Price impact equilibrium with transaction costs and TWAP trading

Eunjung Noh

Florida State University

Department of Mathematics

Kim Weston[1]

Rutgers University

Department of Mathematics

August 24, 2021

Abstract

We prove the existence of an equilibrium in a model with transaction costs and price impact where two agents are incentivized to trade towards a target. The two types of frictions – price impact and transaction costs – lead the agents to two distinct changes in their optimal investment approach: price impact causes agents to continuously trade in smaller amounts, while transaction costs cause the agents to cease trading before the end of the trading period. As the agents lose wealth because of transaction costs, the exchange makes a profit. We prove the existence of a strictly positive optimal transaction cost from the exchange's perspective.

Keywords: Transaction costs, Price impact, Equilibrium, Targeted trading, TWAP, Incompleteness, Trading frictions

[1]The authors would like to thank Xiao Chen, Jetlir Duraj, and Kasper Larsen for helpful discussions on this work. We would also like to thank the anonymous referees and associate editor for their comments, which greatly improved the paper. The second author acknowledges support by the National Science Foundation under Grant No. DMS#1908255 (2019-2022). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation (NSF).

1 Introduction

We study a financial equilibrium model with frictions stemming from both transaction costs and price impact. Two agents are incentivized to trade towards a targeted number of shares throughout the trading period. In equilibrium, the agents seek to maximize their expected wealth minus a penalty for deviating from their targets. Their wealth is further reduced by transaction costs and is affected by the perceived price impact on the stock price from their trades.

Incomplete equilibrium is notoriously difficult to study. When the incompleteness stems from frictions, this difficulty is exacerbated. This work proposes a tractable model for a financial equilibrium with two simultaneous frictions. We answer two questions:

(1) *How do transaction costs and price impact affect prices and strategies in equilibrium?*

(2) *What is the optimal level of transaction costs?*

The main contribution of this work is to show that the effect of transaction costs in equilibrium is distinct from that of price impact. Price impact affects the equilibrium stock price and depresses the rate of trade, whereas transaction costs cause the agents to cease trading early in the trading period. We also prove the existence of a strictly positive optimal level of transaction costs, where optimality is determined from the perspective of the market's exchange. The exchange collects the transaction fees as the agents trade.

The two frictions that we focus on are transaction costs and price impact. When studied individually, transaction cost equilibrium models often resort to simplifications in order to draw conclusions. Our model shares this approach, as it is simple enough to remain tractable while complex enough to capture differing effects of our two frictions. Weston [27] proves the existence of a transaction cost equilibrium in a tractable model with deterministic equilibrium annuity prices. Continuum-of-agent models, where market clearing is averaged over infinitely many agents, are studied in Vayanos and Vila [25], Vayanos [23], Huang [17], and Dávila [12]. Herdegen and Muhle-Karbe [16] study equilibrium with transaction costs where clearing holds approximately, up to a leading order. For proportional transaction costs, Gonon et. al. [14] study equilibrium with an ergodic objective. Lo et. al. [20] and Buss et. al. [8] study the numerics behind transaction cost equilibria without establishing the existence of an equilibrium. In contrast, we prove the existence of an equilibrium with proportional transaction costs (and price impact) in a standard setting with two agents, market clearing, and consumption at a terminal time.

Price impact and optimal liquidation models for a single agent take the price impact form as given and allow agents to consider how the size and timing of their trades will impact the traded asset price and, hence, future wealth. Our equilibrium model endogenizes price impact, which is realized in the stock's drift in the form proposed in Cuoco and Cvitanić [11]. We incorporate price impact in our model by modeling it through perceived off-equilibrium price paths. Perceived off-equilibrium paths describe how prices react to trading strategies even though a given trading strategy may be suboptimal. The optimal trading strategy determines the on-equilibrium price path. Both agents' on-equilibrium price dynamics must coincide with the equilibrium stock price. In this way, price impact affects equilibrium prices, even though it is only modeled through each individual agent's perceived off-equilibrium price paths.

Other models have endogenized price impact in an equilibrium setting using various approaches. Kyle [19] uses a game theoretic framework, where the market maker attempts to filter out private information from the aggregate trades of noise traders and an informed trader. Vayanos [24] and Sannikov and Skrzypacz [21] use private information to endogenize price impact. Choi et. al. [10] endogenize price impact by allowing perceived off-equilibrium price paths to vary in a Cuoco and Cvitanić [11]-sense while also allowing their agents' trading targets to be private.

Our modeling set-up is most similar to Choi et. al. [10], who study targeted trading in Nash equilibria with price impact. We introduce transaction costs into a simple form of their setting in order to compare the effects of both frictions (price impact and transaction costs) in equilibrium. Our model shares similarities with others in terms of single-agent models and equilibrium settings. Linear-quadratic models with trading targets are studied in several works, including Bank et. al. [4], Sannikov and Skrzypacz [21], and Voß [26]. Gonon et. al. [14] use linear-quadratic controls and allow for proportional transaction costs, rather than a quadratic approximation to transaction costs, as in Brunnermeier and Pedersen [7], Gârleanu and Pedersen [13], and Bouchard et. al. [5].

Similar to the two-agent settings of Weston [27], Herdegen and Muhle-Karbe [16], and Gonon et. al. [14], our two agents share a filtration, which reveals each of their trading targets to the other agent. This modeling set-up leads to a fully-revealing equilibrium since the agents' aggregate trading target is revealed at the initial time by the initial equilibrium stock price. Hence, each agent can infer the other agents' target.

Every transaction fee paid by the agents is income for the market's exchange. Higher transaction costs generate more income for the exchange for every share traded, but lower transaction costs induce the agents to trade a higher volume of shares. Consequently, we can prove that there exists a strictly positive level of transaction costs that maximizes the exchange's expected profit. Optimal transaction costs have been of interest starting most notably with the introduction of the Tobin tax in Tobin [22]. Previous equilibrium approaches consider optimality from a welfare perspective. The continuum of agents in Dávila [12] differ in their beliefs about the dividend's distribution. The agents' belief difference versus the central planner's choice of distribution when calculating welfare leads to a strictly positive optimal transaction tax. In Weston [27], the welfare decreases as transaction costs increase, leading to zero as the welfare optimizing transaction cost parameter. In our model, the agents are identical in their beliefs and differ only in their trading targets. Nonetheless, we prove the existence of a strictly positive optimal transaction cost from the exchange's perspective.

The paper is organized as follows. Section 2 describes our model inputs. Section 3 presents our main result, Theorem 3.3, which establishes the existence of a price impact equilibrium with transaction costs. The choice of an optimal transaction cost from the perspective of the exchange is presented in Section 4. The proofs are contained in Section 5.

2 The model

Let $T > 0$ be a fixed time horizon, which we think of as one trading day in length. We work in a continuous-time setting and let $B = \{B_t\}_{t\in[0,T]}$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The market consists of two traded securities: a bank account and a stock. The bank account is a financial asset in zero-net supply with a constant zero interest rate. The stock is in constant positive net supply with the supply denoted by $n > 0$. Since the time horizon is thought to be small, the stock does not pay any dividends over the period $[0, T]$.[2] Instead, the stock has an exogenous volatility denoted by σ. Such models are common in the equilibrium literature; see, for example, Chapter

[2]Even in the case of a longer time horizon, Hartzmark and Solomon [15], Boudoukh et. al. [6], and Atmaz and Basak [3] make the case that the non-dividend paying stocks are prevalent in the stock market and potentially play a prominent role in longer-time horizon asset pricing models.

4 of Karatzas and Shreve [18]. The volatility is progressively measurable with respect to $\{\sigma(B_u : 0 \le u \le t)\}_{t\in[0,T]}$ and $\mathbb{E}\int_0^T \sigma_u^2 du < \infty$. We assume that all prices are denominated in a single consumption good.

Two investors, $i = 1, 2$, trade in the market. They each seek to maximize expected wealth yet are subjected to inventory penalties throughout the trading period. Their wealth is further penalized by transaction costs, which are proportional to the rate of trade at the rate $\lambda > 0$. Their wealth is indirectly penalized by perceived price impact from trades. For every share purchased, the agents perceive that the stock's drift decreases linearly.

Each agent i has a target number of shares a_i she wishes to acquire (or sell off) throughout the trading period. The random variables a_1 and a_2 are assumed to satisfy $\mathbb{E}[a_i^2] < \infty$ and be independent of the Brownian motion B. The filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$ is given by

$$\mathcal{F}_t := \sigma\left(a_1, a_2, B_u : u \in [0, t]\right), \qquad t \in [0, T].$$

All market participants have access to the same filtration. All processes are adapted to $\mathbb{F}$.

