Strategic Execution in the Presence of an Uninformed Arbitrageur

Ciamac C. Moallemi
Graduate School of Business
Columbia University
email: ciamac@gsb.columbia.edu

Beomsoo Park
Electrical Engineering
Stanford University
email: beomsoo@stanford.edu

Benjamin Van Roy
Management Science & Engineering
Electrical Engineering
Stanford University
email: bvr@stanford.edu

Initial Version: January 1, 2008
Current Revision: December 21, 2010

Abstract

We consider a trader who aims to liquidate a large position in the presence of an arbitrageur who hopes to profit from the trader's activity. The arbitrageur is uncertain about the trader's position and learns from observed price fluctuations. This is a dynamic game with asymmetric information. We present an algorithm for computing perfect Bayesian equilibrium behavior and conduct numerical experiments. Our results demonstrate that the trader's strategy differs significantly from one that would be optimal in the absence of the arbitrageur. In particular, the trader must balance the conflicting desires of minimizing price impact and minimizing information that is signaled through trading. Our results demonstrate that accounting for information signaling and the presence of strategic adversaries can greatly reduce execution costs.

1. Introduction

When buying or selling securities, value is lost through execution costs such as exchange fees, commissions, bid-ask spreads, and price impact. The latter can be dramatic and typically dominates other sources of execution cost when trading large blocks, when the security is thinly traded, or when there is an urgent demand for liquidity. Execution algorithms aim to reduce price impact by partitioning the quantity to be traded and placing trades sequentially. Growing recognition for the importance of execution has fueled an academic literature on the topic as well as the formation of specialized groups at investment banks and other organizations to offer execution services.

This is a preprint version of the article. The final version may be found at < https://doi.org/10.1016/j.finmar.2011.11.002 >.

Optimal execution algorithms have been developed for a number of models. In the base model of Bertsimas and Lo (1998), a stock price nominally follows a discrete-time random walk and the market impact of a trade is permanent and linear in trade size. The authors establish that expected cost is minimized by an equipartitioning policy. This policy trades equal amounts over time increments within the trading horizon. Further developments have led to optimal execution algorithms for models that incorporate price predictions (Bertsimas and Lo, 1998), bid-ask spreads and resilience (Obizhaeva and Wang, 2005; Alfonsi et al., 2007a), nonlinear price impact models (Almgren, 2003; Alfonsi et al., 2007b), and risk aversion (Subramanian and Jarrow, 2001; Almgren and Chriss, 2000; Dubil, 2002; Huberman and Stanzl, 2005; Engle and Ferstenberg, 2006; Hora, 2006; Almgren and Lorenz, 2006; Schied and Schönenborn, 2007; Lorenz, 2008).

The aforementioned results offer insight into how one should partition a block and sequence trades under various assumptions about market dynamics and objectives. The resulting algorithms, however, are unrealistic in that they exhibit predictable behavior. Such predictable behavior allows strategic adversaries, which we call arbitrageurs, to "front-run" trades and profit at the expense of increased execution cost. For example, consider liquidating a large block by an equipartitioning policy which sells an equal amount during each minute of a trading day. Trades early in the day generate abnormal price movements. The resulting "information leakage" allows an observing arbitrageur to anticipate further liquidation. If the arbitrageur sells short and closes his position at the end of the day, he profits from expected price decrease. The arbitrageur's actions amplify price impact and therefore increase execution costs. Concern about the increased cost of trading due to information leakage is not academic. Indeed, it is known that many high-frequency statistical arbitrage trading strategies developed by banks and hedge funds profit by exploiting precisely this type of signalling (Duhigg, 2009).

Several recent papers study game-theoretic models of execution in the presence of strategic arbitrageurs (Brunnermeier and Pedersen, 2005; DeMarzo and Urošević, 2006; Carlin et al., 2007; Schönenborn and Schied, 2007; Oehmke, 2010). However, these models involve games with symmetric information, in which arbitrageurs know the position to be liquidated. In more realistic scenarios, this information would be the private knowledge of the trader, and the arbitrageurs would make inferences as to the trader's position based on observed market activity.

This type of information asymmetry is central to effective execution. The fact that his position is unknown to others allows the trader to greatly reduce execution costs. But to do so requires the deliberate management of information leakage, or the signals that are transmitted via trading activity. Further, the desire to minimize information signaling may be at odds with the desire to minimize price impact. A model through which such signaling can be studied must account for uncertainty among arbitrageurs and their ability to learn from observed price fluctuations. In this paper we formulate and study a simple model which we believe to be the first that meets this requirement.

The contributions of this paper are as follows:

1. We formulate the optimal execution problem as a dynamic game with asymmetric information. This game involves a trader and a single arbitrageur. Both agents are risk neutral, and market dynamics evolve according to a linear permanent price impact model. The trader seeks to liquidate his position in a finite time horizon. The arbitrageur attempts to infer the position of the trader by observing market price movements, and seeks to exploit this information for profit.

2. We develop an algorithm that computes perfect Bayesian equilibrium behavior.

3. We demonstrate that the associated equilibrium strategies take on a simple structure: Trades placed by the trader are linear in the trader's position, the arbitrageur's position and the arbitrageur's expectation of the trader's position. Trades placed by the arbitrageur are linear in the arbitrageur's position and his expectation of the trader's position. Equilibrium policies depend on the time horizon and a parameter that we call the "relative volume". This parameter captures the magnitude of the per-period activity of the trader relative to the exogenous fluctuations of the market.

4. We present computational results that make several points about perfect Bayesian equilibrium in our model:

    (a) In the presence of adversaries, there are significant potential benefits to employing perfect Bayesian equilibrium strategies.

(b) Unlike strategies proposed based on prior models in the literature, which exhibit deterministic sequences of trades, trades in a perfect Bayesian equilibrium adaptively respond to price fluctuations; the trader leverages these random outcomes to conceal his activity.

(c) When the relative volume of the trader's activity is low, in equilibrium, the trader can ignore the presence of the arbitrageur and will equipartition to minimize price impact. Alternatively, when the relative volume is high, the trader will concentrate his trading activity in a short time interval so as to minimize signaling.

(d) The presence of the arbitrageur leads to a spill-over effect. That is, the trader's expected loss due to the arbitrageur's presence is larger than the expected profit of the arbitrageur. Hence, other market participants benefit from the arbitrageur's activity.

5. We discuss how the basic model presented can be can be extended to incorporate a number of additional features, such as transient price impact and risk aversion.

Our primary motivation is to carry out the first study of how the presence of arbitrageurs should influence execution rather than to capture any specific kind of trading activity. That said, our model may reflect aspects of block trade execution in the absence of asymmetric information regarding stock value. Though many such block trades are executed through "upstairs markets" in which reputable traders can signal that they are liquidity motivated to reduce price impact, as suggested by findings of Madhavan and Cheng (1997), smaller block trades or trades by less credible parties are more efficiently broken up and executed through the "downstairs market." Our model relates to the latter category of block trades. One might argue that our base model, which assumes permanent price impact, does not suit the study of liquidity motivated execution. Our view is that price impact in such a context should indeed be temporary, and that the permanent price impact model serves as an abstraction that approximates situations where the time constant of price impact significantly exceeds the execution time horizon. Further, we will discuss later in the paper how our approach generalizes to more complex models involving temporary price impact.

Solving for perfect Bayesian equilibrium in dynamic games with asymmetric information is notoriously difficult. What facilitates effective computation in our model is that, in equilibrium, each agent solves a tractable linear-quadratic Gaussian control problem. Similar approaches based on linear-quadratic Gaussian control have previously been used to analyze equilibrium behavior of traders with private information. Here, the private signal typically takes the form of information on the fundamental value of the traded asset. This line of work on "insider trading" or "strategic trading" begins with the seminal paper of Kyle (1985), and includes many subsequent papers (e.g., Back, 1992; Holden and Subrahmanyam, 1992; Foster and Viswanathan, 1994; Holden and Subrahmanyam, 1994; Foster and Viswanathan, 1996; Vayanos, 1999, 2001; Back and Baruch, 2004; Guo and Kyle, 2005; Cao et al., 2006). Among these contributions, Foster and Viswanathan (1994) come closest to the model and method we propose. In the model of that paper, there are two strategic traders, many "noise" traders, and a market maker. The strategic traders possess information that is not initially reflected in market prices. One trader knows more than the other. The more informed trader adapts trades to maximize his expected payoff, and this entails controlling how his private information is revealed through price fluctuations. This model parallels ours if we think of the arbitrageur as the less informed trader. However, in our model there is no private information about future dividends but instead uncertainty about the size of the position to be liquidated. Further, in the model of Foster and Viswanathan (1994), trades influence prices because the market maker tries to infer the traders' private information whereas. In our setting, there is an exogenously specified price impact model. The algorithm we develop bears some similarity to that of Foster and Viswanathan (1994), but requires new features designed to address differences in our model.

