cost of delaying trade on welfare when traders have private valuations. Thus, assuming proper discounting is essential for those models because in the limit as the discount rate approaches zero, the competitive outcomes typically obtain in those models. See Duffie, Gârleanu, and Pedersen (2005), Weill (2020), and Pintér and Üslü (2022), among others.

where B is a standard Brownian motion and $Z_0 = \hat{\theta}_0$. The σ-algebra $\mathcal{F}_t$ is generated by B. Depending on the realization of θ, a_1, and a_2, informed and uninformed traders will have different beliefs about the asset payoff. We will discuss this when we define the equilibrium.

In the spirit of Gârleanu (2009), traders have infrequent access to a continuous competitive market in which they can trade as price takers and without any friction. More specifically, we assume that trader $i \in [0, 1]$ gains a one-time access to the competitive market at date $t = \tau(i)$, where $\tau(i)$ is an exponentially distributed random variable with mean $1/\lambda$. That is, trader i can rebalance her asset position only at the instant $t = \tau(i)$. We also assume the market access times, $\tau(i)$s, are i.i.d. across traders. Conditional on market access at date t, a trader with endowment a and information type j buys $q(a, j, t)$ units of the asset at the market-clearing price P_t. Naturally, if $q(a, j, t) < 0$, it means the trader sells $-q(a, j, t)$ units.

3 Equilibrium

In illiquid market models with homogeneous information sets, such as Gârleanu (2009), Lagos and Rocheteau (2009), and Pagnotta and Philippon (2018), the market price has only the role of clearing the market. Differently from these papers, in our model, traders have different information sets about the common asset value, which gives the equilibrium price two simultaneous roles, clearing the market and revealing information, as in Grossman and Stiglitz (1980).

3.1 Definition

Because the trading horizon is infinite, all traders eventually end up trading with probability one. However, there is uncertainty as to when a trader will gain access to the market. A trader of type j with endowment a who gains access to the market at time t buys

$$q(a, j, t) = \arg\max_q \mathbb{E}\left[-e^{-\gamma\{v(a+q)-P_t q\}} \middle| \mathcal{F}_t^{j,a}\right] \tag{2}$$

shares of the risky asset, where the expectation operator, $\mathbb{E}$, is over the random variable v, and $\mathcal{F}_t^{j,a} \supset \mathcal{F}_t$ denotes the information a type-(j, a) trader gathered up to time t.

Because $N_t^j(\cdot)$ is a conditional cdf, it satisfies

$$\int_{\mathbb{R}} dN_t^j(a) = 1 \tag{3}$$

for all $t \geq 0$ and $j \in \mathcal{J}$. By the law of large numbers, this cdf evolves according to the following

Kolmogorov forward equation:

$$\frac{\partial N_t^j(a)}{\partial t} = -\lambda e^{-\lambda t} N_0^j(a) + \lambda \int_{\mathbb{R}} \mathbb{I}_{\{a'+q(a',j,t)\leq a\}} e^{-\lambda t} dN_0^j(a') \tag{4}$$

for all $t \geq 0$ and $(j, a) \in \mathcal{J} \times \mathbb{R}$. The first and second terms on the RHS of Equation (4) represent respectively the gross outflow from and gross inflow to the mass of j-type traders with an asset holding less than or equal to a. Note that these outflows and inflows can only come from the mass of traders who have not gained access to the market yet. The mass, $e^{-\lambda t} N_0^j(a)$, of j-type traders with an asset holding less than or equal to a who are yet to trade obtains by solving

$$\frac{d}{dt} f_t = -\lambda f_t, \quad f_0 = N_0^j(a).$$

These traders gain access to the market at the Poisson rate λ, which implies a rate of (gross) outflow $\lambda e^{-\lambda t} N_0^j(a)$. The second term means any j-type trader can contribute to the (gross) inflow to the mass $N_t^j(a)$, if she gains access to the market, which happens at rate λ, and if her post-trade asset holding is less than or equal to a.

Finally, the market-clearing condition must ensure that the traders' aggregated demand nets out to zero, at each date $t \geq 0$:

$$\lambda \int_{\mathbb{R}} q(a, \text{inf}, t)\, \alpha\, e^{-\lambda t} dN_0^{\text{inf}}(a) + \lambda \int_{\mathbb{R}} q(a, \text{uninf}, t)\, (1-\alpha)\, e^{-\lambda t} dN_0^{\text{uninf}}(a) = 0, \tag{5}$$

where $\alpha \in (0, 1)$ is the fraction of informed traders.

We have specified all requirements to define the equilibrium. We only have to bring those ingredients together. Formally, a rational expectations equilibrium (REE) is defined as follows.

Definition 1. *A rational expectations equilibrium is a list composed of (i) a trade size function $q(a, j, t)$, (ii) a market-clearing price P_t, (iii) a perceived price function $\tilde{P}_t(\{Z_s\}_{s\in[0,t]}, \theta, a_1, a_2)$, and (iv) a set of distributions for traders' asset positions $N_t^j(a)$ such that*

- *Given (ii) and (iii), (i) solves (2);*
- *Given (i), (iv) satisfies (3) and (4);*
- *Given (i), (ii) clears the market through (5);*
- *Expectations are correct: $\tilde{P}_t(\{Z_s\}_{s\in[0,t]}, \theta, a_1, a_2) = P_t$.*

3.2 Characterization

Let $\mu_t^{j,a} = \mathbb{E}\left[v \,|\, \mathcal{F}_t^{j,a}\right]$ and $S_t^j = \left(\mathbb{V}\left[v \,|\, \mathcal{F}_t^{j,a_1}\right]\right)^{-1} = \left(\mathbb{V}\left[v \,|\, \mathcal{F}_t^{j,a_2}\right]\right)^{-1}$, where $\mathbb{V}$ denotes the variance operator.[4] Assuming that v is *conditionally* normally distributed,[5] (2) can be re-written as

$$q(a,j,t) = \arg\max_q \, e^{-\gamma\left\{-P_t q + (a+q)\mu_t^{j,a} - \frac{\gamma}{2}\frac{(a+q)^2}{S_t^j}\right\}}.$$

Using the FOC and SOC of this unconstrained optimization problem, we obtain the following lemma.

Lemma 1. *Fix a market price P_t and a collection $\left\{\mathcal{F}_t^{j,a}\right\}_{(j,a)\in\mathcal{J}\times\{a_1,a_2\}}$ for the traders' information sets. Assume that v is normally distributed conditional on any of the given information sets. Then, a type-j trader with endowment $a \in \{a_1, a_2\}$ who gains access to the market at time t chooses her optimal trade quantity as*

$$q\,(a,j,t) = \frac{\left(\mu_t^{j,a} - P_t\right) S_t^j}{\gamma} - a, \tag{6}$$

where $\mu_t^{j,a} \equiv \mathbb{E}\left[v \,|\, \mathcal{F}_t^{j,a}\right]$ and $S_t^j \equiv \left(\mathbb{V}\left[v \,|\, \mathcal{F}_t^{j,a_1}\right]\right)^{-1} = \left(\mathbb{V}\left[v \,|\, \mathcal{F}_t^{j,a_2}\right]\right)^{-1}$.

Next, we endogenize the traders' information sets and the market price to complete the characterization of the general equilibrium of the trading game. For uninformed traders, there are two sources of public information to gauge the value of the risky asset: The market price P_t and the public signal process Z_t. First, we derive the mean and variance of θ conditional on observing the public signal process Z_s up to time $s = t$. This mean and variance will be later used in determining the posterior mean and variance of the asset payoff v when an uninformed trader gains access to the market.

An application of Kalman-Bucy filter to (1) implies that

$$\mathbb{E}\left[\theta \,|\, \mathcal{F}_t\right] = \theta + \frac{z}{\sqrt{S_\theta + t/\sigma^2}} \tag{7}$$

and

$$\left(\mathbb{V}\left[\theta \,|\, \mathcal{F}_t\right]\right)^{-1} = S_\theta + t/\sigma^2, \tag{8}$$

[4] $\mathbb{V}\left[v \,|\, \mathcal{F}_t^{j,a_1}\right] = \mathbb{V}\left[v \,|\, \mathcal{F}_t^{j,a_2}\right]$ follows because a_1 and a_2 are i.i.d. random variables. This will become clearer shortly when we endogenize the traders' information sets and derive these posterior variances in closed form.

[5] This assumption is verified ex post. The verification follows because (i) all signals are Gaussian and (ii) the market price, which is the endogenous public signal in the model, is a linear function of Gaussian random variables as is typically assumed in the literature.

where z is a random variable with the standard normal distribution.[6] As time passes, observing the public signal process Z_t for longer makes the conditional mean (7) closer to the informed traders' signal θ, and so, the conditional precision (8) is an increasing function of time, t.

The second source of public signal for uninformed traders is the market price, P_t. As is standard in the literature, we restrict our attention to linear pricing functions:

$$\tilde{P}_t\left(\{Z_s\}_{s\in[0,t]}, \theta, a_1, a_2\right) = \beta_1(t)\,\hat{\theta}_t + \beta_2(t)\,\theta + \beta_3(t)\,(a_1 + a_2)\,, \tag{9}$$

where for $k \in \{1,2,3\}$, $\beta_k : [0,\infty) \to \mathbb{R}$ is a function of time for which traders have perfect foresight and $\hat{\theta}_t \equiv \mathbb{E}[\theta \,|\, \mathcal{F}_t]$ is the posterior mean of the information learned by observing $\{Z_s\}_{s\in[0,t]}$. Naturally, β_1 and β_2 are positive-valued, while β_3 is negative-valued. The only component of the price which is informative about the asset's fundamental value v and not common knowledge among traders is θ. Thus, for an uninformed trader, price is a noisy signal about θ. However, an informed trader already knows θ, and so, observing the market-clearing price, $P_t = \tilde{P}_t\left(\{Z_s\}_{s\in[0,t]}, \theta, a_1, a_2\right)$, reveals a_2 (resp. a_1) to her if her endowment is a_1 (resp. a_2). Therefore, by replacing $\{Z_s\}_{s\in[0,t]}$ with its sufficient statistic $\hat{\theta}_t$ inside the linear pricing function (9), the information sets of an informed trader and of an uninformed trader are

$$\mathcal{F}_t^{\mathrm{inf},a_1} = \mathcal{F}_t^{\mathrm{inf},a_2} = \{\theta, a_1, a_2\} \cup \mathcal{F}_t$$

and

$$\mathcal{F}_t^{\mathrm{uninf},a} = \{a\} \cup \left\{\tilde{P}_t\left(\hat{\theta}_t, \theta, a, a_-\right) = P_t\right\} \cup \mathcal{F}_t,$$

respectively. Then, for an informed trader, the relevant expectations and precisions are as follows:

$$\mu_t^{\mathrm{inf},a_1} = \mu_t^{\mathrm{inf},a_2} = \mathbb{E}[v \,|\, \theta] = \theta,$$
$$S_t^{\mathrm{inf}} = \left(\mathbb{V}[v \,|\, \theta]\right)^{-1} = S_\varepsilon.$$

For uninformed traders, we use (7), (8), and the Bayesian updating formula for normal variable to determine $\mu_t^{\mathrm{uninf},a}$ and S_t^{uninf}. The signal from the price is $\frac{P_t - \beta_1(t)\hat{\theta}_t - \beta_3(t)a}{\beta_2(t)} \sim \mathcal{N}(\theta, \frac{1}{\Phi(t)^2 S_A})$, where $\Phi(t) \equiv \frac{\beta_2(t)}{2\beta_3(t)}$. Then,

$$\frac{1}{S_t^{\mathrm{uninf}}} = \frac{1}{S_\theta + t/\sigma^2 + \Phi(t)^2 S_A} + \frac{1}{S_\varepsilon}$$

[6]See Øksendal (2003, p. 104) for a similar derivation.

and

$$\mu_t^{\text{uninf},a} = \frac{\left(S_\theta + t/\sigma^2\right)\hat{\theta}_t + \Phi\left(t\right)^2 S_A \frac{P_t - \beta_1(t)\hat{\theta}_t - \beta_3(t)a}{\beta_2(t)}}{S_\theta + t/\sigma^2 + \Phi\left(t\right)^2 S_A}.$$

Substituting these posterior means and precisions into (6), we obtain the individual demand of any trader given the market price P_t. Substituting these individual demands into (5),

$$\lambda \int_{\mathbb{R}} \left(\frac{\left(\mu_t^{\text{inf},a} - P_t\right) S_t^{\text{inf}}}{\gamma} - a \right) \alpha\, e^{-\lambda t} dN_0^{\text{inf}}(a) + \lambda \int_{\mathbb{R}} \left(\frac{\left(\mu_t^{\text{uninf},a} - P_t\right) S_t^{\text{uninf}}}{\gamma} - a \right) (1-\alpha)\, e^{-\lambda t} dN_0^{\text{uninf}}(a) = 0.$$

After cancellations, rearranging, and letting $\mu_t^{\text{inf}} = \mu_t^{\text{inf},a_1} = \mu_t^{\text{inf},a_2}$,

$$\alpha \frac{\left(\mu_t^{\text{inf}} - P_t\right) S_t^{\text{inf}}}{\gamma} + (1-\alpha) \frac{\left(\frac{\mu_t^{\text{uninf},a_1} + \mu_t^{\text{uninf},a_2}}{2} - P_t\right) S_t^{\text{uninf}}}{\gamma} = \frac{a_1 + a_2}{2}, \tag{10}$$

which implies that effective market clearing is independent of the level of trading delays. Two features of our model gives rise to this result. First, the individual demand (6) is independent of λ thanks to the lack of re-trade considerations. Indeed, a trader's optimal post-trade position is determined based only on the final payoff derived from the asset position because she gains access to the market at exactly one point in time. Second, although only a subset of the traders have access to the market at any point in time, the composition of those who trade stays the same because trading delays apply uniformly to all traders: The sum of the masses of informed and uninformed traders who gain market access, $\alpha e^{-\lambda t} + (1-\alpha)\, e^{-\lambda t} = e^{-\lambda t}$, becomes smaller and smaller over time but their composition stays the same and this gives rise to an effective market-clearing condition independent of λ.