A trading strategy $\theta = \{\theta_t\}_{t\in[0,T]}$ denotes the number of shares held in stock. We say that θ is *admissible* if it is adapted to $\mathbb{F}$, càdlàg, of finite variation on $[0, T]$ $\mathbb{P}$-a.s.,[3] and satisfies $\mathbb{E}\int_0^T (\sigma_t\theta_t)^2\, dt < \infty$. We write $\mathcal{A}$ to denote the collection of admissible strategies. Agent i is endowed at the beginning of the trading period with $n/2$ shares of stock. We normalize the shares in the bank account to zero. We allow for θ_0 to differ from $n/2$, as the agents may choose to trade a lump sum immediately. In the absence of transaction costs or the penalty term given in (2.3) below, the agents' allocations would be Pareto optimal. However, the presence of frictions and penalties motivates the agents to deviate from their initial positions.

Since $\theta \in \mathcal{A}$ is of finite variation, we can decompose θ into

$$\theta_t = \frac{n}{2} + \theta_t^{\uparrow} - \theta_t^{\downarrow}, \qquad t \in [0, T], \tag{2.1}$$

where $\theta^{\uparrow}, \theta^{\downarrow}$ are adapted to $\mathbb{F}$, càdlàg, nondecreasing, and

$$\{t \in [0, T] : d\theta_t^{\uparrow} > 0\} \cap \{t \in [0, T] : d\theta_t^{\downarrow} > 0\} = \emptyset. \tag{2.2}$$

[3] We restrict strategies to those with finite variation since strategies with infinite first-order variation would result in infinite transaction costs.

A change in trading position is possible at time 0, and we allow for $\theta_0^{\uparrow} > 0$ or $\theta_0^{\downarrow} > 0$ as long as (2.2) holds.

At the close of the trading period, agents consume their acquired wealth. The agents are subjected through their optimization problems to inventory penalties throughout the trading period. For $i \in \{1, 2\}$ and a given $\theta \in \mathcal{A}$, the penalty term, or loss term, for agent i is measured by

$$L_{i,T}^{\theta} := \frac{1}{2} \int_0^T \kappa(t) \left(\gamma(t) \left(a_i - \frac{n}{2} \right) - \left(\theta_t - \frac{n}{2} \right) \right)^2 dt. \tag{2.3}$$

The function $\kappa : (0, T) \to (0, \infty)$ describes the intensity of the penalty, while $\gamma : [0, T] \to [0, 1]$ describes the desired intraday trading target trajectory. Both agents share the same deterministic functions κ and γ. We assume that γ is càdlàg, nonnegative, bounded in absolute value by one, and nondecreasing. Our main example is time-weighted average price (TWAP), where the intraday trajectory function is $\gamma^{\text{TWAP}}(t) := t/T$. We assume that κ is measurable and $\int_0^T \kappa(t)dt$ is finite. This penalty term serves to motivate agent i to trade towards the target a_i. Whereas κ is a weighting function, γ describes how quickly throughout the trading period the agent is expected to move from the initial $n/2$ shares towards a_i shares at the period's end.

The agents perceive a price impact as the result of their trades. For $i = 1, 2$ and a trading strategy $\theta \in \mathcal{A}$, we model this impact via the perceived off-equilibrium stock price's drift by

$$dS_{i,t}^{\theta} = \kappa(t) \left(c_0(t, a_1 + a_2) - c_1 \theta_t + \gamma(t) c_2 \left(a_i - \frac{n}{2} \right) \right) dt + \sigma_t \, dB_t, \quad S_0 \in \mathcal{F}_0. \tag{2.4}$$

The function c_0 and constants c_1, c_2 will be determined in equilibrium and are the same for both agents. The constant c_1 turns out to be a free parameter that determines the level of price impact in the market. The $c_1 = 0$ case corresponds to an equilibrium without price impact, where the agents are price-takers.

For fixed constants c_0, c_1, and c_2, the perceived stock price is determined from a given traded strategy. In equilibrium, we will require each agent's optimal perceived stock price to be consistent with the equilibrium stock price. The perceived price varies with the trading strategy, and so perceived off-equilibrium prices are not forced to agree with the equilibrium stock price or another agent's perceptions. Following the work of Choi et. al. [10] without transaction costs, we work with perceived off-equilibrium stock prices whose martingale term $\int \sigma dB$ and initial value S_0 are independent of θ and i. In particular, this choice implies that the perceived initial stock prices always agree, but the perceived terminal stock prices may not agree.

Price impact in (2.4) is realized through the drift of the stock as in Cuoco and Cvitanić [11]. Larger values for c_1 result in more price impact because the more an agent buys, the more she drives her perceived future prices down. While traditional price impact models, such as Almgren and Chriss [1], affect the stock price directly, our version of price impact affects the future returns of the stock by depressing them when a trader seeks a larger market share.

For agent $i \in \{1,2\}$ and a trading strategy $\theta \in \mathcal{A}$, agent i's perceived wealth process is given by

$$X^{\theta}_{i,t} = \frac{n}{2} S_0 + \int_0^t \theta_u dS^{\theta}_{iu} - \lambda \left(\theta^{\uparrow}_t + \theta^{\downarrow}_t \right), \qquad t \in [0,T]. \tag{2.5}$$

We recall that the decomposition of θ in (2.1) allows for $\theta^{\uparrow}_0$ and $\theta^{\downarrow}_0$ to differ from zero. Both frictions – price impact and transaction costs – are at play in the perceived wealth dynamics. Agent i's objective is

$$\mathbb{E}\left[X^{\theta}_{i,T} - L^{\theta}_{i,T} \,|\, \mathcal{F}_0 \right] \longrightarrow \max$$

over $\theta \in \mathcal{A}$, where $L^{\theta}_{i,T}$ is defined in (2.3) and $X^{\theta}_{i,T}$ in (2.5).

3 Equilibrium

In an equilibrium, the stock price is determined so that markets clear when both agents invest optimally. The equilibrium stock price must agree with both agents' perceived prices when the optimal strategies are applied.

Definition 3.1. Let $\lambda > 0$ be a given transaction cost level. Trading strategies $\theta_1, \theta_2 \in \mathcal{A}$, a price process $\widehat{S} = \{\widehat{S}_t\}_{t\in[0,T]}$, and price impact coefficients c_0, c_1, c_2 form a *price impact equilibrium* if

(a) *Strategies are optimal:* For $i = 1, 2$, we have that

$$\mathbb{E}\left[X^{\theta_i}_{i,T} - L^{\theta_i}_{i,T} \,|\, \mathcal{F}_0 \right] = \sup_{\theta\in\mathcal{A}} \mathbb{E}\left[X^{\theta}_{i,T} - L^{\theta}_{i,T} \,|\, \mathcal{F}_0 \right], \tag{3.1}$$

where $L^{\theta}_{i,T}$ is defined in (2.3), $X^{\theta}_{i,T}$ in (2.5), and the perceived off-equilibrium price impact stock dynamics are given in (2.4) with coefficients c_0, c_1, and c_2.

(b) *Markets clear:* We have $\theta_{1,t} + \theta_{2,t} = n$ for all $t \in [0,T]$.

(c) *Prices are consistent:* The equilibrium stock price process $\widehat{S}$ is an Itô process, and for $t \in \{u \in [0,T] : d\theta^{\uparrow}_{1,u} + d\theta^{\downarrow}_{1,u} > 0\}$, we have that

$$S^{\theta_1}_{1,t} = S^{\theta_2}_{2,t} = \widehat{S}_t.$$

The price impact stock dynamics of $S^{\theta_1}_1$ and $S^{\theta_2}_2$ are given in (2.4) with coefficients c_0, c_1, and c_2.

Even though off-equilibrium, the agents perceive a price impact from their trades, the on-equilibrium stock price must agree with the agents' perceived prices when their optimal strategies are applied. Definition 3.1(c) requires the perceived prices to agree with the realized equilibrium price *when trade occurs*. (Since there are only two agents in the model, trade occurs if and only if agent 1 trades.) Therefore, in equilibrium, the two agents may have different perceived prices at times when they do not trade. This requirement on perceived prices in equilibrium is similar to employing shadow prices in equilibrium since an equilibrium stock price can only be uniquely identified when trade occurs; see Weston [27].

Market clearing in Definition 3.1(b) requires clearing of the stock market, however Walras' Law holds in our model in that the other markets (bank account and real goods) clear as well. For a given strategy $\theta \in \mathcal{A}$ and equilibrium stock price $\widehat{S}$, we define the realized wealth in equilibrium through its self-financing condition (see (2.5)) by

$$\widehat{X}^{\theta}_t := \frac{n}{2}\widehat{S}_0 + \int_0^t \theta_u d\widehat{S}_u - \lambda\left(\theta^{\uparrow}_t + \theta^{\downarrow}_t\right), \qquad t \in [0,T].$$

We define the corresponding holdings in the bank account by

$$\theta^{(0)}_t := \widehat{X}^{\theta}_t - \theta_t \widehat{S}_t, \qquad t \in [0,T]. \tag{3.2}$$

We recall that the interest rate has been taken to be zero, since consumption only occurs at one point in time.