It is also worth discussing how our model differs from that of Vayanos (2001). Both models consider the inventory of a large trader as asymmetric information. The details and the goals of the models are significantly different, however. In particular, Vayanos (2001) seeks a structural model to provide intuition for behavior of a large trader seeking to maximize utility through trade and consumption decisions. We, on the other hand, specialize to the context of minimizing execution costs in a short time horizon (e.g., one day) trade execution problem. We seek to provide specific policy recommendations for this problem, which is known to be of significant practical interest. Vayanos (2001) assumes an implicit price impact that arises endogenously through the a continuum of competitive market makers. We assume an exogenous and explicit price impact. This is important in the context of trade execution problems, since such forms of explicit price impact can be directly estimated. Moreover, the competitive market makers of Vayanos (2001) do not directly trade. Instead, they manipulate prices in anticipation of the large trader, and their effect is diluted through competition. Our model, with a strategic arbitrageur, more directly captures the idea of "front-running."

The remainder of this paper is organized as follows. The next section presents our problem formulation. Section 3 discusses how perfect Bayesian equilibrium in this model is characterized by a dynamic program. A practical algorithm for computing perfect Bayesian equilibrium behavior is developed in Section 4. This algorithm is applied in computational studies, for which results are presented and interpreted in Section 5. Several extensions of this model are discussed in Section 6. Finally, Section 7 makes some closing remarks and suggests directions for future work. Proofs of all theoretical results are presented in the appendices.

2. Problem Formulation

In this section, the optimal execution problem is formulated as a game of asymmetric information. Our formulation makes a number of simplifying assumptions and we omit several factors that are important in the practical implementation of execution strategies, for example, transient price impact and risk aversion. Our goal here is to highlight the strategic and informational aspects of execution in a streamlined fashion. However, these assumptions are discussed in more detail and a number of extensions of this basic model are presented in Section 6.

2.1. Game Structure

Consider a game that evolves over a finite horizon in discrete time steps $t = 0, \ldots, T+1$. There are two players: a trader and an arbitrageur. The trader begins with a position $x_0 \in \mathbb{R}$ in a stock, which he must liquidate by time T. Denote his position at each time t by x_t, and thus require that $x_t = 0$ for $t \geq T$. The arbitrageur begins with a position y_0. Denote his position at each time t by y_t. In general, the arbitrageur has additional flexibility and will not be limited to the same time horizon as the trader. For simplicity, this flexibility is modelled by assuming that the arbitrageur has one additional period of trading activity. In other words, though we do require that $y_{T+1} = 0$, we do not require that $y_T = 0$. This assumption will be revisited in Section 6.1.

2.2. Price Dynamics

Denote the price of the stock at time t by p_t. This price evolves according to the permanent linear price impact model given by

$$p_t = p_{t-1} + \Delta p_t = p_{t-1} + \lambda(u_t + v_t) + \epsilon_t. \tag{1}$$

Here, $\lambda > 0$ is a parameter that reflects the sensitivity of prices to trade size, and u_t and v_t are, respectively, the quantities of stock purchased by the trader and the arbitrageur at time t. Note that, given the horizon of the trader, $u_{T+1} \triangleq 0$. The positions evolve according to

$$x_t = x_{t-1} + u_t, \quad \text{and} \quad y_t = y_{t-1} + v_t.$$

The sequence $\{\epsilon_t\}$ is a normally distributed IID process with $\epsilon_t \sim N(0, \sigma_\epsilon^2)$, for some $\sigma_\epsilon > 0$. This noise sequence represents the random and exogenous fluctuations of market prices. We assume that the trading decisions u_t and v_t are made at time $t-1$, and executed at the price p_t at time t. Note that there is no drift term in the price evolution equation (1). In the intraday horizon of typical optimal execution problems, this is usually a reasonable assumption. This assumption will be revisited in Section 6.3. Further, the price impact in (1) is permanent in the sense that it is long-lived relative to the length of the time horizon T. It is stationary in the sense that the sensitivity λ is constant. In Section 6.3, we will allow for transient price impact as well non-stationary price dynamics.

2.3. Information Structure

The information structure of the game is as follows. The dynamics of the game (in particular, the parameters λ and σ_ϵ) and the time horizon T are mutually known. From the perspective of the arbitrageur, the initial position x_0 of the trader is unknown. Further, the trader's actions u_t are not directly observed. However, the arbitrageur begins with a prior distribution ϕ_0 on the trader's initial position x_0. As the game evolves over time, the arbitrageur observes the price change Δp_t at each time t. The arbitrageur updates his beliefs based on these price movements, at any time t maintaining a posterior distribution ϕ_t of the trader's current position x_t, based on his observation of the history of the game up to and including time t.

From the trader's perspective, it is assumed that everything is known. This is motivated by the fact that the arbitrageur's initial position y_0 will typically be zero and the trader can go through the same inference process as the arbitrageur to arrive at the prior distribution ϕ_0. Given a prescribed policy for the arbitrageur (for example, in equilibrium), the trader can subsequently reconstruct the arbitrageur's positions and beliefs over time, given the public observations of market price movements. We do make the assumption, however, that any deviations on the part of the arbitrageur from his prescribed policy will not mislead the trader. In our context, this assumption is important for tractability. We discuss the situation where this assumption is relaxed, and the trader does not have perfect knowledge of the arbitrageur's positions and beliefs, in Section 7.

2.4. Policies

The trader's purchases are governed by a policy, which is a sequence of functions $\pi = \{\pi_1, \ldots, \pi_T\}$. Each function π_{t+1} maps x_t, y_t, and ϕ_t, to a decision u_{t+1} at time t. Similarly, the arbitrageur follows a policy $\psi = \{\psi_1, \ldots, \psi_{T+1}\}$. Each function ψ_{t+1} maps y_t and ϕ_t to a decision v_{t+1} made at time t. Since policies for the trader and arbitrageur must result in liquidation, we require that $\pi_T(x_{T-1}, y_{T-1}, \phi_{T-1}) = -x_{T-1}$ and $\psi_{T+1}(y_T, \phi_T) = -y_T$. Denote the set of trader policies by Π and the set of arbitrageur policies by Ψ.

Note that implicit in the above description is the restriction to policies that are Markovian in the following sense: the state of the game at time t is summarized for the trader and arbitrageur by the tuples (x_t, y_t, ϕ_t) and (y_t, ϕ_t), respectively, and each player's action is only a function of his state. Further, the policies are pure strategies in the sense that, as a function of the player's state, the actions are deterministic. In general, one may wish to consider policies which determine actions as a function of the entire history of the game up to a given time, and allow randomization over the choice of action. Our assumptions will exclude equilibria from this more general class. However, it will be the case that for the equilibria that we do find, arbitrary deviations that are history dependent and/or randomized will not be profitable.

If the arbitrageur applies an action v_t and assumes the trader uses a policy $\hat{\pi} \in \Pi$, then upon observation of Δp_t at time t, the arbitrageur's beliefs are updated in a Bayesian fashion according to

$$\phi_t(S) = \Pr\left(x_t \in S \mid \phi_{t-1}, y_{t-1}, \lambda(\hat{\pi}_t(x_{t-1}, y_{t-1}, \phi_{t-1}) + v_t) + \epsilon_t = \Delta p_t\right), \tag{2}$$

for all measurable sets $S \subset \mathbb{R}$. Note that Δp_t here is an observed numerical value which could have resulted from a trader action $u_t \neq \hat{\pi}_t(x_{t-1}, y_{t-1}, \phi_{t-1})$. As such, the trader is capable of misleading the arbitrageur to distort his posterior distribution ϕ_t.