The last step is to substitute the posterior means and precisions in (10). The resulting expression is linear in $\hat{\theta}_t$, θ, a_1, and a_2. Each shows up in only one place and is additive. Everything else is just constant coefficients. This confirms the guess of a linear pricing rule. The functions, $\beta_1(t)$, $\beta_2(t)$, and $\beta_3(t)$, can be solved for by matching coefficients, which takes us to the following proposition.

Proposition 1. *Given a share of informed traders α, the unique linear price process that clears the market is*

$$\tilde{P}_t\left(\hat{\theta}_t, \theta, a_1, a_2\right) = \left(1 - \beta\left(t\right)\right)\hat{\theta}_t + \beta\left(t\right)\left(\theta + \frac{a_1 + a_2}{2\Phi}\right), \tag{11}$$

where

$$\Phi = -\frac{\alpha S_\varepsilon}{\gamma}$$

and

$$\beta(t) = \frac{\alpha\left(S_\theta + t/\sigma^2 + S_\varepsilon\right) + \Phi^2 S_A}{S_\theta + t/\sigma^2 + \alpha S_\varepsilon + \Phi^2 S_A}. \tag{12}$$

We use the results of Proposition 1 to calculate forecasting price efficiency (FPE) and revelatory price efficiency (RPE) in the spirit of Bond, Edmans, and Goldstein (2012) and Bai, Philippon, and Savov (2016), where

$$FPE_t \equiv \left(\mathbb{V}\left[\theta \mid P_t\right]\right)^{-1}$$

and

$$RPE_t \equiv \left(\mathbb{V}\left[\theta \mid \{Z_s\}_{s\in[0,t]}, a, P_t\right]\right)^{-1} - \left(\mathbb{V}\left[\theta \mid \{Z_s\}_{s\in[0,t]}, a\right]\right)^{-1}.$$

FPE measures how well the market price reflects the forecastable part, θ, of the asset value $v = \theta + \varepsilon$. However, uninformed traders do not rely only on the market price when making their own forecasts. They use any other public information they possess and the information contained in their asset endowment realization. Thus, there is a need for a measure that reflects the unique information from the market price that is not captured by other information sources. RPE is such a measure. While FPE is only a forecasting concept, RPE measures the precision gain from observing the market price.

Corollary 2. *The forecasting price efficiency (FPE) and the revelatory price efficiency (RPE) are, respectively,*

$$FPE_t = S_\theta + \frac{1}{\frac{(1-\beta(t))^2}{S_\theta + t/\sigma^2} + 2\left(\beta(t)\right)^2 \frac{\gamma^2}{\alpha^2 S_\varepsilon^2 S_A}}$$

and

$$RPE = \frac{\alpha^2 S_\varepsilon^2 S_A}{\gamma^2},$$

where $\beta(t)$ is given by (12). Thus, given a share of informed traders, trading delays ($1/\lambda$) and how long-lived the private information is (σ) are irrelevant to RPE.

Corollary 2 shows that while FPE is a function of time, RPE is constant over time. That FPE is time varying is intuitive because the private information of the informed traders becomes progressively public as everyone observes the stochastic process Z and then the "representative" trader (i.e. the fictitious trader who shows up in the market-clearing condition (10) and has the same marginal valuation as the trading price) has a more precise information about θ. That RPE is constant is a less obvious result. One way to understand this RPE result above is that for uninformed traders, price is observationally equivalent to a signal of the form $\theta + \frac{a_-}{2\Phi}$ with the precision

$$\left(\mathbb{V}\left[\theta + \frac{a_-}{2\Phi} \,|\, \theta\right]\right)^{-1} = \frac{\alpha^2 S_\varepsilon^2 S_A}{\gamma^2}.$$

Because the unique information from observing the price after eliminating its parts relating to $\hat{\theta}_t$ and a is independent of t, its precision that we show is equal to RPE is independent of t as well.

Perhaps an even more surprising implication of Corollary 2 is that, given a share of informed traders, trading delays $(1/\lambda)$ and how long-lived the private information is (σ) are irrelevant to the precision of the price signal. Similar to the case where the effective market clearing (10) is independent of λ and σ that we explain above, this again arises because the frictions apply uniformly to all traders. Thus, at any date, a subset of traders that is representative of the entire pool of traders gain market access, which leaves RPE unaffected by the friction parameters σ and λ. It is important to note, however, this irrelevance result is a conditional irrelevance result because the fraction of informed traders is taken exogenously. In Section 4, we endogenize the fraction of informed traders by studying the traders' endogenous information acquisition decision. We show that because trading delays lead to a possibility of not being able to benefit from private information, traders have a reduced incentive to become informed when there are trading delays. This insight encapsulates the main difference between our model and Grossman and Stiglitz (1980).

As a preview for the results in Section 4, consider the left panel of Figure 1 that depicts two sample paths for the equilibrium price under different assumptions for the expected lifespan of the private information. Over time, the informational edge of informed traders reduces as uninformed traders observe Z_s for longer. If the degree, σ, of noise in this public signal process is lower, the equilibrium price quickly approaches the informed traders' signal, which, in turn, reduces the profitability of informed trading. In the Figure, the blue path for the equilibrium price ($\sigma = 0.5$) approaches the informed traders' signal, θ, a lot quicker than the red path

($\sigma = 2$) approaches. This means that given a level, $1/\lambda > 0$, of trading delays, traders are more likely to acquire costly private signals if they anticipate a sample path like the red path rather than the blue one. Moreover, given σ, a trader is more likely to acquire costly private signals if trading delays are less severe because less severe trading delays mean, in expectation, an informed trader can trade when the difference between the equilibrium price and the informed trader's signal is larger.

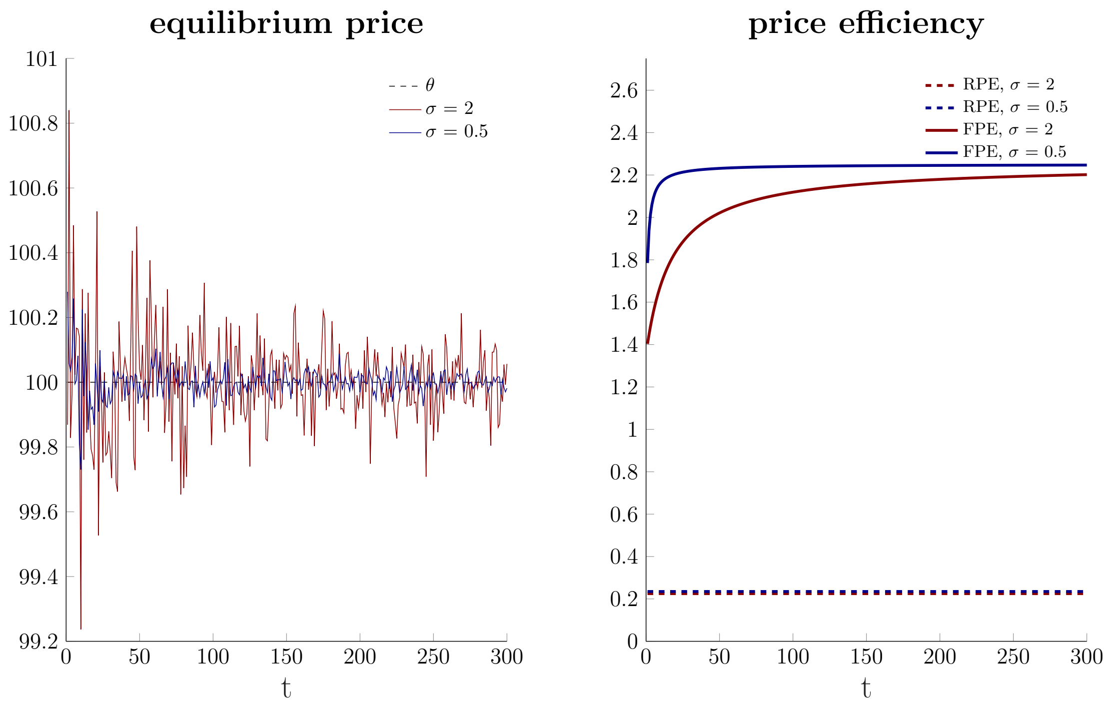


Figure 1: Sample path of equilibrium price on the left panel, and the price efficiency measures on the right panel. Parameter choices: $\alpha = 0.3$, $\gamma = 2$, $S_\theta = S_\varepsilon = 1$, and $S_A = 10$. Realized values for random variables: $\theta = 100$ and $a_1 = a_2 = 0$.

3.3 Equilibrium price dynamics

Because different traders who gain access to the market at different points in time face different market-clearing prices, they end up trading at different prices. This leads to economic questions: How does the *average* transaction price compare to the informed traders' signal, θ? How *variable* are the transaction prices across traders who trade at different points in time? The following Proposition offers answers to these questions.

Proposition 3. *Assume $S_\theta = 0$. Let $\delta = \alpha S_\varepsilon \left(1 + \frac{\alpha S_A S_\varepsilon}{\gamma^2}\right)$ and let $\mathbb{E}_C[P]$ and $\mathbb{V}_C[P]$ denote, respectively, the average and the variance of the transaction prices conditional on θ, a_1, and a_2.*

Then, the deviation of the average transaction price from the informed traders' signal per unit of supply is

$$\frac{\mathbb{E}_C[P]-\theta}{(a_1+a_2)/2}=-\frac{\gamma}{S_\varepsilon}\left\{1+\frac{1-\alpha}{\alpha}\lambda\sigma^2\delta e^{\lambda\sigma^2\delta}\Gamma\left(0,\lambda\sigma^2\delta\right)\right\} \tag{13}$$

and the variance of transaction prices is

$$\mathbb{V}_C[P]=(1-\alpha)^2\lambda\sigma^2\left\{-1+e^{\lambda\sigma^2\delta}\left(1+\lambda\sigma^2\delta\right)\Gamma\left(0,\lambda\sigma^2\delta\right)\right\}+o\left(a_1+a_2\right), \tag{14}$$

where $\Gamma(0,y)\equiv\int_y^\infty x^{-1}e^{-x}dx$ *is the upper incomplete gamma function.*

Let us refer to the deviation of the average transaction price from the informed traders' signal per unit of supply as average price discount, in short. Proposition 3 provides us with closed form formulas for the average price discount and a first-order approximation of the price variance. One can use these formulas to understand how various factors influence average price discount and price variance. For example, the following Corollary establishes results for trading delays.

Corollary 4. *Assume* $\gamma, S_\varepsilon, S_A, \sigma>0$. *Average price discount is monotone decreasing in* $\lambda>0$ *for all* $\alpha\in(0,1)$. *Price variance is increasing in* $\lambda\simeq 0$ *and decreasing in* $\lambda\simeq\infty$ *for all* $\alpha\in[0,1)$.

Figure 2 illustrates the results in Corollary 4 for a particular set of parameters. When $\lambda=0$, trading delays are infinite and none of the traders can trade until $t=\infty$ at which point the public signal process reveals the informed traders' signal θ. This means that everyone trades at the same price, and so, price variance is zero. Also, when $\lambda=0$, price discount is the smallest in terms of its absolute value, 2 (because $\gamma/S_\varepsilon=2$). As λ gets larger, trading delays decline and traders gain access to the market earlier than $t=\infty$. Because informed traders know θ, trading early does not change their willingness to pay. However, uninformed traders' willingness to pay is smaller because they trade when the public signal is less informative about θ. As a result, average price discount increases in its absolute value as trading delays get smaller.

The right panel of Figure 2 shows that price variance is hump-shaped in λ. As λ starts increasing from 0, traders start trading at random times earlier than $t=\infty$, and this increases price variance at first. But once λ reaches some high value, further increase in λ starts depressing price variance because then traders starts trading at times very close to $t=0$ and this makes their willingness to pay closer together because they trade before a meaningful precision has been obtained from the public signal process. Consistent with Corollary 4, this implies that the maximum level of price variance obtains for some intermediate value of λ.

To our knowledge, the non-monotonic relationship between trading delays and price variance is a novel result in the literature. The existing models of equilibrium price variance utilizing random trading times find that mitigating trading delays lead to a monotone decline in price variance. See, for example, Afonso and Lagos (2015), Hugonnier, Lester, and Weill (2015), Üslü and Velioğlu (2019), and Wong and Zhang (2022). These models feature symmetric information, and the reason behind price variance is traders' bilateral negotiation. As trading delays mitigate, traders' outside options become closer to each other, in turn making transaction prices closer together. In our model, instead, trade is multilateral with price taking traders, but information is asymmetric. Because the dispersion level of information that traders possess at their respective time of trading is maximized at an intermediate level of trading delays, a non-monotonic relationship between trading delays and price variance arises.

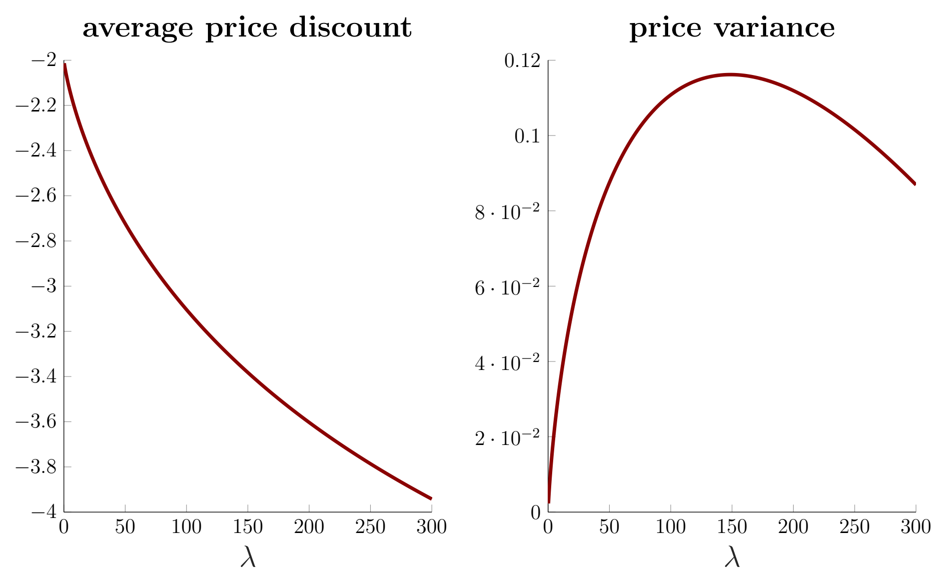


Figure 2: Average price discount as given by (13) and the price variance when $a_1 = a_2 = 0$ as given by (14), plotted as functions of λ. Parameter choices: $\alpha = 0.3$, $\gamma = 2$, $\sigma = 0.05$, $S_\theta = 0$, $S_\varepsilon = 1$, and $S_A = 10$.