Lemma 3.2 presents a version of Walras' Law applied to a price impact equilibrium with transaction costs. Its proof is presented in Section 5. Lemma 3.2 shows that the bank account provides the mechanism by which transaction costs exit the economy. The transaction costs paid in equation (3.3) below go to the exchange.

Lemma 3.2. *For a given transaction cost $\lambda > 0$, let a price impact equilibrium satisfying Definition 3.1 be given with optimal stock holdings θ_1, θ_2 and equilibrium stock price $\widehat{S}$. For $i = 1, 2$, we let $\theta^{(0)}_{i,t}$ correspond to the equilibrium bank account holdings at t of agent i with stock market strategy θ_i, as in* (3.2)*. Then the bank clears in the sense that*

$$\theta^{(0)}_{1,t} + \theta^{(0)}_{2,t} = -\lambda \left(\theta^{\uparrow}_{1,t} + \theta^{\downarrow}_{1,t} + \theta^{\uparrow}_{2,t} + \theta^{\downarrow}_{2,t} \right), \qquad t \in [0, T]. \tag{3.3}$$

To begin constructing our equilibrium, for each $i \in \{1, 2\}$, we let

$$a_\Sigma := a_1 + a_2 \quad \text{and} \quad A_i := a_i - \frac{1}{2} a_\Sigma.$$

The random variables A_i describe the deviation of the trading targets a_i from the aggregate target a_Σ. We note that $A_1 + A_2 = 0$.

The presence of transaction costs causes the agents to stop trading before the end of the trading period. To this end, we define the last trading time τ by

$$\tau := \inf \left\{ t \in [0, T] : \, |A_1| \frac{1 + 2c_1}{1 + c_1} \int_t^T \kappa(u) \Big(\gamma(u) - \gamma(t) \Big) du \leq \lambda \right\}. \tag{3.4}$$

The time τ is a random variable valued in $[0, T)$. We provide the motivation and context for τ's definition in Section 5 below. We also define the random variable χ by

$$\chi := \frac{|A_1|(1 + 2c_1)}{1 + c_1} \int_0^T \kappa(u) \gamma(u) du. \tag{3.5}$$

The magnitude of χ will determine if trade occurs in the model or if the agents are deterred from trading by prohibitively high transaction costs.

The following theorem is our main result. The proof of Theorem 3.3 can be found in Section 5.

Theorem 3.3. *Let $\lambda > 0$ and $c_1 > -\frac{1}{2}$ be given. Suppose that $\kappa : (0, T) \to (0, \infty)$ is measurable with $\int_0^T \kappa(u) du < \infty$ and that $\gamma : [0, T] \to [0, 1]$ is càdlàg, nonnegative, bounded by one, and nondecreasing. There exists a price impact equilibrium where the* *price impact stock dynamics in* (2.4) *have coefficients* c_0 *and* c_2 *given in terms of* c_1 *by*

$$c_0(t, a_\Sigma) := c_1 n - \frac{1+2c_1}{2(1+c_1)}\gamma(t)(a_\Sigma - n) \quad \textit{and} \quad c_2 := \frac{c_1}{1+c_1}. \tag{3.6}$$

For $i = 1, 2$*, the equilibrium holdings* $\theta_i \in \mathcal{A}$ *are given for* $t \in [0, T]$ *by*

$$\theta_{i,t} := \begin{cases} \frac{n}{2} + \frac{A_i}{1+c_1}\gamma(t), & t < \tau, \\ \frac{n}{2} + \frac{1}{\int_\tau^T \kappa(u)du}\left\{\int_\tau^T \kappa(u)\gamma(u)\frac{A_i}{1+c_1}du - \frac{\lambda sign(A_i)}{1+2c_1}\right\}, & t \geq \tau \textit{ and } \chi > \lambda, \\ \frac{n}{2}, & t \geq \tau \textit{ and } \chi \leq \lambda, \end{cases} \tag{3.7}$$

where τ *is defined in* (3.4)*,* χ *is defined in* (3.5)*, and the sign function convention is* $sign(0) = 0$ *so that*

$$sign(x) := \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ 1, & x > 0. \end{cases}$$

We let $\widehat{S} = \{\widehat{S}_t\}_{t\in[0,T]}$ *be defined by*

$$\widehat{S}_t := \int_0^t \sigma_u dB_u + \frac{1}{2}\int_t^T \kappa(u)\Big(\gamma(u)(a_\Sigma - n) - c_1 n\Big)du. \tag{3.8}$$

Then, $\widehat{S}$*,* θ_1*,* θ_2*,* c_0*,* c_1*, and* c_2 *form a price impact equilibrium.*

For a given transaction cost parameter $\lambda > 0$, Theorem 3.3 shows that equilibrium is not unique. Indeed, there exists a distinct equilibrium for every choice of price impact coefficient $c_1 > -\frac{1}{2}$. When both $\lambda > 0$ and $c_1 > -\frac{1}{2}$ are fixed, the equilibrium is unique within the class of models that are fully revealing and have perceived stock price dynamics as in (2.4). Uniqueness is not immediately obvious because there appears to be ambiguity in the stock price (and perceived stock prices) when trade does not occur. The argument follows from the requirement that the equilibrium stock price drift and the resulting last trading time τ must be chosen symmetrically in both agents in order for their optimization problems to be satisfied simultaneously.

Remark 3.1 (About assumptions). Our model is simple enough to be tractable but complex enough to say something nontrivial. To strike this balance, we impose a number of assumptions, some of which are needed for Theorem 3.3, while others are merely convenient. One of our model's most restrictive assumptions is to constrain the number of agents to two. This is not an assumption of convenience, nor is it isolated to our model. Indeed, all continuous time equilibria with (proportional) transaction costs either consist of 2-agent economies, continuum-of-agent economies, or do not prove an existence result. In our model, with $c_1 \neq 0$ and $\lambda > 0$, there is no equilibrium with three or more agents when the perceived stock price's drift reacts linearly to the trading strategy. With $c_1 = 0$ and $\lambda > 0$, it is possible to include a third agent into the model, but four or more agents are not possible in equilibrium.[4]

The initial holdings of the agents are identical, which is not strictly necessary. This assumption can be relaxed to some extent, but difficulty arises because some initial holding configurations may make it impossible for the agents' optimization problems to be satisfied due to an initial jump in holdings. The problematic configurations are difficult to characterize in terms of model inputs.

We also assume that a_1 and a_2 are independent of the Brownian motion B. This assumption is used to prove admissibility of the optimal trading strategies but could be relaxed at the expense of a lengthier proof.

3.1 Effects of frictions in equilibrium

Both transaction costs and price impact affect equilibrium, and each friction has its own distinct modeling characteristics and equilibrium effects. Both frictions penalize the agents through their wealth reduction. Transaction costs do so by directly subtracting transaction fees from wealth, while price impact does so indirectly by depressing the stock's drift with each increase in the number of shares held.

From a modeling perspective, agents are held accountable for transaction costs in equilibrium through market clearing, and their wealth decreases as a result. Price impact frictions appear only as *perceived* changes in perceived off-equilibrium asset prices and wealth in the individual optimization problems. Price impact is not incorporated explicitly into the market clearing condition, but the perceived prices must align with the realized prices in equilibrium by Definition 3.1(c).

Equilibrium effects of the two frictions are similar in that more frictions lead to less trade. However, each friction has its own mechanism by which it impacts equilibrium

[4]We would like to thank Jetlir Duraj for discussions on this issue.

outcomes. First, we consider the effects of price impact, and we suppose that the level of transaction costs is fixed. Higher (and positive) levels of the price impact coefficient c_1 cause the agents to trade less, while also causing a linear decrease in the equilibrium stock price. The decrease in the perceived stock prices transfers over to the realized equilibrium stock price $\widehat{S}$ in (3.8). The case when $c_1 = 0$ corresponds to a price-taking equilibrium, in which the agents' perceived stock dynamics are not impacted by trade.

The equilibrium effects from transaction costs are perhaps more subtle. At trading times, the equilibrium trading strategies are unaffected by transaction costs. However, transaction costs determine how long agents are willing to trade by affecting τ in (3.4). Larger values of λ produce smaller values of τ, meaning that agents are not willing to continue trading if the penalty incurred by transactions costs is sufficiently large. This behavior contrasts equilibrium results with quadratic penalties, such as Bouchard et. al. [5], which exhibit a decrease in trade due to penalties rather than stopping trade. Quadratic penalties are often viewed as a (more) mathematically tractable approximation to (proportional) transaction costs. Yet the stark qualitative differences shown in Theorem 3.3 make quadratic penalties a poor substitute for transaction costs in equilibrium.