2.5. Objectives

Assume that both the trader and the arbitrageur are risk neutral and seek to maximize their expected profits (this assumption will be revisited in Section 6.2). Profit is computed according to the change of book value, which is the sum of a player's cash position and asset position, valued at the prevailing market price. Hence, the profits generated by the trader and arbitrageur between time t and time $t+1$ are, respectively,

$$p_{t+1}x_{t+1} - p_{t+1}u_{t+1} - p_t x_t = \Delta p_{t+1} x_t, \quad \text{and} \quad p_{t+1}y_{t+1} - p_{t+1}v_{t+1} - p_t y_t = \Delta p_{t+1} y_t.$$

If the trader uses policy π and the arbitrageur uses policy ψ and assumes the trader uses policy $\hat{\pi}$, the trader expects profits

$$U_t^{\pi,(\psi,\hat{\pi})}(x_t, y_t, \phi_t) \triangleq \mathsf{E}^{\pi,(\psi,\hat{\pi})}\left[\sum_{\tau=t}^{T-1} \Delta p_{\tau+1} x_\tau \;\middle|\; x_t, y_t, \phi_t\right],$$

over times $\tau = t+1, \ldots, T$. Here, the superscripts indicate that trades are executed based on π and ψ, while beliefs are updated based on $\hat{\pi}$. Similarly, the arbitrageur expects profits

$$V_t^{(\psi,\hat{\pi}),\pi}(y_t, \phi_t) \triangleq \mathsf{E}^{\pi,(\psi,\hat{\pi})}\left[\sum_{\tau=t}^{T} \Delta p_{\tau+1} y_\tau \;\middle|\; y_t, \phi_t\right],$$

over times $\tau = t+1, \ldots, T+1$. Here, the conditioning in the expectation implicitly assumes that x_t is distributed according to ϕ_t.

Note that $-U_t^{\pi,(\psi,\hat{\pi})}(x_0, y_0, \phi_0)$ is the trader's expected execution cost. For practical choices of π, ψ, and $\hat{\pi}$, we expect this quantity to be positive since the trader is likely to sell his shares for less than the initial price. To compress notation, for any π, ψ, and t, let

$$U_t^{\pi,\psi} \triangleq U_t^{\pi,(\psi,\pi)}, \quad \text{and} \quad V_t^{\psi,\pi} \triangleq V_t^{(\psi,\pi),\pi}.$$

2.6. Equilibrium Concept

As a solution concept, we consider perfect Bayesian equilibrium (Fudenberg and Tirole, 1991). This is a refinement of Nash equilibrium that rules out implausible outcomes by requiring subgame perfection and consistency with Bayesian belief updates. In particular, a policy $\pi \in \Pi$ is a *best response* to $(\psi, \hat{\pi}) \in \Psi \times \Pi$ if

$$U_t^{\pi,(\psi,\hat{\pi})}(x_t, y_t, \phi_t) = \max_{\pi' \in \Pi} \; U_t^{\pi',(\psi,\hat{\pi})}(x_t, y_t, \phi_t), \tag{3}$$

for all t, x_t, y_t, and ϕ_t. Similarly, a policy $\psi \in \Psi$ is a *best response* to $\pi \in \Pi$ if

$$V_t^{\psi,\pi}(y_t, \phi_t) = \max_{\psi' \in \Psi} \; V_t^{\psi',\pi}(y_t, \phi_t), \tag{4}$$

for all t, y_t, and ϕ_t. We define perfect Bayesian equilibrium, specialized to our context, as follows:

Definition 1. *A* **perfect Bayesian equilibrium** *(PBE) is a pair of policies* $(\pi^*, \psi^*) \in \Pi \times \Psi$ *such that:*

1. π^* *is a best response to* (ψ^*, π^*)*;*
2. ψ^* *is a best response to* π^**.*

In a PBE, each player's action at time t depends on positions x_t and/or y_t and the belief distribution ϕ_t. These arguments, especially the distribution, make computation and representation of a PBE challenging. We will settle for a more modest goal. We compute policy actions only for cases where ϕ_t is Gaussian. When the initial distribution ϕ_0 is Gaussian and players employ these PBE policies, we require that subsequent belief distributions ϕ_t determined by Bayes' rule (2) also be Gaussian. As such, computation of PBE policies over the restricted domain of Gaussian distributions is sufficient to characterize equilibrium behavior given any initial conditions involving a Gaussian prior. To formalize our approach, we now define a solution concept.

Definition 2. *A policy $\pi \in \Pi$ (or $\psi \in \Psi$) is a* **Gaussian best response** *to $(\psi, \hat{\pi}) \in \Psi \times \Pi$ (or $\pi \in \Pi$) if (3) (or (4)) holds for all t, x_t, y_t, and Gaussian ϕ_t. A* **Gaussian perfect Bayesian equilibrium** *is a pair $(\pi^*, \psi^*) \in \Pi \times \Psi$ of policies such that*

1. *π^* is a Gaussian best response to (ψ^*, π^*);*
2. *ψ^* is a Gaussian best response to π^*;*
3. *if ϕ_0 is Gaussian and arbitrageur assumes the trader uses π^* then, independent of the true actions of the trader, the beliefs $\phi_1, \dots, \phi_{T-1}$ are Gaussian.*

Note that when Gaussian PBE policies are used and the prior ϕ_0 is Gaussian, the system behavior is indistinguishable from that of a PBE since the policies produce actions that concur with PBE policies at all states that are visited.

Given a belief distribution ϕ_t, define the quantities

$$\mu_t \triangleq \mathsf{E}[x_t \mid \phi_t], \qquad \sigma_t^2 \triangleq \mathsf{E}\left[(x_t - \mu_t)^2 \,\middle|\, \phi_t\right], \quad \text{and} \quad \rho_t \triangleq \lambda\sigma_t/\sigma_\epsilon.$$

Since λ and σ_ϵ are constants, ρ_t is simply a scaled version of the standard deviation σ_t. The ratio λ/σ_ϵ acts as a normalizing constant that accounts for the informativeness of observations. The reason we consider this scaling is that it highlights certain invariants across problem instances. In Section 5.2, we will interpret the value of ρ_0 as the relative volume of the trader's activity in the marketplace. For the moment, it is sufficient to observe that if the distribution ϕ_t is Gaussian, it is characterized by (μ_t, ρ_t).

3. Dynamic Programming Analysis

In this section, we develop abstract dynamic programming algorithms for computing PBE and Gaussian PBE. We also discuss structural properties of associated value functions. The dynamic programming recursion relies on the computation of equilibria for single-stage games, and we also discuss the existence of such equilibria. The algorithms of this section are not implementable, but their treatment motivates the design of a practical algorithm that will be presented in the next section.

3.1. Stage-Wise Decomposition

The process of computing a PBE and the corresponding value functions can be decomposed into a series of single-stage equilibrium problems via a dynamic programming backward recursion. We begin by defining some notation. For each π_t, ψ_t, and u_t, define a dynamic programming operator $F_{u_t}^{(\psi_t,\hat{\pi}_t)}$ by

$$(F_{u_t}^{(\psi_t,\hat{\pi}_t)}U)(x_{t-1},y_{t-1},\phi_{t-1}) \triangleq \mathsf{E}_{u_t}^{(\psi_t,\hat{\pi}_t)}\big[\, \lambda(u_t+v_t)x_{t-1} + U(x_t,y_t,\phi_t) \;\big|\; x_{t-1},y_{t-1},\phi_{t-1}\big],$$

for all functions U, where $x_t = x_{t-1} + u_t$, $y_t = y_{t-1} + v_t$, $v_t = \psi_t(y_{t-1},\phi_{t-1})$, and ϕ_t results from the Bayesian update (2) given that the arbitrageur assumes the trader trades $\hat{\pi}_t(x_{t-1},y_{t-1},\phi_{t-1})$ while the trader actually trades u_t. Similarly, for each π_t and v_t, define a dynamic programming operator $G_{v_t}^{\pi_t}$ by

$$(G_{v_t}^{\pi_t}V)(y_{t-1},\phi_{t-1}) \triangleq \mathsf{E}_{v_t}^{\pi_t}\big[\, \lambda(u_t+v_t)y_{t-1} + V(y_t,\phi_t) \;\big|\; y_{t-1},\phi_{t-1}\big],$$

for all functions V, where $y_t = y_{t-1} + v_t$, $u_t = \pi_t(x_{t-1},y_{t-1},\phi_{t-1})$, x_{t-1} is distributed according to the belief ϕ_{t-1}, and ϕ_t results from the Bayesian update (2) given that the arbitrageur correctly assumes the trader trades u_t.