Again, in this analysis, the exogeneity of α is key to obtain the patterns discussed above. However, as trading delays change, naturally traders' incentive to acquire information changes, and so, the *equilibrium* α would change, potentially changing the patterns above. Indeed, in Section 4, we show that once α is endogenized, the pattern regarding average price discount changes altogether becoming non-monotone. The pattern regarding price variance stays qualitatively the same but the magnitude of the association, i.e., the slopes, changes at different regions, becoming stronger on certain regions and weaker on others.

3.3.1 A discussion of the variance of the equilibrium transaction prices: price dispersion or price volatility?

Because of our assumption that all traders have access to the market exactly once during the trading session, in their strict interpretation, price dispersion and price volatility are the same in our model. Thus, how the variance of the equilibrium transaction prices maps to price dispersion and price volatility will depend on how one maps our model's trading session to trading in practice. Our leading mapping is as follows.

We consider the asset payoff as the sum of a learnable (θ) and an unlearnable news (ε). The learnable news is apparent to informed investors at the beginning of the trading session ($t = 0$). The unlearnable news will be apparent the everyone at the end of the trading session ($t = \infty$). No news arrive at the market during the trading session. Thus, our model's (infinite) time horizon maps in practice to a (finite) trading session during which no news about the asset's fundamentals arrive at the market. In the empirical literature, the variance (or standard deviation) of prices observed during a trading session without any news arrival is named price dispersion. See, Jankowitsch, Nashikkar, and Subrahmanyam (2011), Jankowitsch, Nagler, and Subrahmanyam (2014), and Pintér and Üslü (2022), among others.

Differently from the search-based models of price dispersion, the "representative" trader has different information at different points during the trading session in our model. Thus, our model generates price dispersion even in the absence of bilateral negotiations. Then, a natural question is how the representative trader's information state evolves in the absence of news. This question relates to how the diffusion process (1) for Z_t is interpreted. In the spirit of Amador and Weill (2012), we interpret the process (1) as progressively revealing the news at $t = 0$ to broader sets of traders over time rather being additional arrival of "new" information.

4 Information acquisition

So far, we have studied the equilibrium of the trading game by taking as given the fraction of informed traders. In this section, we study the traders' endogenous information acquisition decisions. Assume that a trader becomes an informed trader if she pays a cost of C at time $t = 0_{-}$ before her initial endowment realizes. She becomes an uninformed trader otherwise.

Using Lemma 1, a trader decides to be informed if

$$-e^{\gamma C}\mathbb{E}\left[e^{-\frac{1}{2}\left(\mu_t^{\text{inf}}-\tilde{P}_t\left(\hat{\theta}_t,\theta,a_1,a_2\right)\right)^2 S_t^{\text{inf}}}\right] \geq -\mathbb{E}\left[e^{-\frac{1}{2}\left(\mu_t^{\text{uninf},a}-\tilde{P}_t\left(\hat{\theta}_t,\theta,a,a_-\right)\right)^2 S_t^{\text{uninf}}}\right], \tag{15}$$

where the expectation is taken over the random variables θ, a_1, a_2, $\hat{\theta}_t$, and t. The following Lemma derives an interpretable version of this optimality condition.

Lemma 2. *Emphasizing its dependence on α through Φ and $\beta(t)$, let*

$$\omega(\alpha,t) = \sqrt{\frac{S_\theta + t/\sigma^2}{S_\theta + t/\sigma^2 + S_\varepsilon + \left(1 + \frac{S_\theta + t/\sigma^2}{\Phi^2 S_A}\right)(\beta(t))^2 S_\varepsilon}} \tag{16}$$

denote a measure of expected benefit of accessing the market at time t as an uninformed trader when the fraction of informed traders is α. Then, a trader finds it weakly optimal to be informed if

$$e^{\gamma C} \leq \frac{\int_0^\infty \sqrt{1 + \frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta + \gamma^2 t/\sigma^2 + \alpha^2 S_\varepsilon^2 S_A}}\,\omega(\alpha,t)\,\lambda e^{-\lambda t}dt}{\int_0^\infty \omega(\alpha,t)\,\lambda e^{-\lambda t}dt}, \tag{17}$$

and strictly optimal if the inequality is strict.

This optimality condition (17) reveals how traders will make their information acquisition decisions. The LHS is the utility cost of acquiring information, while the RHS is its utility benefit. Due to the exponential utility assumption, the latter (RHS) is the *ratio* of the expected (gross) utility of trading as an informed trader to the expected utility of trading as an uninformed trader.[7] If the LHS of (17) is strictly larger than its RHS, no one will acquire information and the equilibrium fraction of informed traders will be zero, $\alpha = 0$. If the LHS is strictly smaller, however, everyone will acquire information and, then, $\alpha = 1$. As in Grossman and Stiglitz (1980), the only possibility to obtain an equilibrium with the co-existence of informed and uninformed traders is that the LHS and the RHS are equal to each other and every trader has a mixed strategy, that is, they choose to be informed with probability $\alpha \in (0,1)$ such that

$$e^{\gamma C} = \frac{\int_0^\infty \sqrt{1 + \frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta + \gamma^2 t/\sigma^2 + \alpha^2 S_\varepsilon^2 S_A}}\,\omega(\alpha,t)\,\lambda e^{-\lambda t}dt}{\int_0^\infty \omega(\alpha,t)\,\lambda e^{-\lambda t}dt}. \tag{18}$$

[7] Proving Lemma 2 amounts to moving the expectation on the LHS of (15) to the RHS and multiplying both sides with -1, which yields the optimality condition (17).

When $\sigma = \infty$ or $\lambda = \infty$, we are at the Grossman and Stiglitz (1980) case and ω does not affect the indifference condition (18). As a result, the RHS in this special case is obviously strictly decreasing in α, which guarantees a unique equilibrium α:

$$e^{\gamma C} = \sqrt{1 + \frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta + \alpha^2 S_\varepsilon^2 S_A}}. \tag{19}$$

Intuitively, if information is not short-lived or if there are no trading delays, the utility benefit of acquiring information is not affected by the time of trading. If information is not short-lived, the economy, and hence, the resulting market price is stationary as the "representative" trader has the same information at any point in time. If there are no trading delays, everyone trades immediately at $t = 0$ at the same price before any information dispersion arises amongst uninformed traders. Thus, in both cases, the integrals in the numerator and the denominator of (18) are "degenerate" and so ωs cancel each other out, which leads to (19). However, in our model with short-lived information and trading delays, $\sigma < \infty$ and $\lambda < \infty$, and so, traders face a non-degenerate distribution of trading times over an infinite horizon, when making their information acquisition decisions at $t = 0$.

Proposition 5. *Suppose $\gamma^2 < \frac{S_\theta S_\varepsilon S_A}{S_\theta + S_\varepsilon}$. Then, there exists a unique equilibrium fraction $\alpha^\star \in [0, 1]$ of informed traders. It satisfies the following comparative statics:*

$$\frac{\partial \alpha^\star}{\partial \sigma} \geq 0, \quad \frac{\partial \alpha^\star}{\partial \lambda} \geq 0, \quad \textit{and} \quad \frac{\partial \alpha^\star}{\partial C} \leq 0.$$

These inequalities are strict for $\alpha^\star \in (0, 1)$.

Proposition 5 states that, if traders' risk aversion is not too high, there exists a unique equilibrium fraction of informed traders.[8] As C increases, naturally, a lower fraction of traders become informed as in Grossman and Stiglitz (1980). The novel comparative statics implied by our model are with respect to how long-lived the informational advantage is (σ) and liquidity (λ). Intuitively, if the informational advantage is shorter-lived, a lower fraction traders choose to be informed because of their worry that their private information may become effectively public before they are able to benefit from it, because trading is not instantaneous in this illiquid market. Relatedly, keeping the expected lifespan of the information constant, if market liquidity improves, the Proposition implies that a larger fraction of traders will become informed.

[8]This condition on risk aversion is a sufficient condition for unique $\alpha^\star$ because it guarantees that the net benefit of becoming informed is monotone for all $\alpha \in [0, 1]$. Numerical examples suggest that even if this condition is severely violated (e.g., for γ a hundred times larger than the stated upper bound), the net benefit of becoming informed is monotone for a large spectrum of αs apart from very small $\alpha \simeq 0$, which means that the Proposition holds for a larger spectrum of parameters than implied by the sufficient condition in its statement.

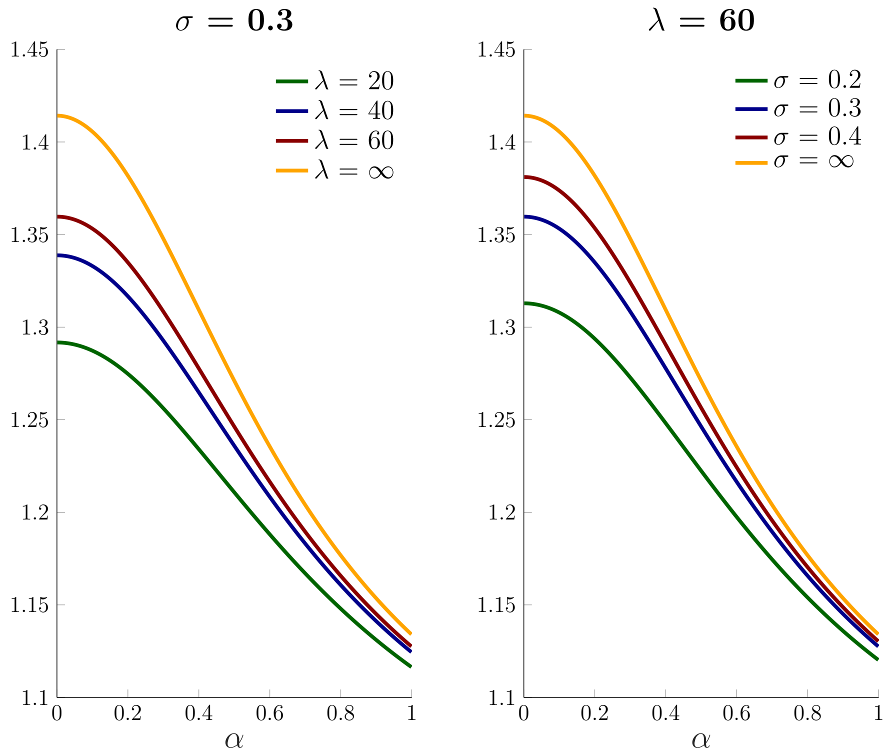


Figure 3: Benefit of becoming informed as measured by the RHS of Equation (18) for varying levels of λ on the left panel and for varying levels of σ on the right panel. Parameter choices: $\gamma = 2$, $S_\theta = S_\epsilon = 1$, and $S_A = 10$.

Figure 3 offers a graphical portrayal of Proposition 5. Given the effective cost $e^{\gamma C}$ of acquiring information, one can use the benefit functions depicted in Figure 3 to determine the equilibrium fraction $\alpha^\star$ of informed traders for varying levels of λ and σ. For example, if $e^{\gamma C} = 1.3$ and $\lambda = \infty$, the left panel implies a large $\alpha^\star \approx 0.47$, which coincides with the fraction of informed traders that would obtain in the Grossman and Stiglitz (1980) environment. If λ is reduced to 60 by keeping $e^{\gamma C} = 1.3$, now a smaller fraction of traders become informed, $\alpha^\star \approx 0.4 < 0.47$. If λ is further reduced to 20 or below, none of the traders become informed, $\alpha^\star = 0$. Similarly, by inspecting the right panel of Figure 3, one sees that the Grossman and Stiglitz (1980) result obtains when $\sigma = \infty$ in this case, and the equilibrium fraction of informed traders decreases as σ decreases, consistent with Proposition 5.

4.1 Market efficiency

Corollary 2 of Section 3.2 offers intuitive comparative statics about the market efficiency (RPE). The market is informationally more efficient if the fraction of informed traders (α) is higher, the variability of payoff after purchasing information is lower (i.e., S_ε is higher), uncertainty regarding the asset endowment of other traders is lower (i.e., S_A is higher), and traders are less risk averse. We argued in Section 3.2 that it was striking that trading delays and the short-lived nature of the private information do not matter for the price informativeness. Now, the new results with endogenous information acquisition imply that the market efficiency is affected by trading delays and the short-lived nature of the private information through the equilibrium level of informed trading, $\alpha^\star$.

Bai, Philippon, and Savov (2016) find that RPE increased for Standard & Poor's (S&P) 500 firms between 1960 to 2014. They further analyze the factors behind this increase and find that (i) it is not caused by changes in return predictability because the predictable component of returns is quite stable, and (ii) it is likely caused by information production. Through the lens of our model, the findings of Bai, Philippon, and Savov (2016) indicate that what is behind the increase in RPE is an increase in $\alpha^\star$, and not an increase in S_ε. Our model offers theoretical insights into these findings because $\alpha^\star$ is endogenous in our model. An important determinant of $\alpha^\star$ in our model is λ as shown by Proposition 5. It can be argued that λ increased during 1960-2014 because "[t]rading costs have fallen, and liquidity has increased by the orders of magnitude" (Bai, Philippon, and Savov, 2016, p. 625). Accordingly, our model implies that $\alpha^\star$ increased during the same period (keeping everything else constant).