Moreover, the on-equilibrium stock price $\widehat{S}$ in (3.8) is unaffected by transaction costs. The apparent lack of an effect for $\widehat{S}$ occurs because the on-equilibrium stock price can only be uniquely determined when trade occurs. When trade does not occur, as is the case at the end of the trading period under transaction costs, $\widehat{S}$ is consistent with equilibrium in that the agents will still agree not to trade using the price $\widehat{S}$. See Dávila [12] and Weston [27] for a similar phenomenon in equilibrium models with transaction costs.

4 Optimal transaction cost

The agents pay transaction costs at a rate $\lambda > 0$ proportional to the number of shares traded. As the agents lose the wealth paid out in transaction costs, an exchange collects the fees. (See (3.3).) Higher values of λ mean that the exchange will receive more fees with every share traded. However, higher values of λ cause the agents to stop trading sooner, resulting in fewer shares traded. Given a distributional estimate on the trading targets (i.e., priors), the exchange can find a strictly positive optimal transaction cost proportion that maximizes its expected profit.

This application is intended to serve as an example for how an optimal transaction cost level can be calculated in equilibrium outside the spectrum of welfare optimization, such as in Dávila [12] or Weston [27]. Though one may expect to see several exchanges competing over their transaction cost levels for traders' business, here we consider just one exchange. A single exchange with two price-taking agents could be thought of as a national healthcare insurance market with the government acting as the exchange and only two insurance providers. Here, we assume that the exchange sets the transaction cost level ex ante in the sense that the agents reveal their targets at the beginning of the trading period.

In this section, we make the dependence on λ explicit in the notation by denoting $\theta_i = \theta_i^{(\lambda)}$ and $\tau = \tau(\lambda)$. We fix a price impact coefficient $c_1 > -\frac{1}{2}$. By Theorem 3.3, an equilibrium exists, and the total profit received by the exchange in that equilibrium is given by

$$\text{Profit}(\lambda) := \lambda\left(\left(\theta_{1,T}^{(\lambda)}\right)^{\uparrow} + \left(\theta_{1,T}^{(\lambda)}\right)^{\downarrow} + \left(\theta_{2,T}^{(\lambda)}\right)^{\uparrow} + \left(\theta_{2,T}^{(\lambda)}\right)^{\downarrow}\right).$$

By market clearing and the monotonicity of the optimal trading strategies, we see that

$$\text{Profit}(\lambda) = 2\lambda\left|\theta_{1,T}^{(\lambda)} - \frac{n}{2}\right|.$$

Since the exchange does not have advanced knowledge of the agents' trading targets when selecting the transaction cost level, it must estimate the targets by using its ex ante priors.

Proposition 4.1 asserts that a strictly positive optimal transaction cost level exists for the exchange. The proof of Proposition 4.1 is presented in Section 5.

Proposition 4.1. *Let $c_1 > -\frac{1}{2}$ be given. Suppose that $\kappa : (0,T) \to (0,\infty)$ is measurable with $\int_0^T \kappa(u)du < \infty$ and that $\gamma : [0,T] \to [0,1]$ is càdlàg, nonnegative, bounded by one, nondecreasing, and there exists $t \in [0,T)$ so that $\gamma(t) > 0$. Also, suppose that $0 < \mathbb{E}\left[A_1^2\right] < \infty$. Then, there exists $\widehat{\lambda} > 0$ so that*

$$\widehat{\lambda} \in \mathit{Argmax}\left\{\lambda > 0 : \mathbb{E}\left[\mathit{Profit}(\lambda)\right]\right\}.$$

Example 4.2. To illustrate the exchange's choice of optimal transaction cost level, we consider an example with TWAP traders. We take $T := 1$, $n := 100$, $\kappa(t) := 1$, $\gamma(t) := \gamma^{\text{TWAP}}(t) = t$, and we allow the price impact coefficient $c_1 > -\frac{1}{2}$ to be arbitrary. For a

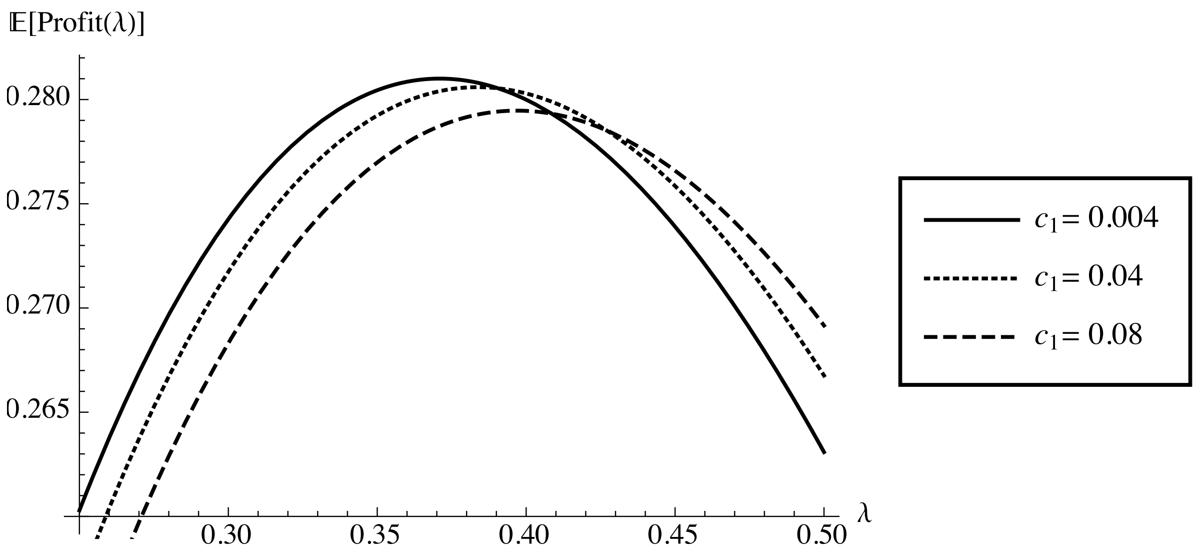

Figure 1: The exchange's expected profit in equilibrium is plotted as a function of the transaction cost $\lambda > 0$. Each plot corresponds to a different level of price impact, which is measured by varying the perceived price impact parameter $c_1 > -\frac{1}{2}$.

given transaction cost level $\lambda > 0$, we have that

$$\text{Profit}(\lambda) = 2 \max\left(0, \frac{\lambda A_1}{1+c_1} - \lambda\sqrt{\frac{2\lambda A_1}{(1+c_1)(1+2c_1)}}\right).$$

Each agent begins trading with 50 shares. The exchange estimates that agent 1 will seek to obtain a targeted number of shares that is uniformly distributed on $(50, 55)$, while it believes that agent 2 will target exactly 50 shares. The exchange's expected profit can be computed by

$$\mathbb{E}\left[\text{Profit}(\lambda)\right] = \frac{8}{5}\int_0^{\sqrt{2.5}} \frac{\lambda y^2}{1+c_1} \max\left(0, y - \sqrt{\frac{2\lambda(1+c_1)}{1+2c_1}}\right) dy.$$

Figure 1 plots the exchange's expected profit as a function of λ. The three plots vary in the degree of the agents' perceived price impact, which is measured by the parameter c_1. The case $c_1 = 0$ corresponds to a price-taking equilibrium with no price impact. When $c_1 > 0$, the agents perceive a non-zero level of price impact.

In this example, the exchange's choice of optimal transaction cost increases with increasing price impact parameter c_1. However, the exchange's optimal expected profit decreases as the perceived price impact increases.

For $\lambda > 0$ and $c_1 > -\frac{1}{2}$, equilibrium stock drift is given by

$$\text{drift}\left(\widehat{S}\right)_t = 50c_1 - A_1 t.$$

We recall in this example that the random variable A_1 is uniformly distributed on $(0, 2.5)$ and is measurable at time zero with respect to the filtration generated by the equilibrium stock price. The last trading time is given by

$$\tau = \max\left(0, 1 - \sqrt{\frac{2\lambda}{A_1} \cdot \frac{1+c_1}{1+2c_1}}\right),$$

which yields an optimal strategy for agent 1 of $\theta_{1,t} = 50 + \frac{A_1}{1+c_1}(t \wedge \tau)$. We note that the last trading time and the optimal strategy depend on both λ and c_1, while the drift of the equilibrium stock price depends on c_1 but not λ. We plot agent 1's optimal trading strategy as a function of time in Figure 2.