Consider Algorithm 1 for computing a PBE. In Step 1, the algorithm begins by initializing the terminal value functions U_{T-1}^* and V_{T-1}^*. These terminal value functions have a simple closed form in equilibrium. This is because, at time T, the trader must liquidate his position, hence

1: Initialize the terminal value functions U^*_{T-1} and V^*_{T-1} according to (5)–(6)
2: **for** $t = T-1, T-2, \ldots, 1$ **do**
3: Compute (π^*_t, ψ^*_t) such that for all x_{t-1}, y_{t-1}, and ϕ_{t-1},

$$\pi^*_t(x_{t-1}, y_{t-1}, \phi_{t-1}) \in \operatorname*{argmax}_{u_t} \left(F^{(\psi^*_t, \pi^*_t)}_{u_t} U^*_t \right)(x_{t-1}, y_{t-1}, \phi_{t-1})$$

$$\psi^*_t(y_{t-1}, \phi_{t-1}) \in \operatorname*{argmax}_{v_t} \left(G^{\pi^*_t}_{v_t} V^*_t \right)(y_{t-1}, \phi_{t-1})$$

4: Compute the value functions at the previous time step by setting, for all x_{t-1}, y_{t-1}, and ϕ_{t-1},

$$U^*_{t-1}(x_{t-1}, y_{t-1}, \phi_{t-1}) \leftarrow \left(F^{(\psi^*_t, \pi^*_t)}_{\pi^*_t} U^*_t \right)(x_{t-1}, y_{t-1}, \phi_{t-1})$$

$$V^*_{t-1}(y_{t-1}, \phi_{t-1}) \leftarrow \left(G^{\pi^*_t}_{\psi^*_t} V^*_t \right)(y_{t-1}, \phi_{t-1})$$

5: **end for**

Algorithm 1: PBE Solver

$\pi^*_T(x_{T-1}, y_{T-1}, \phi_{T-1}) = -x_{T-1}$. Similarly, arbitrageur must liquidate his position over times T and $T+1$. In equilibrium, he will do so optimally, thus his value function takes the form

$$\begin{aligned} V^*_{T-1}(y_{T-1}, \phi_{T-1}) &= \max_{v_T} \; \mathsf{E}\left[\lambda(-x_{T-1} + v_T) y_{T-1} - \lambda(y_{T-1} + v_T)^2 \;\middle|\; y_{T-1}, \phi_{T-1} \right] \\ &= -\lambda \left(\mu_{T-1} + \tfrac{3}{4} y_{T-1} \right) y_{T-1}, \end{aligned} \tag{5}$$

where the optimizing decision is $\psi^*_T(y_{T-1}, \phi_{T-1}) = -\frac{1}{2} y_{T-1}$. It is straightforward to derive the corresponding expression of the trader's value function,

$$\begin{aligned} U^*_{T-1}(x_{T-1}, y_{T-1}, \phi_{T-1}) &= \mathsf{E}\left[\lambda \left(-x_{T-1} - \tfrac{1}{2} y_{T-1} \right) x_{T-1} \;\middle|\; x_{T-1}, y_{T-1}, \phi_{T-1} \right] \\ &= -\lambda \left(x_{T-1} + \tfrac{1}{2} y_{T-1} \right) x_{T-1}. \end{aligned} \tag{6}$$

At each time $t < T$, equilibrium policies must satisfy the best-response conditions (3)–(4). Given the value functions U^*_t and V^*_t, these conditions decompose recursively according to to Step 3. Given such a pair (π^*_t, ψ^*_t), the value functions U^*_{t-1} and V^*_{t-1} for the prior time period are, in turn, computed in Step 4.

It is easy to see that, so long as Step 3 is carried out successfully each time it is invoked, the algorithm produces a PBE (π^*, ϕ^*) along with value functions $U^*_t = U^{\pi^*, \psi^*}_t$ and $V^*_t = V^{\psi^*, \pi^*}_t$. However, the algorithm is not implementable. For starters, the functions π^*_t, ψ^*_t, U^*_{t-1}, and V^*_{t-1}, which must be computed and stored, have infinite domains.

3.2. Linear Policies

Consider the following class of policies:

Definition 3. *A function π_t is* **linear** *if there are coefficients $a_{x,t}^{\rho_{t-1}}$, $a_{y,t}^{\rho_{t-1}}$ and $a_{\mu,t}^{\rho_{t-1}}$, which are functions of ρ_{t-1}, such that*

$$\pi_t(x_{t-1}, y_{t-1}, \phi_{t-1}) = a_{x,t}^{\rho_{t-1}} x_{t-1} + a_{y,t}^{\rho_{t-1}} y_{t-1} + a_{\mu,t}^{\rho_{t-1}} \mu_{t-1}, \tag{7}$$

for all x_{t-1}, y_{t-1}, and ϕ_{t-1}. Similarly, function ψ_t is **linear** *if there is a coefficients $b_{y,t}^{\rho_{t-1}}$ and $b_{\mu,t}^{\rho_{t-1}}$, which is a function of ρ_{t-1}, such that*

$$\psi_t(y_{t-1}, \phi_{t-1}) = b_{y,t}^{\rho_{t-1}} y_{t-1} + b_{\mu,t}^{\rho_{t-1}} \mu_{t-1}, \tag{8}$$

for all y_{t-1} and ϕ_{t-1}. A policy is linear if the component functions associated with times $1, \ldots, T-1$ are linear.

By restricting attention to linear policies and Gaussian beliefs, we can apply an algorithm similar to that presented in the previous section to compute a Gaussian PBE. In particular, consider Algorithm 2. This algorithm aims to computes a single-stage equilibrium that is linear. Further, actions and values are only computed and stored for elements of the domain for which ϕ_{t-1} is Gaussian. This is only viable if the iterates U_t^* and V_t^*, which are computed only for Gaussian ϕ_t, provide sufficient information for subsequent computations. This is indeed the case, as a consequence of the following result.

Theorem 1. *If the belief distribution ϕ_{t-1} at time is Gaussian, and the arbitrageur assumes that the trader's policy $\hat{\pi}_t$ is linear with $\hat{\pi}_t(x_{t-1}, y_{t-1}, \phi_{t-1}) = \hat{a}_{x,t}^{\rho_{t-1}} x_{t-1} + \hat{a}_{y,t}^{\rho_{t-1}} y_{t-1} + \hat{a}_{\mu,t}^{\rho_{t-1}} \mu_{t-1}$, then the belief distribution ϕ_t is also Gaussian. The mean μ_t is a linear function of y_{t-1}, μ_{t-1}, and the observed price change Δp_t, with coefficients that are deterministic functions of the scaled variance ρ_{t-1}. The scaled variance ρ_t evolves according to*

$$\rho_t^2 = \left(1 + \hat{a}_{x,t}^{\rho_{t-1}}\right)^2 \left(\frac{1}{\rho_{t-1}^2} + (\hat{a}_{x,t}^{\rho_{t-1}})^2\right)^{-1}. \tag{9}$$

In particular, ρ_t is a deterministic function of ρ_{t-1}.

It follows from this result that if π^* is linear then, for Gaussian ϕ_{t-1}, $F_{u_t}^{(\psi^*,\pi^*)}U_t^*$ only depends on values of U_t^* evaluated at Gaussian ϕ_t. Similarly, if π^* is linear then, for Gaussian ϕ_{t-1}, $G_{v_t}^{\pi^*}V_t^*$ only depends on values of V_t^* evaluated at Gaussian ϕ_t. It also follows from this theorem that Algorithm 2, which only computes actions and values for Gaussian beliefs, results in a Gaussian PBE (π^*, ψ^*). We should mention, though, that Algorithm 2 is still not implementable since the restricted domains of U_t^* and V_t^* remain infinite.