In addition to the time-series evidence highlighted above, Bai, Philippon, and Savov (2016) have indirect cross-sectional evidence for the positive relationship between λ and informational efficiency implied by our model. They show that FPE is higher for high-turnover firms compared to low-turnover firms. Because FPE is a measure that subsumes RPE, FPE tends to increase as $\alpha^\star$ increases as well. Thus, this can be considered an indirect evidence for higher λ being the reason behind higher RPE.[9] In Section 5 of our paper, we, instead, provide direct cross-sectional evidence for this channel of our model by using transaction-level data from the UK corporate bond market. We use two proxies for RPE that are commonly employed in the market microstructure literature. We show that bonds with larger average number of daily transactions, which is the direct counterpart of λ in our model, exhibit higher informational efficiency implied by both proxies.[10]

[9] Bai, Philippon, and Savov (2016) do not use turnover in their formal analysis of RPE. Instead, they conduct their formal analysis of turnover by looking at FPE, and make an indirect inference for RPE: "These results show that higher liquidity is associated with higher price informativeness. As liquidity has risen significantly over time, this finding supports the view that rising liquidity has contributed to the observed rise in price informativeness and an increase in RPE" (p. 644).

[10] The two proxies are the return autocorrelation measure of Hendershott and Jones (2005) and the price delay measure of Hou and Moskowitz (2005), which are both inversely related with RPE. See Section 5.2 for more information.

While trading technologies and liquidity have improved substantially, which reduced trading delays during 1960-2014, these are not the only big changes happening in the financial sector. Indeed, the last decade has witnessed a substantial growth of "big data" financial technology. Thus, it is natural to ask how the market efficiency changes in response to traders' easier access to useful information. We contrast two possible scenarios: a decline in the information acquisition cost, C, representing traders' access to better datasets and more sophisticated data analysis technologies, or a decline in the noise of information revelation, σ, representing the shorter-lived nature of private information due to the technology allowing traders to extract others' information, which is a common concern about big data financial technologies as stated by Farboodi and Veldkamp (2019).

Proposition 5 and Corollary 2 indicate that while market efficiency improves in response to a decline in the information acquisition cost, it worsens in response to private information becoming shorter-lived. Hence, the impact of the growth of financial technology on market efficiency is ambiguous because both effects are, in principle, present for any asset albeit with different strengths. Bai, Philippon, and Savov (2016) present evidence consistent with this ambiguity: stock market efficiency exhibits a secular decline for the universe of all firms, while it exhibits a secular increase for the S&P500 firms. Our model rationalizes their findings if the improvement in market efficiency from the decline in C is dominated by (resp. dominates) the decline in market efficiency from the decline in σ for the universe of all firms (resp. the S&P500 firms). Table C.2 of Bai, Philippon, and Savov (2016, p. 653) shows that the magnitude of the decline in market efficiency for the universe of all firms is highest for the 3-year horizon during the 2010-2014 period of their dataset, when the big data arguably plays the most prominent role in their dataset.

4.2 Equilibrium price dynamics with information acquisition

In Section 3.3, we studied how the average price discount and the price variance depend on trading delays by taking as given the fraction of informed traders. However, the theoretical and empirical discussions above indicate that the level of trading delays is a first-order determinant of equilibrium information acquisition incentives, and so, as trading delays vary, the equilibrium fraction of informed traders vary as well, at least in the medium or long run. Thus, an important question is whether the patterns in Figure 2 are qualitatively robust to endogenous information acquisition. To this end, we reproduce below the results of Figure 2 with equilibrium information acquisition, presented along with the fixed exogenous $\alpha = 0.3$ as earlier.

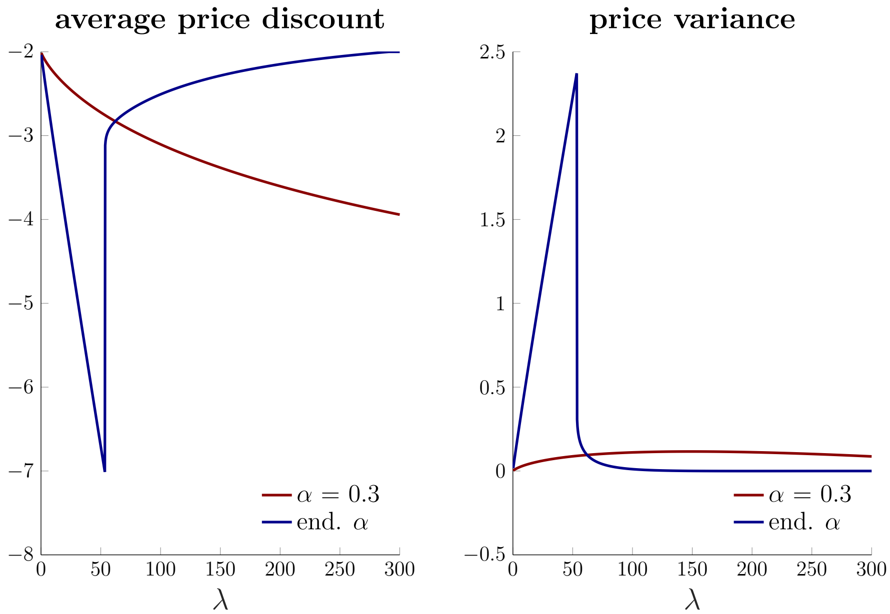


Figure 4: Average price discount as given by (13) and the price variance when $a_1 = a_2 = 0$ as given by (14), plotted as functions of λ. Parameter choices: $C = 0.054$, $\gamma = 2$, $\sigma = 0.05$, $S_\theta = 0$, $S_\varepsilon = 1$, and $S_A = 10$.

For the same parametrization in Figure 2, Figure 4 reveals that the effect of trading delay on price dynamics is very different when trading delays are severe enough to prohibit all information acquisition ($\lambda < 50$, $\alpha^\star = 0$) and when trading delays are moderate so that there are some informed traders ($\lambda > 50$, $\alpha^\star > 0$, and $\alpha^\star$ is increasing in λ). Similar to Figure 2, when trading delays are very severe, a mitigation of delays increases average price discount in its absolute value but this time the effect is even stronger. This is because, in Figure 4, the entire population of traders is uninformed. While only a fraction 0.7 of traders start trading early with worse precision in Figure 2, now the entire population (fraction one) of traders start trading with worse precision, and so, the price discount deepens very quickly. The worst discount is reached around $\lambda = 50$, when trading delays are severe enough to still prohibit information acquisition, but moderate enough to allow uninformed traders to trade early on average when their signal precision is weak. The right panel of Figure 4 also shows that the price variance reaches its maximum around the same level because this large population of uninformed traders gain access to the market at dispersed times.

Once λ exceeds a certain threshold both the average price discount and the price variance change direction very quickly. As λ increases further from this threshold, the average price discount gets smaller and smaller because informed traders with higher information precision start trading. This also increases the informativeness of the equilibrium price process, which increases the precision of uninformed trading as well. Similarly, price variance exhibits a sharp decline because traders' valuations become closer together thanks to the information that informed traders bring to the market.

One can decompose the average price discount (13) as the negative of the sum of a payoff risk premium (γ/S_ε) and a liquidity risk premium, because the former is independent of λ and the latter depends on λ. Then, a broader message of the example in Figure 4 is that illiquidity premium and price dispersion can be non-monotonic functions of λ in a model with endogenous information acquisition. To our knowledge, this is a novel result and in sharp contrast with the existing illiquid market models with exogenous information, which predict that illiquidity premium[11] and price dispersion[12] are monotone increasing in trading delays. In Section 5, we provide novel empirical evidence for the non-monotonic relationship between trading delays and price dispersion by using transaction-level data from the UK corporate bond market. We leave the empirical analysis of illiquidity premium for future work because disentangling payoff risk premium and illiquidity premium requires a structural approach, which goes beyond the scope of the reduced-form empirical analysis we provide in Section 5.[13]

5 Novel empirical evidence

The theory developed in previous sections yields two testable predictions, summarized as follows.

Prediction 1 (Informational efficiency)**.** *Informational efficiency is negatively related to trading delays, as implied by Proposition 5 and Corollary 2.*

Prediction 2 (Price dispersion)**.** *Price dispersion is non-monotonic in trading delays, i.e., price dispersion is lowest for the very thinly and the very actively traded assets, as implied by Corollary 4.*

5.1 Data and measurement

To test these two predictions, we use a transaction-level data from the UK corporate bond market, sourced by the UK Financial Conduct Authority. The dataset contains information on

[11]See, for example, Duffie, Gârleanu, and Pedersen (2005), Duffie, Gârleanu, and Pedersen (2007), and He and Milbradt (2014).

[12]See, for example, Afonso and Lagos (2015), Hugonnier, Lester, and Weill (2015), and Üslü and Velioğlu (2019).

[13]Among others, see d'Avernas (2017) and Chen, Cui, He, and Milbradt (2018) for such structural work.

the identities of counterparties, the transaction time, the transaction price and quantity, the International Securities Identification Number, and buyer-seller flags.[14] Our sample covers the period between August 2011 and December 2017. We include all available transactions (e.g., including client-dealer as well as interdealer trades). We follow Czech and Pinter (2020) in matching our transaction-level data with information on time-to-maturity as well as bond ratings from Thomson Reuters Eikon, covering the three major rating agencies Moody's, Standard & Poor's (S&P), and Fitch.[15] For each bond, we take the modal value of the rating during the lifetime of the given bond in our sample.

To measure trading delays at the bond-day level, we use the trade-level data to compute the number of daily transactions in the given bond. For each bond, we compute the average number of daily transactions across the trading days in order to measure trading frequency (the inverse of trading delays) at the bond-level. Moreover, we compute measures of price dispersion in the spirit of Jankowitsch, Nashikkar, and Subrahmanyam (2011). Specifically, we compute for each transaction k in bond i the following measure:

$$\sigma_k = \left(\log\left(p_k\right) - \log\left(m_{i,t}\right)\right)^2, \tag{20}$$

where p_k is the transaction price corresponding to trade k, and $m_{i,t}$ is the average daily transaction in bond i on day t. We then compute price dispersion, Σ_i, as the square root of the average values of σ_k in each bond i, expressed in basis points. We construct dispersion both at the bond- and the bond-day levels.

In Table 1, we present summary statistics for the variables used in the regression analyses below. Panel A shows the results at the bond-level. Across the 2,776 bonds, average trading frequency amounts to about 4-5 daily number of transactions. The mean and median of price dispersion are around 16-17 bps and 19-20 bps, respectively, with weighted and unweighted measures yielding similar values. The mean of year-to-maturity (9.4 years) is above the median (5.6 years), which is due to some corporate bonds in the sample having very long (more than 50 years) maturity. The median bond has a credit rating of BBB-.

[14] All transaction prices are clean prices (so they do not include accrued interest). For further details on the Zen dataset, see the Transaction Reporting User Pack: https://www.fca.org.uk/publication/finalised-guidance/fg15-03.pdf. Recent applications of the dataset can be found in Czech and Roberts-Sklar (2019), Czech, Huang, Lou, and Wang (2021), and Pintér and Üslü (2022) among others.

[15] Our default option is to use ratings from Moody's due to the firm's vast market coverage, and to use S&P ratings if ratings from Moody's are unavailable for a certain bond. Fitch ratings as a third option if necessary. Credit ratings in our dataset take 24 numerical values (1=AAA, 2=AA+, 3=AA, 4=AA-, 5=A+, 6=A, 7=A-, 8=BBB+, 9=BBB, 10=BBB-, 11=BB+, 12=BB, 13=BB-, 14=B+, 15=B, 16=B-, 17=C, 18=C+, 19=CC, 20=CC+, 21=CC-, 22=CCC, 23=CCC+, 24=DDD).

Panel B of Table 1 presents the results at the bond-day level. Trading frequencies continue to average around 4-5 daily number of observations (though they have a larger standard deviation compared to the bond-level, i.e. 7.9 vs 3.9 transactions). The mean (12 bps) and median (16 bps) of price dispersion are smaller than at the bond-level, which is driven by the fact that bonds with low price dispersion tend to trade on more days than bonds with higher dispersion. Similarly, the median credit rating is 8 (corresponding to BBB+ rating) at the bond-day level, which is higher than the median (BBB-) at the bond-level. This is due to investment-grade bonds being traded on more trading days than high-yield bonds.

Table 1: Summary Statistics

	(1) mean	(2) p25	(3) p50	(4) p75	(5) sd	(6) N
Panel A: Bond Level						
Trading Frequency	4.90	3.28	4.05	5.30	3.90	2,776
Price Dispersion (W)	20.65	6.88	16.97	29.04	18.84	2,776
Price Dispersion (U)	19.01	7.13	16.32	27.14	16.42	2,776
Maturity	9.38	2.82	5.56	12.61	10.23	2,776
Credit Rating	12.48	6.00	10.00	23.00	8.24	2,776
Panel B: Bond-Day Level						
Trading Frequency	5.22	2.00	4.00	6.00	7.89	625,925
Price Dispersion (W)	12.18	0.00	3.35	16.18	19.94	625,925
Price Dispersion (U)	12.11	0.00	3.47	16.46	19.41	625,925
Maturity	9.40	3.45	6.35	12.22	10.98	625,925
Credit Rating	9.74	6.00	8.00	11.00	6.66	625,925

Notes: This table reports summary statistics at the bond-level and the day-bond level. Trading frequency is measured as the number of daily transactions in the given bond. Price dispersion is the square root of either the trade-size weighted (W) or unweighted (U) mean of squared price deviations (see formula (20)) in basis points. Maturity is measured by years-to-maturity. Credit ratings take 24 numerical values (1=AAA, 2=AA+, 3=AA, 4=AA-, 5=A+, 6=A, 7=A-, 8=BBB+, 9=BBB, 10=BBB-, 11=BB+, 12=BB, 13=BB-, 14=B+, 15=B, 16=B-, 17=C, 18=C+, 19=CC, 20=CC+, 21=CC-, 22=CCC, 23=CCC+, 24=DDD).