To get a ballpark estimate of an appropriate value for the price impact parameter c_1, we link our model to existing empirical studies in the literature. To the best of our knowledge, empirical work to estimate the price impact in a Cuoco and Cvitanić [11] setting has not been performed. However, Almgren et. al. [2] empirically estimate a nonlinear price impact model, while Chen et. al. [9] linearize [2]'s estimates to empirically estimate price impact in an Almgren and Chriss setting. We bridge the gap from our model to the estimates in [9]. Almgren and Chriss models realize price impact in the current stock price based on changes in the trading rate and holdings, whereas our Cuoco and Cvitanić setting realizes price impact in future prices via the drift dependence on holdings. To relate these two price impact models, we consider trading rates $d\theta_t = \theta'\, dt$ for a constant $\theta' \in \mathbb{R}$. Applying our perceived prices in (2.4) with c_0 and c_2 as in (3.6) to our example (with $\kappa(t) = 1$, $\gamma(t) = t$, $n = 100$, $T = 1$), we have

$$\begin{aligned}
S^{\theta}_{i,1} - S^{\theta}_{i,0} &= \int_0^1 dS^{\theta}_{i,u} \\
&= -\int_0^1 c_1 \theta_u \, du + \int_0^1 \left(100c_1 + \frac{1+2c_1}{1+c_1} A_i u - (a_i - 50)\, u\right) du + M_1 - M_0 \\
&= -\int_0^1 c_1 \theta' u \, du + \int_0^1 \left(50c_1 + \frac{1+2c_1}{1+c_1} A_i u - (a_i - 50)\, u\right) du + M_1 - M_0 \\
&= -\frac{c_1 \theta'}{2} + \int_0^1 \left(50c_1 + \frac{1+2c_1}{1+c_1} A_i u - (a_i - 50)\, u\right) du + M_1 - M_0.
\end{aligned}$$

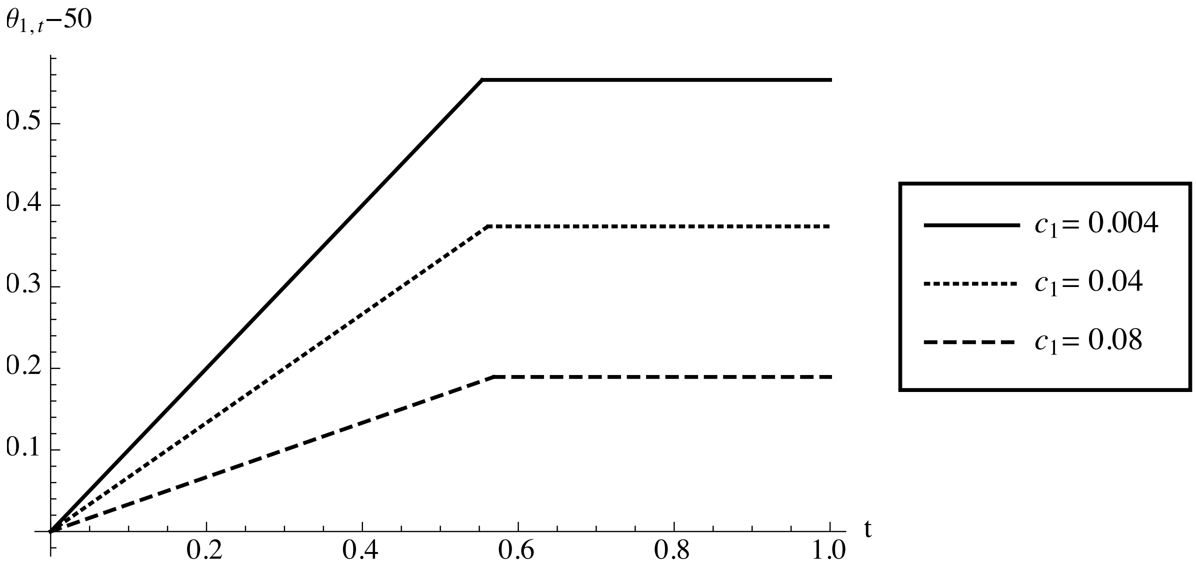

Figure 2: Agent 1's optimal trading strategy (minus the initial holdings) is plotted as a function of time. We fix $A_1 = 1$ and $\lambda = 0.1$. Each plot corresponds to a different level of price impact, which is measured by varying the perceived price impact parameter $c_1 > -\frac{1}{2}$.

We apply [9]'s estimate for the transient price impact coefficient, $0.002 = \frac{c_1}{2}$. Therefore, we estimate that $c_1 = 0.004$ is a ballpark estimate for c_1.

5 Proofs

We begin with a proof of Lemma 3.2, which provides a version of Walras' Law in our setting.

Proof of Lemma 3.2. Let $\lambda > 0$ be given, and let θ_1, θ_2, and $\widehat{S}$ be parameters in a price impact equilibrium satisfying Definition 3.1. The bank account clears for all times $t \in [0, T]$ by

$$\begin{aligned}
&\theta^{(0)}_{1,t} + \theta^{(0)}_{2,t} \\
&\quad = 2 \cdot \frac{n}{2} \widehat{S}_0 + \int_0^t (\theta_{1,u} + \theta_{2,u})\, d\widehat{S}_u - \lambda \left(\theta^{\uparrow}_{1,t} + \theta^{\downarrow}_{1,t} + \theta^{\uparrow}_{2,t} + \theta^{\downarrow}_{2,t} \right) - (\theta_{1,t} + \theta_{2,t})\, \widehat{S}_t \\
&\quad = n\widehat{S}_0 + \int_0^t n\, d\widehat{S}_u - n\widehat{S}_t - \lambda \left(\theta^{\uparrow}_{1,t} + \theta^{\downarrow}_{1,t} + \theta^{\uparrow}_{2,t} + \theta^{\downarrow}_{2,t} \right) \\
&\quad = -\lambda \left(\theta^{\uparrow}_{1,t} + \theta^{\downarrow}_{1,t} + \theta^{\uparrow}_{2,t} + \theta^{\downarrow}_{2,t} \right).
\end{aligned}$$

□

Next, we motivate the proof of our main result, Theorem 3.3. To solve for an equilibrium, each agent must solve a convex optimization problem subject to market consistency requirements (namely, clearing conditions and agreement of the perceived prices at trade times). For $i = 1, 2$, the first-order condition for agent i's optimization problem requires her to seek $\theta \in \mathcal{A}$ so that

$$Y_{i,t}^{\theta} := \mathbb{E}\left[\int_t^T \kappa(u)\left(c_0(u, a_\Sigma) + \frac{n}{2} + \gamma(u)(1+c_2)\left(a_i - \frac{n}{2}\right) - (1+2c_1)\theta_u\right) du \,|\, \mathcal{F}_t\right] \quad (5.1)$$

satisfies $-\lambda \le Y_{i,t}^{\theta} \le \lambda$ for all $t \in [0, T]$ and

$$\int_0^T \left(\lambda - Y_{i,t}^{\theta}\right) d\theta_t^{\uparrow} = \int_0^T \left(\lambda + Y_{i,t}^{\theta}\right) d\theta_t^{\downarrow} = 0. \quad (5.2)$$

Using the clearing and consistency definitions from equilibrium in Definition 3.1 and the form of the first-order condition in (5.1), we deduce the forms of c_0 and c_2 in (3.6). Those calculations yield

$$Y_{i,t}^{\theta} = \mathbb{E}\left[\int_t^T \kappa(u)(1+2c_1)\left(\frac{n}{2} + \frac{\gamma(u)A_i}{1+c_1} - \theta_u\right) du \,|\, \mathcal{F}_t\right]. \quad (5.3)$$

By observing that $Y_{i,T}^{\theta} = 0$ and selecting a right-continuous version of the conditional expectation in (5.1), the reflection condition (5.2) shows that the optimal strategy will not allow for trade at T. Since the agents' perceived prices are only required to agree with the equilibrium stock price at trade times, the perceived prices may differ from the prescribed dividend payment at T.

Coupling the terminal condition $Y_{i,T}^{\theta} = 0$ with the fact that

$$t \mapsto \frac{n}{2} + \frac{\gamma(t)|A_i|}{1+c_1}$$

is increasing and $\mathcal{F}_0$-measurable for all $t \in [0, T]$ tells us that a solution $\theta_i \in \mathcal{A}$ to the first-order condition is to trade early in the period with

$$\theta_{i,\cdot} = \frac{n}{2} + \frac{\gamma(\cdot)A_i}{1+c_1}$$

then stop trade at time $\tau < T$ when the remaining integral $\int_\tau^T \kappa(1+2c_1)\left(\frac{n}{2} + \frac{\gamma A_i}{1+c_1} - \theta_{i,\tau}\right)$ becomes sufficiently small so that $-\lambda \leq Y^{\theta_i}_{i,\cdot} \leq \lambda$ will hold after τ. These observations motivate the definition of τ in (3.4).