Initialize the terminal value functions U_{T-1}^* and V_{T-1}^* according to (5)–(6)
for $t = T-1, T-2, \ldots, 1$ **do**
Compute linear (π_t^*, ψ_t^*) such that for all x_{t-1}, y_{t-1}, and Gaussian ϕ_{t-1},

$$\pi_t^*(x_{t-1}, y_{t-1}, \phi_{t-1}) \in \underset{u_t}{\operatorname{argmax}} \left(F_{u_t}^{(\psi_t^*,\pi_t^*)}U_t^*\right)(x_{t-1}, y_{t-1}, \phi_{t-1})$$

$$\psi_t^*(y_{t-1}, \phi_{t-1}) \in \underset{v_t}{\operatorname{argmax}} \left(G_{v_t}^{\pi_t^*}V_t^*\right)(y_{t-1}, \phi_{t-1})$$

Compute the value functions at the previous time step by setting, for all x_{t-1}, y_{t-1}, and Gaussian ϕ_{t-1},

$$U_{t-1}^*(x_{t-1}, y_{t-1}, \phi_{t-1}) \leftarrow \left(F_{\pi_t^*}^{(\psi_t^*,\pi_t^*)}U_t^*\right)(x_{t-1}, y_{t-1}, \phi_{t-1})$$

$$V_{t-1}^*(y_{t-1}, \phi_{t-1}) \leftarrow \left(G_{\psi_t^*}^{\pi_t^*}V_t^*\right)(y_{t-1}, \phi_{t-1})$$

end for

Algorithm 2: Linear-Gaussian PBE Solver

Motivated by these observations, for the remainder of the paper, we will focus on computing equilibria of the following form:

Definition 4. *A pair of policies $(\pi^*, \psi^*) \in \Pi \times \Psi$ is a* **linear-Gaussian perfect Bayesian equilibrium** *if it is a Gaussian PBE and each policy is linear.*

3.3. Quadratic Value Functions

Closely associated with linear policies are the following class of value functions:

Definition 5. *A function U_t is* **trader-quadratic-decomposable** *(TQD) if there are coefficients*

$c^{\rho_t}_{xx,t}$, $c^{\rho_t}_{yy,t}$, $c^{\rho_t}_{\mu\mu,t}$, $c^{\rho_t}_{xy,t}$, $c^{\rho_t}_{x\mu,t}$, $c^{\rho_t}_{y\mu,t}$ *and* $c^{\rho_t}_{0,t}$, *which are functions of* ρ_t, *such that*

(10)

$$U_t(x_t, y_t, \phi_t) = -\lambda\bigg(\tfrac{1}{2}c^{\rho_t}_{xx,t}x_t^2 + \tfrac{1}{2}c^{\rho_t}_{yy,t}y_t^2 + \tfrac{1}{2}c^{\rho_t}_{\mu\mu,t}\mu_t^2 + c^{\rho_t}_{xy,t}x_t y_t + c^{\rho_t}_{x\mu,t}x_t\mu_t + c^{\rho_t}_{y\mu,t}y_t\mu_t - \frac{\sigma_\epsilon^2}{\lambda^2}c^{\rho_t}_{0,t}\bigg),$$

for all x_t, y_t, *and* ϕ_t. *A function* V_t *as* **arbitrageur-quadratic-decomposable** *(AQD) if there are coefficients* $d^{\rho_t}_{yy,t}$, $d^{\rho_t}_{\mu\mu,t}$, $d^{\rho_t}_{y\mu,t}$ *and* $d^{\rho_t}_{0,t}$, *which are functions of* ρ_t, *such that*

$$V_t(y_t, \phi_t) = -\lambda\left(\tfrac{1}{2}d^{\rho_t}_{yy,t}y_t^2 + \tfrac{1}{2}d^{\rho_t}_{\mu\mu,t}\mu_t^2 + d^{\rho_t}_{y\mu,t}y_t\mu_t - \frac{\sigma_\epsilon^2}{\lambda^2}d^{\rho_t}_{0,t}\right), \tag{11}$$

for all y_t *and* ϕ_t.

In equilibrium, U^*_{T-1} and V^*_{T-1} are given by Step 1 of Algorithm 2, and hence are TQD/AQD. The following theorem captures how TQD and AQD structure preserved in the dynamic programming recursion given linear policies.

Theorem 2. *If* U^*_t *is TQD and* V^*_t *is AQD, and Step 3 of Algorithm 2 produces a linear pair* (π^*_t, ψ^*_t), *then* U^*_{t-1} *and* V^*_{t-1}, *defined by Step 4 of Algorithm 2 are TQD and AQD, respectively.*

Hence, each pair of value functions generated by Algorithm 2 is TQD/AQD. A great benefit of this property comes from the fact that, for a fixed value of ρ_t, each associated value function can be encoded using just a few parameters.

3.4. Simplified Conditions for Equilibrium

Algorithm 2 relies for each t on existence of a pair (π^*_t, ψ^*_t) of linear functions that satisfy single-stage equilibrium conditions. In general, this would require verifying that each policy function is the Gaussian best response for all possible states. The following theorem provides a much simpler set of conditions. In Section 4, we will exploit these conditions in order to compute equilibrium policies.

Theorem 3. *Suppose that* U^*_t *and* V^*_t *and TQD/AQD value functions specified by* (10)–(11), *and* (π^*_t, ψ^*_t) *are linear policies specified by* (7)–(8). *Assume that, for all* ρ_{t-1}, *the policy coefficients* *satisfy the first order conditions*

$$0 = \left(\rho_t^2 c^{\rho_t}_{\mu\mu,t} + 2\rho_t c^{\rho_t}_{x\mu,t} + c^{\rho_t}_{xx,t}\right)\left(a^{\rho_{t-1}}_{x,t}\right)^3 + \left(3c^{\rho_t}_{xx,t} + 3\rho_t c^{\rho_t}_{x\mu,t} - 1\right)\left(a^{\rho_{t-1}}_{x,t}\right)^2 + \left(3c^{\rho_t}_{xx,t} + \rho_t c^{\rho_t}_{x\mu,t} - 2\right)a^{\rho_{t-1}}_{x,t} + c^{\rho_t}_{xx,t} - 1, \tag{12}$$

$$a^{\rho_{t-1}}_{y,t} = -\frac{\left(b^{\rho_{t-1}}_{y,t} + 1\right)\left(c^{\rho_t}_{xy,t} + \alpha_t c^{\rho_t}_{y\mu,t}\right)}{c^{\rho_t}_{xx,t} + (\alpha_t + 1)c^{\rho_t}_{x\mu,t} + \alpha_t c^{\rho_t}_{\mu\mu,t}}, \tag{13}$$

$$a^{\rho_{t-1}}_{\mu,t} = -\frac{a^{\rho_{t-1}}_{x,t} b^{\rho_{t-1}}_{\mu,t}\left(c^{\rho_t}_{xy,t} + \alpha_t c^{\rho_t}_{y\mu,t}\right) + \alpha_t\left(c^{\rho_t}_{x\mu,t} + \alpha_t c^{\rho_t}_{\mu\mu,t}\right)/\rho^2_{t-1}}{a^{\rho_{t-1}}_{x,t}\left(c^{\rho_t}_{xx,t} + (\alpha_t + 1)c^{\rho_t}_{x\mu,t} + \alpha_t c^{\rho_t}_{\mu\mu,t}\right)}, \tag{14}$$

$$b^{\rho_{t-1}}_{y,t} = \frac{1 - d^{\rho_t}_{y\mu,t} a^{\rho_{t-1}}_{y,t}}{d^{\rho_t}_{yy,t}} - 1, \qquad b^{\rho_{t-1}}_{\mu,t} = -\frac{\left(1 + a^{\rho_{t-1}}_{\mu,t} + a^{\rho_{t-1}}_{x,t}\right)d^{\rho_t}_{y\mu,t}}{d^{\rho_t}_{yy,t}}, \tag{15}$$

and the second order conditions

$$c^{\rho_t}_{xx,t} + (\alpha_t + 1)c^{\rho_t}_{x\mu,t} + \alpha_t c^{\rho_t}_{\mu\mu,t} > 0, \qquad d^{\rho_t}_{yy,t} > 0, \tag{16}$$

where the quantities α_t and ρ_t satisfy

$$\alpha_t = \frac{a^{\rho_{t-1}}_{x,t}\left(1 + a^{\rho_{t-1}}_{x,t}\right)}{1/\rho^2_{t-1} + \left(a^{\rho_{t-1}}_{x,t}\right)^2}, \qquad \rho_t^2 = \left(1 + a^{\rho_{t-1}}_{x,t}\right)^2 \left(\frac{1}{\rho^2_{t-1}} + \left(a^{\rho_{t-1}}_{x,t}\right)^2\right)^{-1}. \tag{17}$$

Then, (π_t^*, ψ_t^*) satisfy the single-stage equilibrium conditions

$$\pi_t^*(x_{t-1}, y_{t-1}, \phi_{t-1}) \in \operatorname*{argmax}_{u_t} \left(F_{u_t}^{(\psi_t^*, \pi_t^*)} U_t^*\right)(x_{t-1}, y_{t-1}, \phi_{t-1}),$$
$$\psi_t^*(y_{t-1}, \phi_{t-1}) \in \operatorname*{argmax}_{v_t} \left(G_{v_t}^{\pi_t^*} V_t^*\right)(y_{t-1}, \phi_{t-1}),$$

for all x_{t-1}, y_{t-1}, and Gaussian ϕ_{t-1}.