5.2 Testing Prediction 1

To test the theoretical prediction regarding informational efficiency, we follow Indriawan, Pascual, and Shkilko (2022) in measuring price efficiency with two standard metrics: return autocorrelation as in Hendershott and Jones (2005), and price delay as in Hou and Moskowitz (2005). Both measures build on the notion that time-variation in expected returns is negligible at high frequency, so returns should be unpredictable over a short time horizon if markets are informationally efficient. Deviations from return unpredictability could therefore be used as a proxy for *informational inefficiency*.

To compute the first measure, we estimate the following time-series regression separately for each bond i:

$$r_{i,t} = \alpha_i + \beta_i r_{i,t-1} + \varepsilon_{i,t}, \tag{21}$$

where α_i is a constant and $r_{i,t}$ is the daily return of bond i on day t, with returns computed as the log difference of average daily transaction prices. We then take the absolute values of the estimated β_i coefficients, defined as $B_i \equiv |\beta_i|$. Larger values of B_i suggests greater inefficiency in bond i.

As a preliminary step for the construction of the second measure, we estimate the following time-series regression separately for each bond i:

$$r_{i,t} = a_i + \sum_{n=1}^{3}\sum_{k=0}^{2} b_{i,k,n} f_{n,t-k} + \sum_{k=0}^{2} \gamma_{i,k} r_{m,t-k} + \nu_{i,t} \tag{22}$$

where a_i is a constant, $f_{n,t-k}$ is the daily change in the first three principal components of the UK yield curve, $n \in \{1, 2, 3\}$, corresponding to the level, slope, and curvature of the yield curve.[16] The term $r_{m,t-k}$ is the daily log return on the FTSE stock index. The inclusion of both bond market and stock market returns is motivated by the previous literature (e.g., Hotchkiss and Ronen (2002)) on the relevance of both markets in affecting short horizon return dynamics. We then compute the R^2 statistics from regression (22) as well as from a variant of regression (22) which does not include the lagged values of the regressors. We refer to these as unconstrained, R^2_u, and constrained, R^2_c, statistics. For bond i, the price delay is then defined as $D_i \equiv 1 - R^2_c/R^2_u$, which takes values between zero and one. A larger value suggests greater inefficiency in bond i.

Given our two measures of price efficiency, we then run the following cross-sectional regression:

$$Y_i = c + \delta\lambda_i + controls + u_i, \tag{23}$$

where $Y_i = \{B_i, D_i\}$ is either our measure of absolute autocorrelations or our measure of price delay; c is a constant; and λ_i denotes trading frequency measured by the log of average daily number of transactions in bond i.[17] As controls in (23), we include a bond rating dummy which takes value one if the bond has an investment grade (BBB- or higher) and zero otherwise. In

[16] We use data on zero-coupon bond yields (with maturity of 1-25 years), as constructed by the Bank of England. The data can be downloaded from the Bank of England's website.

[17] We also experiment with alternative ways of measuring trading delays. Section B of the Appendix re-estimates regression (23) after we replace the regressor with the natural logarithm of trading volume—defined as the nominal value of transactions—at the bond-day level. As we can see, all the results are qualitatively the same as our baseline.

addition, we sort bonds into four equal sized buckets in terms of years-to-maturity and include in the regression dummy variables corresponding to each bucket.[18]

Table 2: Informational Efficiency and Trading Delays

	(1)	(2)	(3)	(4)	(5)	(6)
	Return Autocorrelation			Price Delay		
	All Bonds	Less Liquid	More Liquid	All Bonds	Less Liquid	More Liquid
Panel A: No Controls						
λ	-0.175***	-0.186***	-0.149***	-0.128***	-0.080**	-0.058*
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
Constant	0.516***	0.538***	0.465***	0.630***	0.602***	0.471***
	(0.03)	(0.05)	(0.05)	(0.03)	(0.05)	(0.06)
N	1080	541	539	1419	711	708
R^2	0.073	0.059	0.058	0.024	0.008	0.003
Panel B: With Controls						
λ	-0.148***	-0.152***	-0.124***	-0.125***	-0.066**	-0.076**
	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)
Inv. Grade	-0.006	0.024	-0.037**	-0.146***	-0.135***	-0.132***
	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Maturity=Q2	-0.087***	-0.054**	-0.119***	-0.058***	-0.070***	-0.051**
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
Maturity=Q3	-0.105***	-0.074***	-0.142***	-0.100***	-0.120***	-0.093***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Maturity=Q4	-0.156***	-0.147***	-0.169***	-0.143***	-0.165***	-0.138***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	0.563***	0.542***	0.556***	0.802***	0.759***	0.668***
	(0.03)	(0.05)	(0.05)	(0.03)	(0.05)	(0.06)
N	1080	541	539	1419	711	708
R^2	0.174	0.126	0.246	0.176	0.166	0.144

Notes: this table presents results from cross-sectional regressions whereby measures of price efficiency are regressed on trading frequency of corporate bonds. Panel A presents the results without controls, and Panel B presents the results with controls. Columns 1-3 show the results where the absolute value of return autocorrelation (computed by (21)) is used as the left-hand side variables. Columns 4-6 show the results where price delay (computed by (22)) is used as the regressand. Trading frequency (the regressor) is measured as the natural logarithm of the average daily number of transactions in bond i. For a bond to be included in the time-series regressions ((21) or (22)), it has to have at least 20 daily observations. Less (more) liquid bonds are defined as those that have approximately less (more) than the median number of daily transactions. The sample covers the period from August 2011 to December 2017. The trade-level dataset covers about 3.8 million transactions, including both client-dealer and inter-dealer trades. T-statistics in parentheses are based on robust standard errors. Asterisks denote significance levels (* $p<0.1$, ** $p<0.05$, *** $p<0.01$).

Results from regression (23) are presented in Table 2. Panel A shows the results for the regression without control variables. We find strong supporting evidence for the first prediction of the theoretical model: corporates bonds feature lower informational efficiency when they are traded less frequently. Specifically, columns 1-3 of the Table shows that one log point increase in the average number of daily transactions in a bond decreases the absolute value of

[18]The four maturity buckets correspond to 0-3, 4-6, 6-11, and 11+ years to maturity.

autocorrelation by about 0.15-1.19 points. The effect is larger among less liquid bonds, i.e., those bonds that are on average traded less frequently. Columns 4-6 of Table 2 shows that a one log point increase in trading frequency decreases price delays by about 0.06-0.13 points, with the effect being statistically and economically more significant among less liquid bonds.

Panel B of Table 2 presents the results when we control for credit rating as well as maturity. We find that bonds with better rating and higher time-to-maturity tend to be more informationally efficient. Importantly, while the inclusion of these controls somewhat diminishes the estimated effect of trading frequency, the coefficients continue to be statistically significant, (with some increase in the statistical significance of trading frequency in the price delay regressions).

5.3 Testing Prediction 2

To test for the theoretical prediction regarding non-monotonic relationship between price dispersion and trading delays, we compute average price dispersion and trading delays at the day-bond level. Then, we estimate the following panel regression:

$$D_{i,t} = \alpha_i + \mu_t + \iota\lambda_{i,t} + \kappa\lambda_{i,t}^2 + u_{i,t}, \tag{24}$$

where α_i and μ_t are bond- and day-level fixed effects, respectively, c is a constant, and λ_i denotes trading frequency measured by the log of average daily number of transactions in bond i, as used above.[19] The panel specification thereby allows us to control for time-invariant bond characteristics as well as for the average effect of macroeconomic shocks that may confound the dispersion – trading delay relationship. Moreover, we compute average price dispersion in two ways: unweighted average and trade-size-weighted average price dispersion.

Positive estimates of ι and negative estimates of κ would be consistent with a hump-shaped non-monotonic relation, predicted by our theoretical model. Table 3 presents the results from the panel regressions using about 0.63 million bond-day observations. We find strong support for this second empirical prediction of our theoretical model. In all the regression specifications, the linear effect of trading frequency is estimated to be positive, whereas the quadratic effect is negative. This implies a positive relationship between trading frequency and dispersion; however the relationship changes beyond a certain level of trading frequency.

[19]Section B of the Appendix re-estimates regression (24) after we replace the regressor with the natural logarithm of trading volume—defined as the nominal value of transactions—at the bond-day level. As we can see, all the results are qualitatively the same as our baseline.

Table 3: Price Dispersion and Trading Delays

	(1)	(2)	(3)
Panel A: Weighted Dispersion			
λ	16.550***	15.568***	15.854***
	(0.57)	(0.42)	(0.42)
λ^2	-2.154***	-2.123***	-2.255***
	(0.18)	(0.11)	(0.11)
N	625927	625665	625662
R^2	0.102	0.240	0.275
Panel B: Unweighted Dispersion			
λ	17.044***	15.836***	16.139***
	(0.61)	(0.43)	(0.43)
λ^2	-2.281***	-2.204***	-2.339***
	(0.19)	(0.11)	(0.12)
N	625927	625665	625662
R^2	0.107	0.254	0.291
Bond FE	No	Yes	Yes
Day FE	No	No	Yes

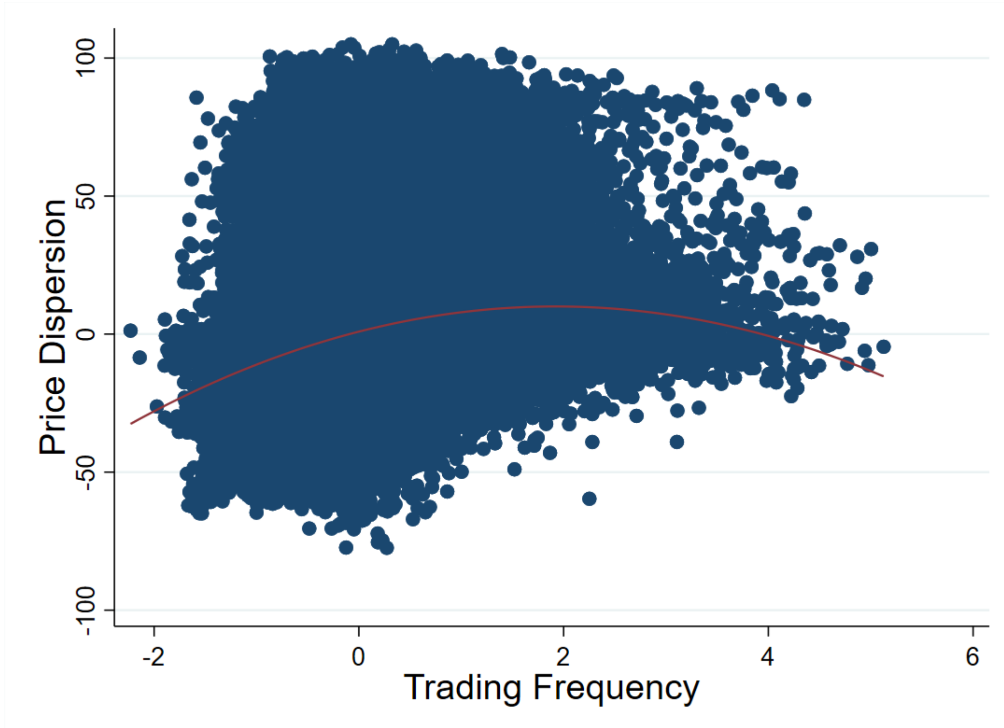
Notes: this table presents results from panel regressions whereby measures of price dispersions are regressed on trading frequency of corporate bonds. The regressor is included in a linear and a quadratic way. The sample covers the period from August 2011 to December 2017. The trade-level dataset covers about 3.8 million transactions, including both client-dealer and inter-dealer trades. T-statistics in parentheses are based on robust standard errors, where we also apply two-way clustering at the bond- and day-level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

To provide a visual illustration of the non-monotonic relationship, we use the most conservative specification (with bond and day fixed effects) to construct a scatter plot as follows. We purge our bond and day fixed effects from unweighted price dispersion and trading frequency, and use the obtained 625,662 residuals to draw a scatter plot as shown by Figure 5. The Figure shows a visibly non-monotonic relation: bond-day observations with intermediate trading delays tend to have higher dispersion compared to bond-day observations with very high or very low delays.

6 Conclusion

We present an asset market model that aims at bridging the gap between the information acquisition literature and the trading delays literature and provide empirical evidence for its novel implications. Our model establishes an inverse relationship between trading delays and price informativeness in equilibrium. Consistent with this implication of our model, we show that the prices of corporate bonds that are easier (resp. more difficult) to trade are informationally more (resp. less) efficient in the UK market. Our model also establishes that equilibrium price

Figure 5: Illustrating the Hump-shaped Relationship between Price Dispersion and Trading Delays



Notes: the scatter plot is constructed using the 625,662 residuals corresponding to column 4 of Panel B in Table 3. That is, to obtain the residuals, we estimate $\Sigma_{i,t} = \alpha_i + \mu_t + \nu_{i,t}$ and $\lambda_{i,t} = \alpha_i + \mu_t + u_{i,t}$. The red line is a fitted quadratic regression line.

dispersion is non-monotonic in trading delays. We, again, empirically confirm this relationship in the cross section of bonds traded in the UK corporate bond market.

Another novel implication of our model is that equilibrium illiquidity premium is non-monotonic in trading delays. Thus, one avenue for future research that our model leads to is to investigate the empirical relationship between illiquidity premium and trading delays. This would require a structural approach because of the necessity of decomposing risk premia into a payoff risk premium and an illiquidity risk premium.

Theoretical introspection leads to further avenues for future research. Because our model focuses on studying market-wide trading delays, it makes the assumption that uninformed and informed traders are subject to the same level of trading delays. One potential extension is to allow for heterogeneity in trading delays across traders. With endogenous information acquisition, this would allow one to study how different traders sort into different information types in equilibrium. Second, because trading delays are an important determinant of traders' information acquisition incentives and implicitly of their trading activity, it is natural to think that, at least in the long run, market makers would adjust their trade execution speed as a strategic response to traders' optimal information acquisition and trading activity. In order to analyze this (long-run) relationship, one could study the market makers' endogenous investment in execution speed.