The first-order condition's reflection condition (5.2) and bounds $-\lambda \leq Y^{\theta}_{i,\cdot} \leq \lambda$ ensure that the agents will trade less (by stopping or pausing trade) with increases in λ. In our model, the agents will stop trade strictly before T and will not resume again. However, any model where agents face a linear-quadratic control problem with proportional transaction costs will also see agents taking breaks from trade when trading optimally. In particular, any potential generalization of this model or the ergodic control problem of Gonon et. al. [14] will exhibit this behavior.

Proof of Theorem 3.3. We break the proof up into steps.

Step 1: Admissibility. We show for $i = 1, 2$, that θ_i defined in (3.7) is admissible; that is, $\theta_i \in \mathcal{A}$. The function γ is càdlàg and nondecreasing, while $\frac{A_i}{1+c_1}$ and τ are $\mathcal{F}_0$-measurable. To verify integrability, we use the square-integrability of A_i and σ as well as their independence to ensure that for some constant $C > 0$, we have

$$\begin{aligned}
\mathbb{E}\int_0^T (\sigma_u \theta_{i,u})^2\, du &\leq C\,\mathbb{E}\int_0^T \sigma_u^2\left(1 + A_i^2\right) du \\
&= C\,\mathbb{E}\left[1 + A_i^2\right]\,\mathbb{E}\int_0^T \sigma_u^2 du \\
&< \infty.
\end{aligned}$$

Thus, $\theta_i \in \mathcal{A}$.

Step 2: Optimality. Next, we check that $\theta_i \in \mathcal{A}$ is optimal in (3.1). For an arbitrary $\theta \in \mathcal{A}$, we define $V_i(\theta)$ by

$$V_i(\theta) := \mathbb{E}\left[X^{\theta}_{i,T} - L^{\theta}_{i,T} \,|\, \mathcal{F}_0\right], \tag{5.4}$$

where $L^{\theta}_{i,T}$ is defined in (2.3), $X^{\theta}_{i,T}$ in (2.5), and the perceived off-equilibrium price dynamics in (2.4) have initial stock price and martingale parts given by the proposed equilibrium stock price $\widehat{S}$ in (3.8).

By the definition of $\widehat{S}$ in (3.8), the martingale part of $\widehat{S}$, S^{θ}_1, and S^{θ}_2 is given by the martingale $\int_0^\cdot \sigma dB$. Since $\theta \in \mathcal{A}$ is adapted and $\mathbb{E}\int_0^T (\sigma_u\theta_u)^2\, du < \infty$, we have that $\left\{\int_0^t \sigma_u \theta_u dB_u\right\}_{t\in[0,T]}$ is a square-integrable martingale with respect to $\mathbb{F}$. We use this property in the calculation of $V_i(\theta)$ below.

For c_0 and c_2 given in (3.6) and the initial stock price given by $\widehat{S}_0$ in (3.8), we have

$$\begin{aligned}
V_i(\theta) &= \mathbb{E}\left[X^\theta_{i,T} - L^\theta_{i,T} \,|\, \mathcal{F}_0\right] \\
&= \frac{n}{2}\widehat{S}_0 - \lambda \mathbb{E}\left[\theta^\uparrow_T + \theta^\downarrow_T \,|\, \mathcal{F}_0\right] \\
&\quad + \mathbb{E}\left[\int_0^T \kappa(t)\left(\left(c_0(t) + \gamma(t)c_2\left(a_i - \frac{n}{2}\right)\right)\theta_t - c_1\theta_t^2 - \frac{1}{2}\left(\theta_t - \frac{n}{2} - \gamma(t)\left(a_i - \frac{n}{2}\right)\right)^2\right) dt \,|\, \mathcal{F}_0\right] \\
&= \frac{n}{2}\widehat{S}_0 - \lambda \mathbb{E}\left[\theta^\uparrow_T + \theta^\downarrow_T \,|\, \mathcal{F}_0\right] - \frac{1}{2}\mathbb{E}\left[\int_0^T \kappa(t)\left(\frac{n}{2} + \gamma(t)\left(a_i - \frac{n}{2}\right)\right)^2 dt \,|\, \mathcal{F}_0\right] \\
&\quad + \mathbb{E}\left[\int_0^T \kappa(t)\left(\left(\left(c_1 + \frac{1}{2}\right)n + \gamma(t)\frac{(1+2c_1)A_i}{1+c_1}\right)\theta_t - \left(c_1 + \frac{1}{2}\right)\theta_t^2\right) dt \,|\, \mathcal{F}_0\right].
\end{aligned}$$

The above calculation reveals that when optimizing over θ in (5.4), the second-order condition requires that $c_1 > -\frac{1}{2}$.

For notational convenience, we define

$$\alpha(t) := \left(c_1 + \frac{1}{2}\right)n + \gamma(t)\frac{(1+2c_1)A_i}{1+c_1} \quad \text{and} \quad \beta := c_1 + \frac{1}{2},$$

so that

$$\begin{aligned}
V_i(\theta) = \frac{n}{2}\widehat{S}_0 &- \lambda \mathbb{E}\left[\theta^\uparrow_T + \theta^\downarrow_T \,|\, \mathcal{F}_0\right] \\
&- \frac{1}{2}\int_0^T \kappa(t)\left(\frac{n}{2} + \gamma(t)\left(a_i - \frac{n}{2}\right)\right)^2 dt + \mathbb{E}\left[\int_0^T \kappa(t)\left(\alpha(t)\theta_t - \beta\theta_t^2\right) dt \,|\, \mathcal{F}_0\right].
\end{aligned}$$

In order to help us with computations, we define $\tilde{\gamma}$ for $t \in [0,T]$ by

$$\tilde{\gamma}_t := \begin{cases} \gamma(t), & t < \tau, \\ \frac{\int_\tau^T \kappa(u)\gamma(u)du - \frac{\lambda(1+c_1)}{|A_1(1+2c_1)|}}{\int_\tau^T \kappa(u)du}, & t \geq \tau \text{ and } \chi > \lambda, \\ 0, & t \geq \tau \text{ and } \chi \leq \lambda. \end{cases} \tag{5.5}$$

The definition of $\tilde{\gamma}$ in comparison with θ_i in (3.7) shows that $\theta_{i,t} = \frac{n}{2} + \frac{A_i}{1+c_1}\tilde{\gamma}_t$. Since $A_1 = 0$ implies that $\chi = 0$ and since $\tau < T$, we have that $\tilde{\gamma}$ is well-defined in the $t \geq \tau$ and $\chi > \lambda$ case.

We now proceed to verify the first-order condition outlined above for our proposed optimal strategy θ_i. By algebraic manipulation and applying (5.7),

$$
\begin{aligned}
&V_i(\theta) - V_i(\theta_i) \\
&\quad = \mathbb{E}\left[\lambda\left(\theta^{\uparrow}_{i,T} + \theta^{\downarrow}_{i,T} - \theta^{\uparrow}_{T} - \theta^{\downarrow}_{T}\right) + \int_0^T \kappa(u)\left(\alpha(u)(\theta_u - \theta_{i,u}) - \beta(\theta_u - \theta_{i,u})(\theta_u + \theta_{i,u})\right) du \,|\, \mathcal{F}_0\right] \\
&\quad = \mathbb{E}\left[\lambda\left(\theta^{\uparrow}_{i,T} + \theta^{\downarrow}_{i,T} - \theta^{\uparrow}_{T} - \theta^{\downarrow}_{T}\right) + \int_0^T \kappa(u)\left((\alpha(u) - 2\beta\theta_{i,u})(\theta_u - \theta_{i,u}) - \beta(\theta_u - \theta_{i,u})^2\right) du \,|\, \mathcal{F}_0\right] \\
&\quad \leq \mathbb{E}\left[\lambda\left(\theta^{\uparrow}_{i,T} + \theta^{\downarrow}_{i,T} - \theta^{\uparrow}_{T} - \theta^{\downarrow}_{T}\right) + \int_0^T \kappa(u)(\alpha(u) - 2\beta\theta_{i,u})(\theta_u - \theta_{i,u}) du \,|\, \mathcal{F}_0\right] \\
&\quad = \mathbb{E}\left[\lambda\left(\theta^{\uparrow}_{i,T} + \theta^{\downarrow}_{i,T} - \theta^{\uparrow}_{T} - \theta^{\downarrow}_{T}\right) + \frac{A_i(1+2c_1)}{1+c_1}\int_0^T \kappa(u)\left(\gamma(u) - \tilde{\gamma}_u\right)(\theta_u - \theta_{i,u}) du \,|\, \mathcal{F}_0\right].
\end{aligned}
$$

We define $Y_i = \{Y_{i,t}\}_{t\in[0,T]}$ by

$$
Y_{i,t} = A_i \frac{1+2c_1}{1+c_1}\int_t^T \kappa(u)\left(\gamma(u) - \tilde{\gamma}_u\right) du.
$$