Note that, while this theorem provides sufficient conditions for linear policies satisfying equilibrium conditions, it does not guarantee the existence or uniqueness of such policies. These remain an open issues. However, we support the plausibility of existence through the following result on Gaussian best responses to linear policies. It asserts that, if ψ_t and $\hat{\pi}_t$ are linear, then there is a linear best-response π_t for the trader in the single-stage game. Similarly, if π_t is linear then there is a linear best-response ψ_t for the arbitrageur in the single-stage game.

Theorem 4. *If U_t is TQD, ψ_t is linear, and $\hat{\pi}_t$ is linear, then there exists a linear π_t such that*

$$\pi_t(x_{t-1}, y_{t-1}, \phi_{t-1}) \in \underset{u_t}{\operatorname{argmax}} \left(F_{u_t}^{(\psi_t, \hat{\pi}_t)} U_t\right)(x_{t-1}, y_{t-1}, \phi_{t-1}),$$

for all x_{t-1}, y_{t-1}, and Gaussian ϕ_{t-1}, so long as the optimization problem is bounded. Similarly, if V_t is AQD and π_t is linear then there exists a linear ψ_t such that

$$\psi_t(y_{t-1}, \phi_{t-1}) \in \underset{v_t}{\operatorname{argmax}} \left(G_{v_t}^{\pi_t} V_t\right)(y_{t-1}, \phi_{t-1}),$$

for all y_{t-1} and Gaussian ϕ_{t-1}, so long as the optimization problem is bounded.

Based on these results, if the trader (arbitrageur) assumes that the arbitrageur (trader) uses a linear policy then it suffices for the trader (arbitrageur) to restrict himself to linear policies. Though not a proof of existence, this observation that the set of linear policies is closed under the operation of best response motivates an aim to compute linear-Gaussian PBE.

4. Algorithm

The previous section presented abstract algorithms and results that lay the groundwork for the development of a practical algorithm which we will present in this section. We begin by discussing a parsimonious representation of policies.

4.1. Representation of Policies

Algorithm 2 takes as input three values that parameterize our model: $(\lambda, \sigma_\epsilon, T)$. The algorithm output can be encoded in terms of coefficients $\{a_{x,t}^{\rho_{t-1}}, a_{y,t}^{\rho_{t-1}}, a_{\mu,t}^{\rho_{t-1}}, b_{y,t}^{\rho_{t-1}}, b_{\mu,t}^{\rho_{t-1}}\}$, for every $\rho_{t-1} > 0$ and each time step[1] $t = 1, \ldots, T-1$. These coefficients parameterize linear-Gaussian PBE policies. Note that the output depends on λ and σ_ϵ only through ρ_t. Hence, given any λ and σ_ϵ with the same ρ_t, the algorithm obtains the same coefficients. This means that the algorithm need only be executed once to obtain solutions for all choices of λ and σ_ϵ.

[1]Recall, from the discussion in Section 3.1, that $a_{x,T}^{\rho_{t-1}} = -1$, $a_{y,T}^{\rho_{t-1}} = a_{\mu,T}^{\rho_{t-1}} = 0$, $b_{y,T}^{\rho_{t-1}} = -1/2$, $b_{\mu,T+1}^{\rho_{t-1}} = b_{\mu,T}^{\rho_{t-1}} = 0$, and $b_{y,T+1}^{\rho_{t-1}} = -1$, for all ρ_{t-1}.

Now, for each t, the policy coefficients are deterministic functions of ρ_{t-1}. For a fixed value of ρ_{t-1}, the coefficients can be stored as five numerical values. However, it is not feasible to simultaneously store coefficients associated with all possible values of ρ_{t-1}. Fortunately, given a linear policy for the trader, Theorem 1 establishes ρ_t is a deterministic function of ρ_{t-1}. Thus, the initial value ρ_0 determines all subsequent values of ρ_t. It follows that, for a fixed value of ρ_0, over the relevant portion of its domain, a linear-Gaussian PBE can be encoded in terms of $5(T-1)$ numerical values. We will design an algorithm that aims to compute these $5(T-1)$ parameters, which we will denote by $\{a_{x,t}, a_{y,t}, a_{\mu,t}, b_{y,t}, b_{\mu,t}\}$, for $t = 1, \ldots, T-1$. These parameters allow us to determine PBE actions at all visited states, so long as the initial value of ρ_0 is fixed.

4.2. Searching for Equilibrium Variances

The parameters $\{a_{x,t}, a_{y,t}, a_{\mu,t}, b_{y,t}, b_{y,t}\}$ characterize linear-Gaussian PBE policies restricted to the sequence $\rho_0, \ldots, \rho_{T-1}$ generated in the linear-Gaussian PBE. We do not know in advance what this sequence will be, and as such, we seek simultaneously compute this sequence alongside the policy parameters.

One way to proceed, reminiscent of the bisection method employed by Kyle (1985) and Foster and Viswanathan (1994) would be to conjecture a value for ρ_{T-1}. Given a candidate value $\hat{\rho}_{T-1}$, the preceding values $\hat{\rho}_{T-2}, \ldots, \hat{\rho}_0$, along with policy parameters for times $T-1, \ldots, 1$, can be computed by sequentially solving the equations (12)–(17) for single-stage equilibria. The resulting policies form a linear-Gaussian PBE, restricted to the sequence $\hat{\rho}_0, \ldots, \hat{\rho}_{T-1}$ that they would generate if $\rho_0 = \hat{\rho}_0$. One can then seek a value of $\hat{\rho}_{T-1}$ such that the resulting $\hat{\rho}_0$ is indeed equal to ρ_0. This can be accomplished, for example, via bisection search.

The bisection method can be numerically unstable, however. This is because, the belief update equation (9) is used to sequentially compute the values $\hat{\rho}_{T-2}, \ldots, \hat{\rho}_0$ backwards in time. When the target value of ρ_0 is very large, small changes in $\hat{\rho}_{T-1}$ can result in very large changes in $\hat{\rho}_0$, making it difficult to match the precisely value of ρ_0.

To avoid this numerical instability, consider Algorithm 3. This algorithm maintains a guess $\hat{\pi}$ of the equilibrium policy of the trader, and, along with the initial value ρ_0, this is used to generate the sequence $\hat{\rho}_1, \ldots, \hat{\rho}_{T-1}$ by applying the belief update equation (9) *forward* in time.

This sequence of values is then used in the single-stage equilibrium conditions to solve for policies (π^*, ψ^*). A sequence of values $\hat{\rho}_1, \ldots, \hat{\rho}_{T-1}$ is then computed forward in time using the policy π^*. If this sequence matches the sequence generated by the guess $\hat{\pi}$, then the algorithm has converged. Otherwise, the algorithm is repeated with a new guess policy that is a convex combination of $\hat{\pi}$ and π^*. Since this algorithm only ever applies the belief equation (9) forward in time, it does not suffer from the numerical instabilities of the bisection method.

Note that Step 6 of the algorithm treats ρ_{t-1} as a free variable that is solved alongside the policy parameters $\{a_{x,t}, a_{y,t}, a_{\mu,t}, b_{y,t}, b_{\mu,t}\}$. These variables are computed by simultaneously solving the system of equations (12)–(17) for single-stage equilibrium. To be precise, $a_{x,t}$ is obtained by solving the cubic polynomial equation (12) numerically. Given a value for $a_{x,t}$, the remaining parameters $\{a_{y,t}, a_{\mu,t}, b_{y,t}, b_{\mu,t}\}$ are be obtained by solving the linear system of equations (13)–(15), while ρ_{t-1} is obtained through (17) . It can then be verified that the second order condition (16) holds. Algorithm 3 is implementable and we use it in computational studies presented in the next section.