References

Anat R. Admati and Paul Pfleiderer. A monopolistic market for information. *Journal of Economic Theory*, 39:400–438, 1986.

Gara Afonso and Ricardo Lagos. Trade dynamics in the market for federal funds. *Econometrica*, 83:263–313, 2015.

Yacine Aït-Sahalia and Mehmet Sağlam. High frequency traders: Taking advantage of speed. Working paper, 2013.

Manuel Amador and Pierre-Olivier Weill. Learning from private and public observations of others' actions. *Journal of Economic Theory*, 147:910–940, 2012.

Ana Babus and Péter Kondor. Trading and information diffusion in over-the-counter markets. *Econometrica*, 86:1727–1769, 2018.

Jennie Bai, Thomas Philippon, and Alexi Savov. Have financial markets become more informative? *Journal of Financial Economics*, 122:625–654, 2016.

Bruno Biais, Thierry Foucault, and Sophie Moinas. Equilibrium fast trading. *Journal of Financial Economics*, 116:292–313, 2015.

Max R. Blouin and Roberto Serrano. A decentralized market with common values uncertainty: Non-steady states. *Review of Economic Studies*, 68:323–346, 2001.

Philip Bond, Alex Edmans, and Itay Goldstein. The real effects of financial markets. *Annual Review of Financial Economics*, 4:339–360, 2012.

Eric Budish, Peter Cramton, and John Shim. The high-frequency trading arms race: Frequent batch auctions as a market design response. *Quarterly Journal of Economics*, 130(4):1547–1621, 2015.

Giovanni Cespa. Information sales and insider trading with long-lived information. *Journal of Finance*, 63:639–672, 2008.

Hui Chen, Rui Cui, Zhiguo He, and Konstantin Milbradt. Quantifying liquidity and default risks of corporate bonds over the business cycle. *Review of Financial Studies*, 31(3):852–897, 2018.

Robert Czech and Gabor Pinter. Informed trading and the dynamics of client-dealer connections in corporate bond markets. Bank of England working papers 895, Bank of England, November 2020. URL https://ideas.repec.org/p/boe/boeewp/0895.html.

Robert Czech and Matt Roberts-Sklar. Investor behaviour and reaching for yield: Evidence from the sterling corporate bond market. *Financial Markets, Institutions & Instruments*, 28 (5):347–379, 2019. doi: https://doi.org/10.1111/fmii.12122. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/fmii.12122.

Robert Czech, Shiyang Huang, Dong Lou, and Tianyu Wang. Informed trading in government bond markets. *Journal of Financial Economics*, 142(3):1253–1274, 2021. URL https://EconPapers.repec.org/RePEc:eee:jfinec:v:142:y:2021:i:3:p:1253-1274.

Tri Vi Dang and Florian Morath. The taxation of bilateral trade with endogenous information. Working paper, 2015.

Adrien d'Avernas. Disentangling credit spreads and equity volatility. Working paper, 2017.

Eduardo Dávila and Cecilia Parlatore. Trading costs and informational efficiency. *Journal of Finance*, Forthcoming, 2020.

Marco Di Maggio, Amir Kermani, and Zhaogang Song. The value of trading relations in turbulent times. *Journal of Financial Economics*, 124(2):266–284, 2017.

Jens Dick-Nielsen, Thomas K. Poulsen, and Obaidur Rehman. Dealer networks and the cost of immediacy. Working paper, 2021.

James Dow and Rohit Rahi. Should speculators be taxed? *Journal of Business*, 73(1):89–107, 2000.

Darrell Duffie, Nicolae Gârleanu, and Lasse H. Pedersen. Over-the-counter markets. *Econometrica*, 73(6):1815–1847, 2005.

Darrell Duffie, Nicolae Gârleanu, and Lasse H. Pedersen. Valuation in over-the-counter markets. *Review of Financial Studies*, 20(6):1865, 2007.

Darrell Duffie, Semyon Malamud, and Gustavo Manso. Information percolation in segmented markets. *Journal of Economic Theory*, 153:1–32, 2014.

Jérôme Dugast and Thierry Foucault. Data abundance and asset price informativeness. *Journal of Financial Economics*, 130(2):367–391, 2018.

Amy K. Edwards, Lawrence E. Harris, and Michael S. Piwowar. Corporate bond market transaction costs and transparency. *Journal of Finance*, 63(2):1421–1451, 2007.

Maryam Farboodi and Laura L. Veldkamp. Long run growth of financial data technology. *American Economic Review*, Forthcoming, 2019.

Maryam Farboodi, Gregor Jarosch, and Robert Shimer. The emergence of market structure. Working paper, 2015.

Thierry Foucault, Johan Hombert, and Ioanid Roşu. News trading and speed. *Journal of Finance*, 71(1):335–382, 2016.

Kenneth D. Garbade and William L. Silber. Price dispersion in the government securities market. *Journal of Political Economy*, 84(4):721–740, 1976.

Nicolae Gârleanu. Portfolio choice and pricing in illiquid markets. *Journal of Economic Theory*, 144(2):532–564, 2009.

Vincent Glode and Christian C. Opp. Asymmetric information and intermediation chains. *American Economic Review*, 106:2699–2721, 2016.

Vincent Glode, Christian C. Opp, and Xingtan Zhang. On the efficiency of long intermediation chains. *Journal of Financial Intermediation*, 38:11–18, 2019.

Mikhail Golosov, Guido Lorenzoni, and Aleh Tsyvinski. Decentralized trading with private information. *Econometrica*, 82(3):1055–1091, 2014.

Sanford J. Grossman and Joseph E. Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70:393–408, 1980.

Zhiguo He and Konstantin Milbradt. Endogenous liquidity and defaultable bonds. *Econometrica*, 82(4):1443–1508, 2014.

Terrence Hendershott and Charles M. Jones. Island Goes Dark: Transparency, Fragmentation, and Regulation. *Review of Financial Studies*, 18(3):743–793, 2005. URL https://ideas.repec.org/a/oup/rfinst/v18y2005i3p743-793.html.

Edith S. Hotchkiss and Tavy Ronen. The Informational Efficiency of the Corporate Bond Market: An Intraday Analysis. *Review of Financial Studies*, 15(5):1325–1354, 2002. URL https://ideas.repec.org/a/oup/rfinst/v15y2002i5p1325-1354.html.

Kewei Hou and Tobias J. Moskowitz. Market frictions, price delay, and the cross-section of expected returns. *The Review of Financial Studies*, 18(3):981–1020, 2005. ISSN 08939454, 14657368. URL http://www.jstor.org/stable/3598084.

Julien Hugonnier, Benjamin Lester, and Pierre-Olivier Weill. Heterogeneity in decentralized asset markets. 2015.

Julien Hugonnier, Benjamin Lester, and Pierre-Olivier Weill. Frictional intermediation in over-the-counter markets. *Review of Economic Studies*, 87(3):1432–1469, 2020.

Vivian Hutson, John S. Pym, and Michael J. Cloud. *Applications of Functional Analysis and Operator Theory*. Elsevier B.V., Amsterdam: the Netherlands, second edition, 2005.

Ivan Indriawan, Roberto Pascual, and Andriy Shkilko. On the Effects of Continuous Trading. working paper, ssrn, 2022.

Rainer Jankowitsch, Amrut Nashikkar, and Marti G. Subrahmanyam. Price dispersion in otc markets: A new measure of liquidity. *Journal of Banking and Finance*, 35(2):343–357, 2011. URL https://EconPapers.repec.org/RePEc:eee:jbfina:v:35:y:2011:i:2:p:343-357.

Rainer Jankowitsch, Florian Nagler, and Marti G. Subrahmanyam. The determinants of recovery rates in the us corporate bond market. *Journal of Financial Economics*, 114(1):155–177, 2014.

Mahyar Kargar, Benjamin Lester, Sebastien Plante, and Pierre-Olivier Weill. Sequential search for corporate bonds. Working paper, 2022.

Ricardo Lagos and Guillaume Rocheteau. Liquidity in asset markets with search frictions. *Econometrica*, 77(2):403–426, 2009.

Stephan Lauermann and Asher Wolinsky. Search with adverse selection. *Econometrica*, 84: 243–315, 2016.

Stephan Lauermann, Wolfram Merzyn, and Gábor Virág. Learning and price discovery in a search market. *Review of Economic Studies*, 85:1159–1192, 2017.

Jongsub Lee, Andy Naranjo, and Güner Velioğlu. When do cds spreads lead? rating events, private entities, and firm-specific information flows. *Journal of Financial Economics*, 130(3): 556–578, 2018.

Samuel Lee. Active investment, liquidity externalities, and markets for information. Working paper, 2013.

Benjamin Lester, Ali Shourideh, Venky Venkateswaran, and Ariel Zetlin-Jones. Market-making with search and information frictions. Working paper, 2018.

Dan Li and Norman Schürhoff. Dealer networks. *Journal of Finance*, 74(1):91–144, 2019.

Wei Li. High frequency trading with speed hierarchies. Working paper, 2018.

Wei Li and Zhaogang Song. Dealers as information intermediaries in over-the-counter markets. Working paper, 2018.

Maureen O'Hara, Yihui Wang, and Xing (Alex) Zhou. The execution quality of corporate bonds. *Journal of Financial Economics*, 130(2):308–326, 2018.

Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag Berlin Heidelberg, New York, Dordrecht, London, 6 edition, 2003.

Emiliano S. Pagnotta and Thomas Philippon. Competing on speed. *Econometrica*, 86(3): 1067–1115, 2018.

Gábor Pintér and Semih Üslü. Comparing search and intermediation frictions across markets. Bank of England working papers 974, Bank of England, April 2022. URL https://ideas.repec.org/p/boe/boeewp/0974.html.

Avanidhar Subrahmanyam. Transaction taxes and financial market equilibrium. *Journal of Business*, 71(1):81–118, 1988.

Semih Üslü. Pricing and liquidity in decentralized asset markets. *Econometrica*, 87(6):2079–2140, 2019.

Semih Üslü and Güner Velioğlu. Liquidity in the cross section of otc assets. Working paper, 2019.

Laura L. Veldkamp. Media frenzies in markets for financial information. *American Economic Review*, 96:577–601, 2006a.

Laura L. Veldkamp. Information markets and comovements of asset prices. *Review of Economic Studies*, 73:823–845, 2006b.

Laura L. Veldkamp. *Information Choice in Macroeconomics and Finance*. Princeton University Press, Princeton, NJ, 2011.

Robert E. Verrecchia. Information acquisition in a noisy rational expectations economy. *Econometrica*, 50:1415–1430, 1982.

Xavier Vives. Endogenous public information and welfare in market games. *Review of Economic Studies*, 84(2):935–963, 2017.

Pierre-Olivier Weill. Leaning against the wind. *Review of Economic Studies*, 74(4):1329, 2007.

Pierre-Olivier Weill. The search theory of otc markets. *Annual Review of Economics*, Forthcoming, 2020.

Asher Wolinsky. Information revelation in a market with pairwise meetings. *Econometrica*, pages 1–23, 1990.

Tsz-Nga Wong and Mengbo Zhang. Disintermediating the federal funds market. Working paper, 2022.

A Proofs

A.1 Proof of Corollary 2

Using (7), (8), (11), and the posterior variance formula for normally distributed random variables (Equation (2.4) of Veldkamp (2011, p. 12)),

$$\mathbb{V}\left[\theta \mid P_t\right] = \frac{(1-\beta(t))^2}{S_\theta + t/\sigma^2} + \frac{2}{S_A}\left(\frac{\beta(t)}{\Phi}\right)^2,$$

$$\mathbb{V}\left[\theta \mid \{Z_s\}_{s\in[0,t]}, a\right] = \left(S_\theta + t/\sigma^2\right)^{-1},$$

and

$$\mathbb{V}\left[\theta \mid \{Z_s\}_{s\in[0,t]}, a, P_t\right] = \left(S_\theta + t/\sigma^2 + \Phi^2 S_A\right)^{-1}.$$

Then, using the definitions

$$FPE_t \equiv \left(\mathbb{V}\left[\theta \mid P_t\right]\right)^{-1}$$

and

$$RPE_t \equiv \left(\mathbb{V}\left[\theta \mid \{Z_s\}_{s\in[0,t]}, a, P_t\right]\right)^{-1} - \left(\mathbb{V}\left[\theta \mid \{Z_s\}_{s\in[0,t]}, a\right]\right)^{-1}.$$

and using the fact that $\Phi = -\frac{\alpha S_\varepsilon}{\gamma}$, the formulas of the Corollary follow.