By the definition of τ in (3.4) and χ in (3.5), we have that $Y_{i,t} = \lambda\,\mathrm{sign}A_i$ for $t \leq \tau$ and $\chi \geq \lambda$, while $Y_{i,t} \in \mathrm{sign}A_i[0,\lambda]$ for all $t \in [0,T]$ and all values of χ. Since θ_i is constant after τ, we see that

$$
\int_{0-}^T Y_{i,u} d\theta_{i,u} = \lambda\left(\theta^{\uparrow}_{i,T} + \theta^{\downarrow}_{i,T}\right). \tag{5.6}
$$

For any $\theta \in \mathcal{A}$,

$$
\begin{aligned}
&\frac{A_i(1+2c_1)}{1+c_1}\int_0^T \kappa(u)\left(\gamma(u) - \tilde{\gamma}_u\right)(\theta_u - \theta_{i,u})\, du - \lambda\left(\theta^{\uparrow}_{T} + \theta^{\downarrow}_{T}\right) \\
&\qquad = \int_{0-}^T Y_{i,u} d\left(\theta - \theta_i\right)_u - Y_{i,T}\left(\theta_T - \theta_{i,T}\right) + Y_{i,0}\left(\frac{n}{2} - \frac{n}{2}\right) - \lambda\left(\theta^{\uparrow}_{T} + \theta^{\downarrow}_{T}\right) \\
&\qquad\qquad \text{by integration by parts} \\
&\qquad = \int_{0-}^T Y_{i,u} d\left(\theta - \theta_i\right)_u - \lambda\left(\theta^{\uparrow}_{T} + \theta^{\downarrow}_{T}\right) \qquad \text{since } Y_{i,T} = 0 \\
&\qquad = \int_{0-}^T Y_{i,u} d\theta_u - \lambda\left(\theta^{\uparrow}_{i,T} + \theta^{\downarrow}_{i,T}\right) - \lambda\left(\theta^{\uparrow}_{T} + \theta^{\downarrow}_{T}\right) \quad \text{by (5.6)} \\
&\qquad = \int_{0-}^T \underbrace{\left(Y_{i,u} - \lambda\right)}_{\leq 0} d\theta^{\uparrow}_u - \int_{0-}^T \underbrace{\left(\lambda + Y_{i,u}\right)}_{\geq 0} d\theta^{\downarrow}_u - \lambda\left(\theta^{\uparrow}_{i,T} + \theta^{\downarrow}_{i,T}\right) \\
&\qquad \leq -\lambda\left(\theta^{\uparrow}_{i,T} + \theta^{\downarrow}_{i,T}\right).
\end{aligned} \tag{5.7}
$$

Finally, we see that

$$\begin{aligned} &V_i(\theta) - V_i(\theta_i) \\ &\quad \leq \mathbb{E}\left[\lambda\left(\theta^{\uparrow}_{i,T} + \theta^{\downarrow}_{i,T}\right) - \lambda\left(\theta^{\uparrow}_{i,T} + \theta^{\downarrow}_{i,T}\right) \mid \mathcal{F}_0\right] \quad \text{by (5.7)} \\ &\quad \leq 0. \end{aligned}$$

Thus, θ_i is optimal for agent i.

Step 3: Markets clear. The stock market clears since for all $t \in [0,T]$, we have $\theta_{1,t} + \theta_{2,t} = n + \frac{\tilde{\gamma}_t}{1+c_1}(A_1 + A_2) = n$.

Step 4: Equilibrium prices are consistent. We verify that for $t \in \{u \in [0,T] : d\theta^{\uparrow}_{1,u} + d\theta^{\downarrow}_{1,u} > 0\}$, we have that

$$S^{\theta_1}_{1,t} = S^{\theta_2}_{2,t} = \widehat{S}_t. \tag{5.8}$$

Since $t \in \{u \in [0,T] : d\theta^{\uparrow}_{1,u} + d\theta^{\downarrow}_{1,u} > 0\} \subseteq [0,\tau]$, it is sufficient to verify that (5.8) holds for $t \in [0,\tau]$.

We use the optimal trading strategy formula (3.7), the equilibrium price impact coefficient formulas (3.6), and the conjectured price impact drift formula (2.4) to calculate the equilibrium drift. For $i \in \{1,2\}$ and $t \in (0,\tau)$, the drift of $S^{\theta_i}_{i,t}$ is given by

$$\begin{aligned} \text{drift}\left(S^{\theta_i}_{i,\cdot}\right)_t &= \kappa(t)\left(c_0(t,a_\Sigma) - c_1\theta_{i,t} + \gamma(t)c_2\left(a_i - \frac{n}{2}\right)\right) \\ &= \kappa(t)\left(\frac{c_1 n}{2} - \frac{1}{2}\gamma(t)\left(a_\Sigma - n\right)\right), \end{aligned}$$

which agrees with the drift of $\widehat{S}$ in (3.8). As per the conjectured price impact drift in (2.4), the martingale components and initial values of $\widehat{S}$, S^{θ}_1, and S^{θ}_2 do not depend on the choice of $i = 1,2$ or $\theta \in \mathcal{A}$. Thus, we conclude that (5.8) holds for all $t \in [0,\tau]$, as desired.

□

Next, we prove Proposition 4.1, which establishes the existence of an optimal transaction cost, $\widehat{\lambda} > 0$.

Proof of Proposition 4.1. Theorem 3.3 establishes the existence of an equilibrium with optimal trading strategies $\theta^{(\lambda)}_i \in \mathcal{A}$ given in (3.7) for every $\lambda > 0$.

As in the proof of Theorem 3.3, we introduce the process $\tilde{\gamma}$ as a function of $t \in [0,T]$

and $\lambda > 0$ by

$$\tilde{\gamma}_t(\lambda) := \begin{cases} \gamma(t), & t < \tau(\lambda), \\ \frac{\int_{\tau(\lambda)}^T \kappa(u)\gamma(u)du - \frac{\lambda(1+c_1)}{|A_1(1+2c_1)|}}{\int_{\tau(\lambda)}^T \kappa(u)du}, & t \geq \tau(\lambda) \text{ and } \chi > \lambda, \\ 0, & t \geq \tau(\lambda) \text{ and } \chi \leq \lambda. \end{cases}$$

Here, χ is defined in (3.5), and we make $\tilde{\gamma}$'s dependence on λ explicit. For any $\lambda > 0$, we use the definition of $\tilde{\gamma}$ above to deduce that

$$\theta_{1,T}^{(\lambda)} - \frac{n}{2} = \theta_{1,\tau(\lambda)}^{(\lambda)} - \frac{n}{2} = \frac{A_1}{1+c_1}\tilde{\gamma}_{\tau(\lambda)}(\lambda).$$

We first seek to show that $\lambda \mapsto \tilde{\gamma}_{\tau(\lambda)}(\lambda)$ is continuous, from which we will conclude that $\lambda \mapsto \text{Profit}(\lambda)$ is continuous. On $\{A_1 = 0\}$, $\chi = 0$ so that $\chi < \lambda$ for all $\lambda > 0$, and thus, $\lambda \mapsto \tilde{\gamma}_{\tau(\lambda)}(\lambda)$ is continuous in this case.

We let $\underline{\lambda} > 0$ be given. On $\{A_1 \neq 0\}$, we consider $0 < \underline{\lambda} \leq \lambda_1 < \lambda_2$. Then, $\tau(\lambda_2) \leq \tau(\lambda_1) \leq \tau(\underline{\lambda}) < T$. For notational simplicity, we define $C := \frac{1+c_1}{|A_1|(1+2c_1)}$ and $\tau_1 := \tau(\lambda_1)$, $\tau_2 := \tau(\lambda_2)$. We have that

$$\begin{aligned} \lambda_1 C &= \int_{\tau_1}^T \kappa(u)\Big(\gamma(u) - \tilde{\gamma}_{\tau_1}(\lambda_1)\Big)du \\ &= \int_{\tau_1}^{\tau_2} \kappa(u)\Big(\gamma(u) - \tilde{\gamma}_{\tau_1}(\lambda_1)\Big)du \\ &\qquad + \underbrace{\int_{\tau_2}^T \kappa(u)\Big(\gamma(u) - \tilde{\gamma}_{\tau_2}(\lambda_2)\Big)du}_{=\lambda_2 C} + \int_{\tau_2}^T \kappa(u)\Big(\tilde{\gamma}_{\tau_2}(\lambda_2) - \tilde{\gamma}_{\tau_1}(\lambda_1)\Big)du \\ &= \int_{\tau_2}^{\tau_1} \kappa(u)\Big(\tilde{\gamma}_{\tau_1}(\lambda_1) - \gamma(u)\Big)du + \lambda_2 C + \Big(\tilde{\gamma}_{\tau_2}(\lambda_2) - \tilde{\gamma}_{\tau_1}(\lambda_1)\Big)\int_{\tau_2}^T \kappa(u)du. \end{aligned}$$