1: Initialize $\hat{\pi}$ to an equipartitioning policy
2: **for** $k = 1, 2, \ldots$ **do**
3: Compute $\hat{\rho}_1, \ldots, \hat{\rho}_{T-1}$ according to the initial value ρ_0 and the policy $\hat{\pi}$ by (9)
4: Initialize the terminal value functions U^*_{T-1} and V^*_{T-1} according to (5)–(6)
5: **for** $t = T-1, T-2, \ldots, 1$ **do**
6: Compute linear (π^*_t, ψ^*_t) and ρ_{t-1} solving the single-stage equilibrium conditions (12)–(17), assuming that $\rho_t = \hat{\rho}_t$
7: Compute the value functions U^*_{t-1} and V^*_{t-1} at the previous time step given (π^*_t, ψ^*_t)
8: **end for**
9: Compute $\tilde{\rho}_1, \ldots, \tilde{\rho}_{T-1}$ according to the initial value ρ_0 and the policy π^* by (9)
10: **if** $\hat{\rho} = \tilde{\rho}$ **then**
11: **return**
12: **else**
13: Set $\hat{\pi} \leftarrow \gamma_k \hat{\pi} + (1-\gamma_k)\pi^*$, where $\gamma_k \in [0, 1)$ is a step-size
14: **end if**
15: **end for**

Algorithm 3: Linear-Gaussian PBE Solver with Variance Search

5. Computational Results

In this section, we present computational results generated using Algorithm 3. In Section 5.1, we introduce some alternative, intuitive policies which will serve as a basis of comparison to the linear-Gaussian PBE policy. In Section 5.2, we discuss the importance of the parameter $\rho_0 \triangleq \lambda\sigma_0/\sigma_\epsilon$ in the qualitative behavior of the Gaussian PBE policy and interpret ρ_0 as a measure of the "relative volume" of the trader's activity in the marketplace. In Section 5.3, we discuss the relative performance of the policies from the perspective of the execution cost of the trader. Here, we demonstrate experimentally that the Gaussian PBE policy can offer substantial benefits. In Section 5.4, we examine the signaling that occurs through price movements. Finally, in Section 5.5, we highlight the fact that the PBE policy is adaptive and dynamic, and seeks to exploit exogenous market fluctuations in order to minimize execution costs.

5.1. Alternative Policies

In order to understand the behavior of linear-Gaussian PBE policies, we first define two alternative policies for the trader for the purpose of comparison. In the absence of an arbitrageur, it is optimal for the trader to minimize execution costs by partitioning his position into T equally sized blocks and liquidating them sequentially over the T time periods, as established by Bertsimas and Lo (1998). We refer to the resulting policy π^{EQ} as an *equipartitioning* policy. It is defined by

$$\pi_t^{\mathrm{EQ}}(x_{t-1}, y_{t-1}, \phi_{t-1}) \triangleq -\frac{1}{T-t+1} x_{t-1},$$

for all t, x_{t-1}, y_{t-1}, and ϕ_{t-1}.

Alternatively, the trader may wish to liquidate his position in a way so as to reveal as little information as possible to the arbitrageur. Trading during the final two time periods $T-1$ and T does not reveal information to the arbitrageur in a fashion that can be exploited. This is because, as discussed in Section 3.1, the arbitrageur's optimal trades at time T and $T+1$ are $v_T = -y_{T-1}/2$ and $v_{T+1} = -y_T$, respectively, and these are independent of any belief of the arbitrageur with respect to the trader's position. Given that the trader is free to trade over these two time periods without any information leakage, it is natural to minimize execution cost by equipartitioning over these two time periods. Hence, define the *minimum revelation* policy π^{MR} to be a policy that liquidates the trader's position evenly across only the last two time periods. That is,

$$\pi_t^{\mathrm{MR}}(x_{t-1}, y_{t-1}, \phi_{t-1}) \triangleq \begin{cases} 0 & \text{if } t < T-1, \\ -\frac{1}{2}x_{t-1} & \text{if } t = T-1, \\ -x_{t-1} & \text{if } t = T, \end{cases}$$

for all t, x_{t-1}, y_{t-1}, and ϕ_{t-1}.

5.2. Relative Volume

Observed in Section 4.1, linear-Gaussian PBE policies are determined as a function of the composite parameter $\rho_0 \triangleq \lambda\sigma_0/\sigma_\epsilon$. In order to interpret this parameter, consider the dynamics of price changes,

$$\Delta p_t = \lambda(u_t + v_t) + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2).$$

Here, ϵ_t is interpreted as the exogenous, random component of price changes. Alternatively, one can imagine the random component of price changes are arising from the price impact of "noise traders". Denote by z_t the total order flow from noise traders at time t, and consider a model where

$$\Delta p_t = \lambda(u_t + v_t + z_t), \quad z_t \sim N(0, \sigma_z^2).$$

If $\sigma_\epsilon = \lambda\sigma_z$, these two models are equivalent. In that case,

$$\rho_0 \triangleq \frac{\lambda\sigma_0}{\sigma_\epsilon} = \frac{\sigma_0}{\sigma_z}.$$

In other words, ρ_0 can be interpreted as the ratio of the uncertainty of the total volume of the trader's activity to the per period volume of noise trading. As such, we refer to ρ_0 as the *relative volume.*

We shall see in the following sections that, qualitatively, the performance and behavior of Gaussian PBE policies are determined by the magnitude of ρ_0. In the high relative volume regime, when ρ_0 is large, either the initial position uncertainty σ_0 is very large or the volatility σ_z of the noise traders is very small. In these cases, from the perspective of the arbitrageur, the trader's activity contributes a significant informative signal which can be decoded in the context of less significant exogenous random noise. Hence, the trader's activity early in the time horizon reveals significant information which can be exploited by the arbitrageur. Thus, it may be better for the trader to defer his liquidation until the end of the time horizon.

Alternatively, in the low relative regime, when ρ_0 is small, the arbitrageur cannot effectively distinguish the activity of the trader from the noise traders in the market. Hence, the trader is free to distribute his trades across the time horizon so as to minimize market impact, without fear of front-running by the arbitrageur.

5.3. Policy Performance

Consider a pair of policies (π, ψ), and assume that the arbitrageur begins with a position $y_0 = 0$ and an initial belief $\phi_0 = N(0, \sigma_0^2)$. Given an initial position x_0, the trader's expected profit is $U_0^{\pi,\psi}(x_0, 0, \phi_0)$. One might imagine, however, that the initial position x_0 represents one of many different trials where the trader liquidates positions. It makes sense for this distribution of x_0 over trials to be consistent with the arbitrageurs belief ϕ_0, since this belief could be based on past trials. Given this distribution, averaging over trials results in expected profit $\mathsf{E}[U_0^{\pi,\psi}(x_0, 0, \phi_0) \mid \phi_0]$. Alternatively, if the trader liquidates his entire position immediately, the expected profit becomes $\mathsf{E}[-\lambda x_0^2 \mid \phi_0] = -\lambda \sigma_0^2$. We define the *trader's normalized expected profit* $\bar{U}(\pi, \psi)$ to be the ratio of these two quantities. When the trader's value function is TQD, this takes the form

$$\bar{U}(\pi, \psi) \triangleq \frac{\mathsf{E}\left[U_0^{\pi,\psi}(x_0, 0, \phi_0) \;\middle|\; \phi_0 \right]}{\lambda \sigma_0^2} = -\frac{1}{2} c_{xx,0}^{\rho_0} + \frac{1}{\rho_0^2} c_{0,0}^{\rho_0},$$

where $c_{xx,0}^{\rho_0}$ and $c_{0,0}^{\rho_0}$ are the trader's appropriate value function coefficients at time $t = 0$.

Analogously, the *arbitrageur's normalized expected profit* $\bar{V}(\pi, \psi)$ is defined to be the expected profit of the arbitrageur normalized by the expected immediate liquidating cost of the trader. When the arbitrageur's value function is AQD, this takes the form

$$\bar{V}(\pi, \psi) \triangleq \frac{\mathsf{E}\left[V_0^{\psi,\pi}(x_0, 0, \phi_0) \;\middle|\; \phi_0 \right]}{\lambda \sigma_0^2} = \frac{1}{\rho_0^2} d_{0,0}^{\rho_0},$$

Now, let (π^*, ψ^*) denote a linear-Gaussian PBE. Since the corresponding value functions are TQD/AQD, the normalized expected profits depend on the parameters $\{\sigma_0, \lambda, \sigma_\epsilon\}$ only through the relative volume parameter $\rho_0 \triangleq \lambda\sigma_0/\sigma_\epsilon$.

Similarly, given the equipartitioning policy π^{EQ}, define ψ^{EQ} to be the optimal response of the arbitrageur to the trader's policy π^{EQ}. This best response policy can be computed by solving the linear-quadratic control problem corresponding to (4), via dynamic programming. The policy takes the form

$$\psi_t^{\text{EQ}}(y_{t-1}, \mu_{t-1}) = \begin{cases} \frac{-1}{T+2-t} y_{t-1} - \frac{(T-t)(T-t+3)}{2(T+1-t)(T+2-t)} \mu_{t-1} & \text{if } 1 \leq t \leq T, \\ -y_T & \text{otherwise.} \end{cases}$$

Using a similar argument as above, it is easy to see that $\bar{U}(\pi^{\text{EQ}}, \psi^{\text{EQ}})$ and $\bar{V}(\pi^{\text{EQ}}, \psi^{\text{EQ}})$ are also functions of the parameter ρ_0.