A.2 Proof of Proposition 3

$$\begin{aligned}
\mathbb{E}_C\left[P\right] &\equiv \mathbb{E}\left[\tilde{P}_t\left(\hat{\theta}_t, \theta, a_1, a_2\right) \mid \theta, a_1, a_2\right] = \mathbb{E}\left[\mathbb{E}\left[\tilde{P}_t\left(\hat{\theta}_t, \theta, a_1, a_2\right) \mid t, \theta, a_1, a_2\right] \mid \theta, a_1, a_2\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[(1-\beta(t))\,\hat{\theta}_t + \beta(t)\left(\theta + \frac{a_1+a_2}{2\Phi}\right) \mid t, \theta, a_1, a_2\right] \mid \theta, a_1, a_2\right] \\
&= \mathbb{E}\left[(1-\beta(t))\,\mathbb{E}\left[\hat{\theta}_t \mid t, \theta\right] + \beta(t)\left(\theta + \frac{a_1+a_2}{2\Phi}\right) \mid \theta, a_1, a_2\right] \\
&= \mathbb{E}\left[(1-\beta(t))\,\theta + \beta(t)\left(\theta + \frac{a_1+a_2}{2\Phi}\right) \mid \theta, a_1, a_2\right] \\
&= \theta + \frac{a_1+a_2}{2\Phi}\mathbb{E}\left[\beta(t)\right] = \theta + \frac{a_1+a_2}{2\Phi}\int_0^\infty \frac{\alpha\left(t/\sigma^2 + S_\varepsilon\right) + \Phi^2 S_A}{t/\sigma^2 + \alpha S_\varepsilon + \Phi^2 S_A}\lambda e^{-\lambda t} dt \\
&= \theta + \frac{a_1+a_2}{2\Phi}\int_0^\infty \left[\alpha + \frac{(1-\alpha)\left(\alpha S_\varepsilon + \Phi^2 S_A\right)}{t/\sigma^2 + \alpha S_\varepsilon + \Phi^2 S_A}\right]\lambda e^{-\lambda t} dt \\
&= \theta + \frac{a_1+a_2}{2\Phi}\int_0^\infty \left[\alpha + \frac{(1-\alpha)\,\delta}{x+\delta}\right]\lambda\sigma^2 e^{-\lambda\sigma^2 x} dx,
\end{aligned}$$

which implies (13) because $\Phi = -\frac{\alpha S_\varepsilon}{\gamma}$ and

$$\int_0^\infty \frac{\delta}{x+\delta} \lambda\sigma^2 e^{-\lambda\sigma^2 x} dx = \lambda\sigma^2 \delta e^{\lambda\sigma^2\delta} \Gamma\left(0, \lambda\sigma^2\delta\right).$$

The first line above follows from the law of iterated expectations, the second line from (11), the third and fourth lines from (7), the fifth line because t is exponentially distributed with parameter λ, the sixth line by rearranging, and the last line using $\delta = \alpha S_\varepsilon + \Phi^2 S_A$ and with the change of variable $x = t/\sigma^2$.

Following similar steps,

$$\begin{aligned}
&\mathbb{V}_C[P] \equiv \mathbb{E}\left[\left\{\tilde{P}_t\left(\hat{\theta}_t, \theta, a_1, a_2\right) - \mathbb{E}_C[P]\right\}^2 \mid \theta, a_1, a_2\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\left\{\tilde{P}_t\left(\hat{\theta}_t, \theta, a_1, a_2\right) - \mathbb{E}_C[P]\right\}^2 \mid t, \theta, a_1, a_2\right] \mid \theta, a_1, a_2\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\left\{(1-\beta(t))\hat{\theta}_t + \beta(t)\left(\theta + \frac{a_1+a_2}{2\Phi}\right)\right.\right.\right. \\
&\left.\left.\left. -\theta - \frac{a_1+a_2}{2\Phi}(1-\alpha)\lambda\sigma^2\delta e^{\lambda\sigma^2\delta}\Gamma\left(0,\lambda\sigma^2\delta\right)\right\}^2 \mid t, \theta, a_1, a_2\right] \mid \theta, a_1, a_2\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\left\{(1-\beta(t))\frac{z}{\sqrt{t/\sigma^2}}\right.\right.\right. \\
&\left.\left.\left. +\frac{a_1+a_2}{2\Phi}\left[\beta(t) - (1-\alpha)\lambda\sigma^2\delta e^{\lambda\sigma^2\delta}\Gamma\left(0,\lambda\sigma^2\delta\right)\right]\right\}^2 \mid t, \theta, a_1, a_2\right] \mid \theta, a_1, a_2\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[((1-\beta(t)))^2 \frac{\sigma^2}{t} + o(a_1+a_2) \mid t, \theta, a_1, a_2\right] \mid \theta, a_1, a_2\right] \\
&= \sigma^2 \mathbb{E}\left[((1-\beta(t)))^2 \frac{1}{t}\right] + o(a_1+a_2) \\
&= (1-\alpha)^2 \mathbb{E}\left[\left(\frac{t/\sigma^2}{t/\sigma^2 + \alpha S_\varepsilon + \Phi^2 S_A}\right)^2 \frac{1}{t/\sigma^2}\right] + o(a_1+a_2) \\
&= (1-\alpha)^2 \int_0^\infty \left[\left(\frac{t/\sigma^2}{t/\sigma^2 + \alpha S_\varepsilon + \Phi^2 S_A}\right)^2 \frac{1}{t/\sigma^2}\right] \lambda e^{-\lambda t} dt + o(a_1+a_2) \\
&\qquad = (1-\alpha)^2 \int_0^\infty \frac{x}{(x+\delta)^2} \lambda\sigma^2 e^{-\lambda\sigma^2 x} dx + o(a_1+a_2),
\end{aligned}$$

which implies (14) because

$$\int_0^\infty \frac{x}{(x+\delta)^2} \lambda\sigma^2 e^{-\lambda\sigma^2 x} dx = \lambda\sigma^2 \left\{-1 + e^{\lambda\sigma^2\delta}\left(1+\lambda\sigma^2\delta\right)\Gamma\left(0,\lambda\sigma^2\delta\right)\right\}.$$

A.3 Proof of Corollary 4

Define $B = \lambda\sigma^2\delta$ and $A = Be^B\Gamma(0, B)$. Then,

$$\frac{\partial A}{\partial B}(B) = e^B\Gamma(0,B) + Be^B\Gamma(0,B) + Be^B\frac{\partial\Gamma}{\partial B}(0,B) = (1+B)e^B\Gamma(0,B) - 1$$

because $\frac{\partial\Gamma}{\partial B}(0,B) = -\frac{1}{Be^B}$. Then, an application of the chain rule implies

$$\frac{\partial A}{\partial\lambda}(\lambda) = \sigma^2\delta\left\{\left(1+\lambda\sigma^2\delta\right)e^{\lambda\sigma^2\delta}\Gamma\left(0,\lambda\sigma^2\delta\right) - 1\right\},$$

and, in turn,

$$\frac{\partial APD}{\partial\lambda}(\lambda) = -\frac{\gamma}{S_\varepsilon}\frac{1-\alpha}{\alpha}\sigma^2\delta\left\{\left(1+\lambda\sigma^2\delta\right)e^{\lambda\sigma^2\delta}\Gamma\left(0,\lambda\sigma^2\delta\right) - 1\right\} < 0,$$

where $APD = \frac{\mathbb{E}_C[P]-\theta}{(a_1+a_2)/2}$ and the inequality follows because the term inside the curly bracket is positive for $\lambda > 0$.

Rewrite (14):

$$\mathbb{V}_C[P] = (1-\alpha)^2\frac{B}{\delta}\left\{-1 + e^B(1+B)\Gamma(0,B)\right\} + o(a_1+a_2).$$

Then,

$$\begin{aligned}
&\frac{\partial\mathbb{V}_C[P]}{\partial B}(B) = (1-\alpha)^2\frac{1}{\delta}\left\{-1 + e^B(1+B)\Gamma(0,B)\right\} \\
&+ (1-\alpha)^2\frac{B}{\delta}\left\{e^B(1+B)\Gamma(0,B) + e^B\Gamma(0,B) + e^B(1+B)\frac{\partial\Gamma}{\partial B}(0,B)\right\} + o(a_1+a_2) \\
&= (1-\alpha)^2\frac{1}{\delta}\left\{-1 + e^B(1+B)\Gamma(0,B)\right\} \\
&+ (1-\alpha)^2\frac{B}{\delta}\left\{e^B(2+B)\Gamma(0,B) - \left(1+\frac{1}{B}\right)\right\} + o(a_1+a_2) \\
&= (1-\alpha)^2\frac{1}{\delta}\left\{e^B(1+B)\Gamma(0,B) + (2+B)\left[-1 + Be^B\Gamma(0,B)\right]\right\} + o(a_1+a_2) \\
&= (1-\alpha)^2\frac{1}{\delta}\left\{-e^B\Gamma(0,B) + (2+B)\left[-1 + (1+B)e^B\Gamma(0,B)\right]\right\} + o(a_1+a_2) \\
&= (1-\alpha)^2\frac{1}{\delta}\left\{-2 + (1+2B)e^B\Gamma(0,B) + B\left[-1 + (1+B)e^B\Gamma(0,B)\right]\right\} + o(a_1+a_2).
\end{aligned}$$

Then, one can verify that

$$\lim_{B\to 0}\frac{\partial\mathbb{V}_C[P]}{\partial B}(B) = \infty$$

and that $\frac{\partial\mathbb{V}_C[P]}{\partial B}(B) < 0$ for $B \simeq \infty$ with

$$\lim_{B\to\infty}\frac{\partial\mathbb{V}_C[P]}{\partial B}(B) = 0,$$

which implies the Corollary because $B = \lambda\sigma^2\delta$.

A.4 Proof of Lemma 2

Using the fact that θ is normally distributed, one can show that

$$\begin{aligned}&\mathbb{E}\left[e^{-\frac{1}{2}\left(\mu_t^{\text{inf}}-\tilde{P}_t\left(\hat{\theta}_t,\theta,a_1,a_2\right)\right)^2 S_t^{\text{inf}}}\mid \hat{\theta}_t,t,\left\{\tilde{P}_t\left(\hat{\theta}_t,\theta,a_1,a_2\right)=P_t\right\}\right]\\&=\sqrt{\frac{S_\theta+t/\sigma^2+\Phi^2 S_A}{S_\theta+t/\sigma^2+S_\varepsilon+\Phi^2 S_A}}\mathbb{E}\left[e^{-\frac{1}{2}\left(\mu_t^{\text{uninf},a}-\tilde{P}_t\left(\hat{\theta}_t,\theta,a,a_-\right)\right)^2 S_t^{\text{uninf}}}\mid \hat{\theta}_t,t,\left\{\tilde{P}_t\left(\hat{\theta}_t,\theta,a_1,a_2\right)=P_t\right\}\right].\end{aligned}$$ [20]

Then, applying the law of iterated expectations,

$$\begin{aligned}&\mathbb{E}\left[e^{-\frac{1}{2}\left(\mu_t^{\text{inf}}-\tilde{P}_t\left(\hat{\theta}_t,\theta,a_1,a_2\right)\right)^2 S_t^{\text{inf}}}\mid t\right]\\&\qquad=\sqrt{\frac{S_\theta+t/\sigma^2+\Phi^2 S_A}{S_\theta+t/\sigma^2+S_\varepsilon+\Phi^2 S_A}}\mathbb{E}\left[e^{-\frac{1}{2}\left(\mu_t^{\text{uninf},a}-\tilde{P}_t\left(\hat{\theta}_t,\theta,a,a_-\right)\right)^2 S_t^{\text{uninf}}}\mid t\right].\end{aligned}$$

Substituting into (15) and further applying the law of iterated expectations, (15) becomes

$$\begin{aligned}&e^{\gamma C}\mathbb{E}\left[\mathbb{E}\left[e^{-\frac{1}{2}\left(\mu_t^{\text{inf}}-\tilde{P}_t\left(\hat{\theta}_t,\theta,a_1,a_2\right)\right)^2 S_t^{\text{inf}}}\mid t\right]\right]\\&\qquad\le\mathbb{E}\left[\sqrt{\frac{S_\theta+t/\sigma^2+S_\varepsilon+\Phi^2 S_A}{S_\theta+t/\sigma^2+\Phi^2 S_A}}\mathbb{E}\left[e^{-\frac{1}{2}\left(\mu_t^{\text{inf}}-\tilde{P}_t\left(\hat{\theta}_t,\theta,a_1,a_2\right)\right)^2 S_t^{\text{inf}}}\mid t\right]\right],\end{aligned}$$

where the outer expectation is taken over t.

Because the sum of normally distributed random variables is normally distributed as well, $\tilde{P}_t\left(\hat{\theta}_t,\theta,a_1,a_2\right)$ is normally distributed conditional on t. Using the fact that $\tilde{P}_t\left(\hat{\theta}_t,\theta,a_1,a_2\right)$ is conditionally normally distributed and using the law of iterated expectations,

$$\mathbb{E}\left[e^{-\frac{1}{2}\left(\mu_t^{\text{inf}}-\tilde{P}_t\left(\hat{\theta}_t,\theta,a_1,a_2\right)\right)^2 S_t^{\text{inf}}}\mid t\right]=\sqrt{\frac{S_\theta+t/\sigma^2}{S_\theta+t/\sigma^2+S_\varepsilon+\left(1+\frac{S_\theta+t/\sigma^2}{\Phi^2 S_A}\right)\left(\beta\left(t\right)\right)^2 S_\varepsilon}}.$$

Since t is exponentially distributed with parameter λ, the inequality stated in the Lemma obtains.

A.5 Proof of Proposition 5

Let

$$H=\frac{\int_0^\infty\sqrt{1+\frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta+\gamma^2 t/\sigma^2+\alpha^2 S_\varepsilon^2 S_A}}\omega\left(\alpha,t\right)\lambda e^{-\lambda t}dt}{\int_0^\infty\omega\left(\alpha,t\right)\lambda e^{-\lambda t}dt}-e^{\gamma C}.$$

[20] See Veldkamp (2011, p. 90) for a similar derivation.