Using that γ is nondecreasing, we see that $\tilde{\gamma}_{\tau_2}(\lambda_2) \leq \gamma(u)$ for all $u \in (\tau_2, \tau_1)$. By rearranging terms, we arrive at

$$\begin{aligned} 0 &\leq \Big(\tilde{\gamma}_{\tau_1}(\lambda_1) - \tilde{\gamma}_{\tau_2}(\lambda_2)\Big)\int_{\tau_2}^T \kappa(u)du \\ &= C(\lambda_2 - \lambda_1) + \int_{\tau_2}^{\tau_1} \kappa(u)\Big(\tilde{\gamma}_{\tau_1}(\lambda_1) - \gamma(u)\Big)du \\ &\leq C(\lambda_2 - \lambda_1) + \Big(\tilde{\gamma}_{\tau_1}(\lambda_1) - \tilde{\gamma}_{\tau_2}(\lambda_2)\Big)\int_{\tau_2}^{\tau_1} \kappa(u)du. \end{aligned}$$

Therefore, on $\{A_1 \neq 0\}$, we have that

$$0 \leq \tilde{\gamma}_{\tau_1}(\lambda_1) - \tilde{\gamma}_{\tau_2}(\lambda_2) \leq \frac{C}{\int_{\tau_1 \vee \tau_2}^T \kappa(u)du}(\lambda_2 - \lambda_1) \leq \frac{C}{\int_{\tau(\underline{\lambda})}^T \kappa(u)du}(\lambda_2 - \lambda_1), \tag{5.9}$$

from which we conclude that $\lambda \mapsto \tilde{\gamma}_{\tau(\lambda)}(\lambda)$ is continuous for $\lambda > \underline{\lambda}$. Since $\underline{\lambda} > 0$ is arbitrary, we have that $\lambda \mapsto \tilde{\gamma}_{\tau(\lambda)}(\lambda)$ is continuous for $\lambda > 0$. Thus, $\lambda \mapsto \theta_{1,T}^{(\lambda)} - \frac{n}{2}$ and $\lambda \mapsto \text{Profit}(\lambda)$ are continuous for $\lambda > 0$. Since $\mathbb{E}[a_i^2] < \infty$, we also have that $\lambda \mapsto \mathbb{E}[\text{Profit}(\lambda)]$ is continuous for $\lambda > 0$.

Finally, we show that $\lambda \mapsto \mathbb{E}[\text{Profit}(\lambda)]$ achieves a maximum for $\lambda > 0$. The assumption that $\mathbb{E}[A_1^2] > 0$ and that there exists $t \in [0, T)$ for which $\gamma(t) > 0$ ensures that there exists $\lambda > 0$ such that $\mathbb{E}[\text{Profit}(\lambda)] > 0$. Since $\tilde{\gamma}$ is bounded by one, we have that

$$0 \leq \lim_{\lambda \downarrow 0} \mathbb{E}\left[\text{Profit}(\lambda)\right] \leq \lim_{\lambda \downarrow 0} 2\lambda\, \mathbb{E}\left[\frac{|A_1|}{1 + c_1}\right] = 0.$$

We define the random variable $\lambda^{\max} := 2|A_1| \int_0^T \kappa(u)du$. For $\lambda \geq \lambda^{\max}$, we have that $\chi < \lambda$, where χ is defined in (3.5). In this case, trade does not occur, and $\left|\theta_{1,\tau(\lambda)}^{(\lambda)} - \frac{n}{2}\right| = 0$. For any $\lambda > 0$, we bound $\text{Profit}(\lambda)$ by

$$\begin{aligned}
\text{Profit}(\lambda) &\leq \frac{2\lambda|A_1|}{1 + c_1}\gamma(T)\mathbb{I}_{\{\lambda \leq \lambda^{\max}\}} \\
&\leq \frac{2\lambda^{\max}|A_1|}{1 + c_1} \\
&\leq A_1^2 \cdot \frac{4\int_0^T \kappa(u)du}{1 + c_1}.
\end{aligned}$$

Since $\mathbb{E}[A_1^2] < \infty$, we apply the dominated convergence theorem to obtain that

$$\lim_{\lambda \to \infty} \mathbb{E}[\text{Profit}(\lambda)] = 0.$$

Thus, $\lambda \mapsto \mathbb{E}[\text{Profit}(\lambda)]$ achieves a maximum for $\lambda > 0$.

□

References

[1] Robert Almgren and Neil Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40, 2001.

[2] Robert Almgren, Chee Thum, Emmanuel Hauptmann, and Hong Li. Direct estimation of equity market impact. *Risk*, 18:1–18, 2005.

[3] Adem Atmaz and Suleyman Basak. Stock market and no-dividend stocks. Working paper, January 2021.

[4] Peter Bank, H. Mete Soner, and Moritz Voß. Hedging with temporary price impact. *Mathematics and Financial Economics*, 11(2):215–239, 2017.

[5] Bruno Bouchard, Masaaki Fukasawa, Martin Herdegen, and Johannes Muhle-Karbe. Equilibrium returns with transaction costs. *Finance and Stochastics*, 22(3):569–601, 2018.

[6] Jacob Boudoukh, Roni Michaely, Matthew Richardson, and Michael R. Roberts. On the importance of measuring payout yield: Implications for empirical asset pricing. *Journal of Finance*, 62:877–915, 2007.

[7] Markus K. Brunnermeier and Lasse Heje Pedersen. Predatory trading. *Journal of Finance*, 60(4):1825–1863, August 2005.

[8] Adrian Buss, Raman Uppal, and Grigory Vilkov. Asset prices in general equilibrium with recursive utility and illiquidity induced by transactions costs. Working paper, January 2014.

[9] Xiao Chen, Jin Hyuk Choi, Kasper Larsen, and Duane J. Seppi. Resolving asset pricing puzzles using price-impact. Working paper, June 2020.

[10] Jin Hyuk Choi, Kasper Larsen, and Duane Seppi. Equilibrium effects of intraday order-splitting benchmarks. *Mathematics and Financial Economics*, 15(2):315–352, 2021.

[11] Domenico Cuoco and Jakša Cvitanić. Optimal consumption choices for a "large" investor. *Journal of Economic Dynamics and Control*, 22:401–436, 1998.

[12] Eduardo Dávila. Optimal financial transaction taxes. Working paper, September 2020.

[13] Nicolae Gârleanu and Lasse Heje Pedersen. Dynamic portfolio choice with frictions. *Journal of Economic Theory*, 165:487–516, 2016.

[14] Lukas Gonon, Johannes Muhle-Karbe, and Xiaofei Shi. Asset pricing with general transaction costs: Theory and numerics. *Mathematical Finance*, 31:595–648, April 2021.

[15] Samuel M. Hartzmark and David H. Solomon. The dividend month premium. *Journal of Financial Economics*, 109(3):640–660, 2013.

[16] Martin Herdegen and Johannes Muhle-Karbe. Stability of Radner equilibria with respect to small frictions. *Finance and Stochastics*, 22(2):443–502, April 2018.

[17] Ming Huang. Liquidity shocks and equilibrium liquidity premia. *Journal of Economic Theory*, 109(1):104–129, March 2003.

[18] Ioannis Karatzas and Steven Shreve. *Methods of Mathematical Finance*. Applications of Mathematics Series. Springer, 1998.

[19] Albert S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1336, November 1985.

[20] Andrew W. Lo, Harry Mamaysky, and Jiang Wang. Asset prices and trading volume under fixed transactions costs. *Journal of Political Economy*, 112(5), 2004.

[21] Yuliy Sannikov and Andrzej Skrzypacz. Dynamic trading: Price inertia and front-running. Working paper, December 2016.

[22] James Tobin. A proposal for international monetary reform. *Eastern Economic Journal*, 4(3-4):153–159, 1978.

[23] Dimitri Vayanos. Transaction costs and asset prices: A dynamic equilibrium model. *Review of Financial Studies*, 11(1):1–58, January 1998.

[24] Dimitri Vayanos. Strategic trading and welfare in a dynamic market. *Review of Economic Studies*, 66:219–254, 1999.

[25] Dimitri Vayanos and Jean-Luc Vila. Equilibrium interest rate and liquidity premium with transaction costs. *Economic Theory*, 13(3), 1999.

[26] Moritz Voß. A two-player price impact game. Working paper, November 2019.

[27] Kim Weston. Existence of a Radner equilibrium in a model with transaction costs. *Mathematics and Financial Economics*, 12(4):517–539, September 2018.