To prove that there exists a unique equilibrium $\alpha^\star$, it suffices to show that H is a strictly decreasing function of α. To begin with, we calculate the first derivative of H with respect to α, applying *Lebesgue dominated convergence theorem*[21] whenever necessary:

$$\frac{\partial H}{\partial \alpha} = -\frac{1}{2}\frac{\int_0^\infty \left(1+\frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta+\gamma^2 t/\sigma^2+\alpha^2 S_\varepsilon^2 S_A}\right)^{-1/2}\frac{2\alpha\gamma^2 S_\varepsilon^3 S_A}{(\gamma^2 S_\theta+\gamma^2 t/\sigma^2+\alpha^2 S_\varepsilon^2 S_A)^2}\omega(\alpha,t)\lambda e^{-\lambda t}dt}{\int_0^\infty \omega(\alpha,t)\lambda e^{-\lambda t}dt}$$
$$+\frac{\int_0^\infty \sqrt{1+\frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta+\gamma^2 t/\sigma^2+\alpha^2 S_\varepsilon^2 S_A}}\,\omega_\alpha(\alpha,t)\lambda e^{-\lambda t}dt}{\int_0^\infty \omega(\alpha,t)\lambda e^{-\lambda t}dt}$$
$$-\frac{\int_0^\infty \sqrt{1+\frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta+\gamma^2 t/\sigma^2+\alpha^2 S_\varepsilon^2 S_A}}\,\omega(\alpha,t)\lambda e^{-\lambda t}dt\int_0^\infty \omega_\alpha(\alpha,t)\lambda e^{-\lambda t}dt}{\left(\int_0^\infty \omega(\alpha,t)\lambda e^{-\lambda t}dt\right)^2}$$
$$\leq -\frac{1}{2}\frac{\int_0^\infty \left(1+\frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta+\gamma^2 t/\sigma^2+\alpha^2 S_\varepsilon^2 S_A}\right)^{-1/2}\frac{2\alpha\gamma^2 S_\varepsilon^3 S_A}{(\gamma^2 S_\theta+\gamma^2 t/\sigma^2+\alpha^2 S_\varepsilon^2 S_A)^2}\omega(\alpha,t)\lambda e^{-\lambda t}dt}{\int_0^\infty \omega(\alpha,t)\lambda e^{-\lambda t}dt}$$
$$+\frac{\sqrt{1+\frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta+\alpha^2 S_\varepsilon^2 S_A}}\int_0^\infty \omega_\alpha(\alpha,t)\lambda e^{-\lambda t}dt}{\int_0^\infty \omega(\alpha,t)\lambda e^{-\lambda t}dt}-\frac{\int_0^\infty \omega_\alpha(\alpha,t)\lambda e^{-\lambda t}dt}{\int_0^\infty \omega(\alpha,t)\lambda e^{-\lambda t}dt},$$

where $\omega_\alpha(\cdot,\cdot)$ refers to the first derivative of $\omega(\cdot,\cdot)$ with respect to its first argument. Thus, to establish $\frac{\partial H}{\partial \alpha}<0$, it suffices to show that $\omega_\alpha(\cdot,t)<0$ for all $t\geq 0$.

Using (16) and Proposition 1,

$$\omega_\alpha(\alpha,t)=\frac{(\omega(\alpha,t))^3(\beta(t))^2}{(S_\theta+t/\sigma^2)\alpha^3 S_\varepsilon S_A}\left(\gamma^2\left(S_\theta+t/\sigma^2\right)-\alpha\left[\gamma^2\left(S_\theta+t/\sigma^2\right)+\alpha^2 S_\varepsilon^2 S_A\right]\frac{d\log\beta(t)}{d\alpha}\right)$$

and

$$\begin{aligned}\frac{d\log\beta(t)}{d\alpha}=&\frac{\gamma^2(S_\theta+t/\sigma^2+S_\varepsilon)+2\alpha S_\varepsilon^2 S_A}{\alpha\gamma^2(S_\theta+t/\sigma^2+S_\varepsilon)+\alpha^2 S_\varepsilon^2 S_A}-\frac{\gamma^2 S_\varepsilon+2\alpha S_\varepsilon^2 S_A}{\gamma^2(S_\theta+t/\sigma^2+\alpha S_\varepsilon)+\alpha^2 S_\varepsilon^2 S_A}\\=&\frac{\gamma^2(S_\theta+t/\sigma^2)(\alpha\gamma^2(S_\theta+t/\sigma^2+S_\varepsilon)+(2-\alpha)\alpha^2 S_\varepsilon^2 S_A)}{\alpha(\alpha\gamma^2(S_\theta+t/\sigma^2+S_\varepsilon)+\alpha^2 S_\varepsilon^2 S_A)(\gamma^2(S_\theta+t/\sigma^2+\alpha S_\varepsilon)+\alpha^2 S_\varepsilon^2 S_A)}.\end{aligned}$$

[21] See, for a reference, Hutson, Pym, and Cloud (2005, p. 55).

In turn, it suffices to show that

$$\omega_\alpha(\alpha,t) = \frac{(\omega(\alpha,t))^3 (\beta(t))^2 \gamma^2}{\alpha^2 S_A} \times \frac{\gamma^4 (S_\theta + t/\sigma^2 + S_\varepsilon) + \gamma^2 S_\varepsilon S_A (\alpha S_\varepsilon - (1-\alpha)(S_\theta + t/\sigma^2)) - (1-\alpha)\alpha^2 S_\varepsilon^3 S_A^2}{(\gamma^2 (S_\theta + t/\sigma^2 + S_\varepsilon) + \alpha S_\varepsilon^2 S_A)(\gamma^2 (S_\theta + t/\sigma^2 + \alpha S_\varepsilon) + \alpha^2 S_\varepsilon^2 S_A)} < 0.$$

Because we want that this inequality is satisfied for all $\alpha \in (0,1)$, via continuity, a necessary condition emerges when $\alpha = 0$:

$$\frac{S_\varepsilon S_A (S_\theta + t/\sigma^2)}{S_\theta + t/\sigma^2 + S_\varepsilon} > \gamma^2,$$

which is implied by the condition stated in the Proposition as the LHS is increasing in t, establishing the uniqueness of the equilibrium $\alpha^\star$:

$$\alpha^\star = \begin{cases} 0 & \text{if } \alpha_s \leq 0 \\ \alpha_s & \text{if } \alpha_s \in (0,1) \\ 1 & \text{if } \alpha_s \geq 1 \end{cases}$$

where α_s solves $H = 0$.

To obtain the comparative statics in the Proposition, we make use of the implicit function theorem:

$$\frac{\partial \alpha}{\partial \sigma} = -\frac{\frac{\partial H}{\partial \sigma}}{\frac{\partial H}{\partial \alpha}}, \quad \frac{\partial \alpha}{\partial \lambda} = -\frac{\frac{\partial H}{\partial \lambda}}{\frac{\partial H}{\partial \alpha}}, \text{ and } \frac{\partial \alpha}{\partial C} = -\frac{\frac{\partial H}{\partial C}}{\frac{\partial H}{\partial \alpha}}.$$

We have already established that $\frac{\partial H}{\partial \alpha} < 0$. To complete the comparative statics, we calculate the remaining derivatives, applying Lebesgue dominated convergence theorem whenever necessary:

$$\frac{\partial H}{\partial \sigma} = \frac{\int_0^\infty \left(1 + \frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta + \gamma^2 t/\sigma^2 + \alpha^2 S_\varepsilon^2 S_A}\right)^{-1/2} \frac{\gamma^2 S_\varepsilon}{(\gamma^2 S_\theta + \gamma^2 t/\sigma^2 + \alpha^2 S_\varepsilon^2 S_A)^2} \frac{\gamma^2 t}{\sigma^3} \omega(\alpha,t) e^{-\lambda t} dt}{\int_0^\infty \omega(\alpha,t) e^{-\lambda t} dt}$$

$$+ \frac{\int_0^\infty \sqrt{1 + \frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta + \gamma^2 t/\sigma^2 + \alpha^2 S_\varepsilon^2 S_A}} \frac{d\omega(\alpha,t)}{d\sigma} e^{-\lambda t} dt \int_0^\infty \omega(\alpha,t) e^{-\lambda t} dt}{\left(\int_0^\infty \omega(\alpha,t) e^{-\lambda t} dt\right)^2}$$

$$- \frac{\int_0^\infty \sqrt{1 + \frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta + \gamma^2 t/\sigma^2 + \alpha^2 S_\varepsilon^2 S_A}} \omega(\alpha,t) e^{-\lambda t} dt \int_0^\infty \frac{d\omega(\alpha,t)}{d\sigma} e^{-\lambda t} dt}{\left(\int_0^\infty \omega(\alpha,t) e^{-\lambda t} dt\right)^2} > 0,$$

where the inequality follows because $\frac{d\omega(\alpha,t)}{d\sigma} < 0$, which is immediately obvious from the following re-written version of $\omega(\alpha,t)$:

$$\omega(\alpha,t) = \sqrt{\frac{1}{1 + \frac{S_\varepsilon}{S_\theta + t/\sigma^2} + \left(\frac{\gamma^2}{S_\varepsilon S_A} + \frac{\alpha^2 S_\varepsilon}{S_\theta + t/\sigma^2}\right)\left(\frac{\alpha S_\varepsilon^2 S_A + \gamma^2 (S_\theta + t/\sigma^2 + S_\varepsilon)}{\alpha S_\varepsilon^2 S_A + \gamma^2 (S_\theta + t/\sigma^2 + \alpha S_\varepsilon)}\right)^2}},$$

$$\begin{aligned}
\frac{\partial H}{\partial \lambda} =& \frac{-\int_0^\infty \sqrt{1 + \frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta + \gamma^2 t/\sigma^2 + \alpha^2 S_\varepsilon^2 S_A}}\,\omega(\alpha,t)\, e^{-\lambda t} t\, dt \int_0^\infty \omega(\alpha,t')\, e^{-\lambda t'} dt'}{\left(\int_0^\infty \omega(\alpha,t)\, e^{-\lambda t} dt\right)^2} \\
&+ \frac{\int_0^\infty \sqrt{1 + \frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta + \gamma^2 t/\sigma^2 + \alpha^2 S_\varepsilon^2 S_A}}\,\omega(\alpha,t)\, e^{-\lambda t} dt \int_0^\infty \omega(\alpha,t')\, e^{-\lambda t'} t' dt'}{\left(\int_0^\infty \omega(\alpha,t)\, e^{-\lambda t} dt\right)^2} \\
=& \frac{\int_0^\infty \sqrt{1 + \frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta + \gamma^2 t/\sigma^2 + \alpha^2 S_\varepsilon^2 S_A}}\,\omega(\alpha,t)\, e^{-\lambda t} \left[\int_0^\infty \omega(\alpha,t')\, e^{-\lambda t'} (t' - t)\, dt'\right] dt}{\left(\int_0^\infty \omega(\alpha,t)\, e^{-\lambda t} dt\right)^2} > 0
\end{aligned}$$

because $\sqrt{1 + \frac{\gamma^2 S_\varepsilon}{\gamma^2 S_\theta + \gamma^2 t/\sigma^2 + \alpha^2 S_\varepsilon^2 S_A}}$ is decreasing in t, and

$$\frac{\partial H}{\partial C} = -\gamma e^{\gamma C} < 0,$$

which gives us the comparative statics stated in the Proposition.

B Additional tables

Table 4: Informational Efficiency and Trade Volume

	(1)	(2)	(3)	(4)	(5)	(6)
	Return Autocorrelation			Price Delay		
	All Bonds	Less Liquid	More Liquid	All Bonds	Less Liquid	More Liquid
Panel A: No Controls						
λ	-0.035***	-0.046***	-0.034***	-0.048***	-0.035***	-0.073***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Constant	0.757***	0.948***	0.724***	1.154***	1.025***	1.471***
	(0.08)	(0.16)	(0.10)	(0.09)	(0.15)	(0.12)
N	1080	541	539	1419	711	708
R^2	0.037	0.035	0.061	0.040	0.016	0.125
Panel B: With Controls						
λ	-0.033***	-0.046***	-0.033***	-0.032***	-0.024**	-0.059***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Inv. Grade	0.012	0.041**	-0.005	-0.131***	-0.124***	-0.097***
	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Maturity=Q2	-0.111***	-0.078***	-0.138***	-0.082***	-0.082***	-0.074***
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
Maturity=Q3	-0.129***	-0.104***	-0.160***	-0.122***	-0.137***	-0.104***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Maturity=Q4	-0.163***	-0.165***	-0.169***	-0.146***	-0.174***	-0.124***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	0.826***	1.024***	0.828***	1.098***	1.027***	1.409***
	(0.08)	(0.15)	(0.09)	(0.10)	(0.15)	(0.12)
N	1080	541	539	1419	711	708
R^2	0.156	0.125	0.258	0.170	0.167	0.211

Notes: this table presents results from cross-sectional regressions whereby measures of price efficiency are regressed on trade volume of corporate bonds. Panel A presents the results without controls, and Panel B presents the results with controls. Columns 1-3 show the results where the absolute value of return autocorrelation (computed by (21)) is used as the left-hand side variables. Columns 4-6 show the results where price delay (computed by (22)) is used as the regressand. Trade volume (the regressor) is measured as the natural logarithm of the daily sum of nominal value of transactions in bond i. For a bond to be included in the time-series regressions ((21) or (22)), it has to have at least 20 daily observations. Less (more) liquid bonds are defined as those that have approximately less (more) than the median number of daily transactions. The sample covers the period from August 2011 to December 2017. The trade-level dataset covers about 3.8 million transactions, including both client-dealer and inter-dealer trades. T-statistics in parentheses are based on robust standard errors. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 5: Price Dispersion and Trade Volume

	(1)	(2)	(3)
Panel A: Weighted Dispersion			
λ	0.688***	1.992***	1.986***
	(0.10)	(0.07)	(0.07)
λ^2	-0.000***	-0.000***	-0.000***
	(0.00)	(0.00)	(0.00)
N	625927	625665	625662
R^2	0.004	0.192	0.231
Panel B: Unweighted Dispersion			
λ	0.652***	2.022***	2.018***
	(0.10)	(0.06)	(0.06)
λ^2	-0.000***	-0.000***	-0.000***
	(0.00)	(0.00)	(0.00)
N	625927	625665	625662
R^2	0.004	0.206	0.246
Bond FE	No	Yes	Yes
Day FE	No	No	Yes

Notes: this table presents results from panel regressions whereby measures of price dispersions are regressed on trade volume of corporate bonds. Trade volume (the regressor) is measured as the natural logarithm of the daily sum of nominal value of transactions in bond i. The regressor is included in a linear and a quadratic way. The sample covers the period from August 2011 to December 2017. The trade-level dataset covers about 3.8 million transactions, including both client-dealer and inter-dealer trades. T-statistics in parentheses are based on robust standard errors, where we also apply two-way clustering at the bond- and day-level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).