Finally, given the minimum revelation policy π^{MR}, define ψ^{MR} to be the optimal response of the arbitrageur to the trader's policy π^{MR}. It can be shown that, when $y_0 = 0$ and $\mu_0 = 0$, the best response of the arbitrageur to the minimum revelation policy is to do nothing–since no information is revealed by the trader in a useful fashion, there is no opportunity to front-run. Hence,

$$\bar{U}(\pi^{\text{MR}}, \psi^{\text{MR}}) = \frac{\mathsf{E}\left[-\frac{1}{2}\lambda x_0^2 - \frac{1}{4}\lambda x_0^2 \;\middle|\; \phi_0 \right]}{\lambda \sigma_0^2} = -\frac{3}{4}, \qquad \bar{V}(\pi^{\text{MR}}, \psi^{\text{MR}}) = 0.$$

In Figure 1, the normalized expected profits of various policies are plotted as functions of the relative volume ρ_0, for a time horizon $T = 20$. In all scenarios, as one might expect, the trader's profit is negative while the arbitrageur's profit is positive. In all cases, the trader's profit under the Gaussian PBE policy dominates that under either the equipartitioning policy or the minimum revelation policy. This difference is significant in moderate to high relative volume regimes.

In the high relative volume regime, the equipartitioning policy fares particularly badly from the perspective of the trader, performing up to a factor of 2 worse than the Gaussian PBE policy.

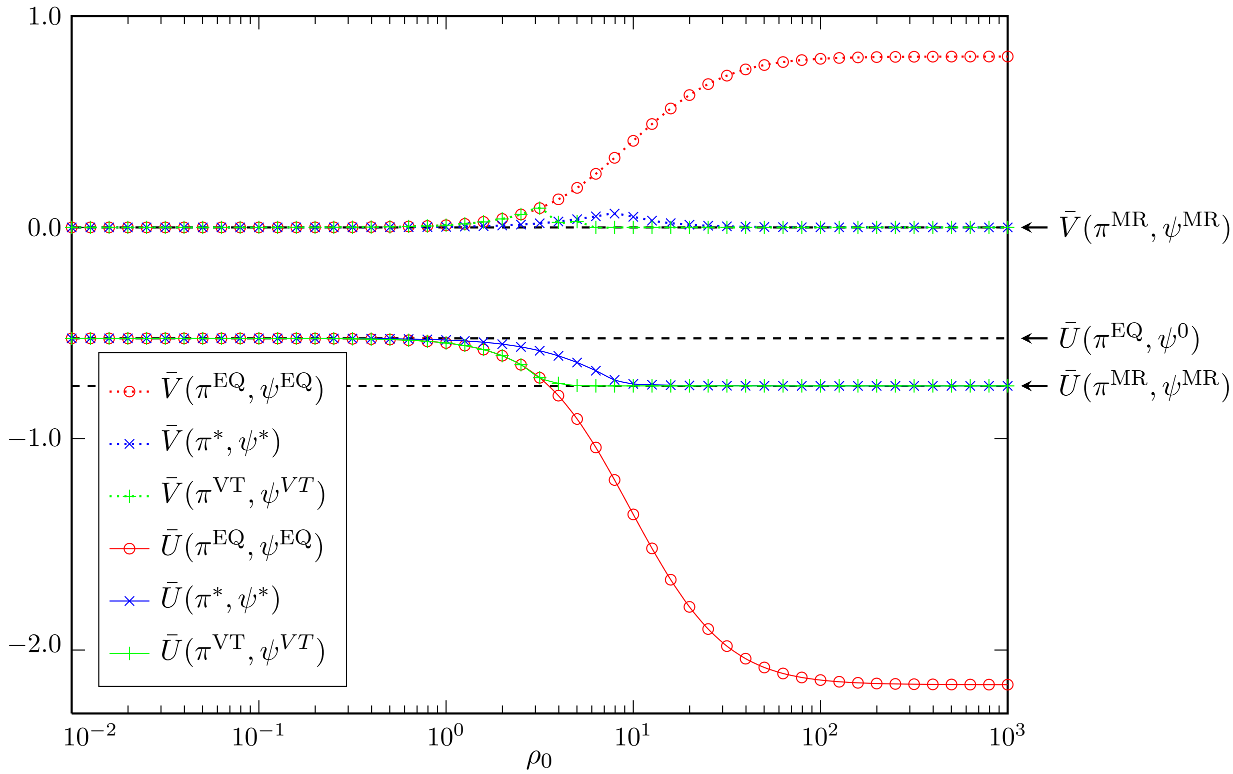


Figure 1: The normalized expected profit of trading strategies for the time horizon $T = 20$.

This effect becomes more pronounced over longer time horizons. The minimum revelation policy performs about as well as the PBE policy. Asymptotically as $\rho_0 \uparrow \infty$, these policies offer equivalent performance in the sense that $\bar{U}(\pi^*, \psi^*) \uparrow \bar{U}(\pi^{\mathrm{MR}}, \psi^{\mathrm{MR}}) = 3/4$.

On the other hand, in the low relative volume regime, the equipartitioning policy and the PBE policy perform comparably. Indeed, define ψ^0 by $\psi_t^0 \triangleq 0$ for all t (that is, no trading by the arbitrageur). In the absence of an arbitrageur, equipartitioning is the optimal policy for the trader, and backward recursion can be used to show that

$$\bar{U}(\pi^{\mathrm{EQ}}, \psi^0) = \frac{T+1}{2T} \approx \frac{1}{2}.$$

Asymptotically as $\rho_0 \downarrow 0$, $\bar{U}(\pi^{\mathrm{EQ}}, \psi^{\mathrm{EQ}}) \downarrow \bar{U}(\pi^{\mathrm{EQ}}, \psi^0)$ and $\bar{U}(\pi^*, \psi^*) \downarrow \bar{U}(\pi^{\mathrm{EQ}}, \psi^0)$. Thus, when the relative volume is low, the effect of the arbitrageur becomes negligible when ρ_0 is sufficiently small.

From the perspective of the arbitrageur in equilibrium, $\bar{V}(\pi^*, \psi^*) \to 0$ as $\rho \to \pm\infty$. In the low relative volume regime, the arbitrageur cannot distinguish the past activity of the trader from noise, and hence is not able to profitably predict and exploit the trader's future activity. In the high relative volume regime, as we shall see in Section 5.5, the trader conceals his position from the arbitrageur by deferring trading until the end of the horizon. Here, as with the minimum revelation policy, the arbitrageur is not able to profitably exploit the trader. Since the arbitrageur can choose not to trade at each period, his best response to any trading strategy should lead to non-negative expected profit. In light of these observations, we can easily infer that in equilibrium the arbitrageur's profit curve should have at least one local maximum.

Both the equipartitioning and minimum revelation policies trade at a constant rate, but over different, extremal time intervals: the equipartitioning policy uses the entire time horizon, while the minimum revelation policy uses only the last two time periods. A fairer benchmark policy might consider optimizing the choice of time interval. Define the *variable time* policy π^{VT} as follows: given the value ρ_0, select the τ such that trading at a constant rate $u_t = -\frac{x_0}{\tau}$ over the last τ time periods results in the highest expected profit for the trader, assuming that the arbitrageur uses a best response policy. Define ψ^{VT} to be the best response of the arbitrageur to π^{VT}. The variable time policy partially accounts for the presence of an arbitrary, and the expected profit with the variable time strategy will always be better that of equipartitioning or minimum revelation. This is demonstrated by the $\bar{U}(\pi^{\mathrm{VT}}, \psi^{\mathrm{VT}})$ curve in Figure 1. However, the trader still fares better with an equilibrium policy, particularly in the intermediate relative volume range, where the difference is close to 20%.[2]

Examining Figure 1, it is clear that, in equilibrium, the sum of the normalized profits of the trader and the arbitrageur is negative, and the magnitude of sum is larger than the magnitude of the loss incurred by the trader in the absence of the arbitrageur. Define the *spill-over* to be the quantity

$$\bar{U}(\pi^{\mathrm{EQ}}, \psi^0) - \Big(\bar{U}(\pi^*, \psi^*) + \bar{V}(\pi^*, \psi^*)\Big) .$$

This is the difference between the normalized expected profit of the trader in the absence of the arbitrageur, under the optimal equipartitioning policy, and the combined normalized expected profits of the trader and arbitrageur in equilibrium. The spill-over measures the benefit of the

[2] In practice, improvements of as low as 0.01% are considered significant.

arbitrageur's presence to the other participants of the system. Note that this benefit is positive, and it is most significant in the high relative volume regime.

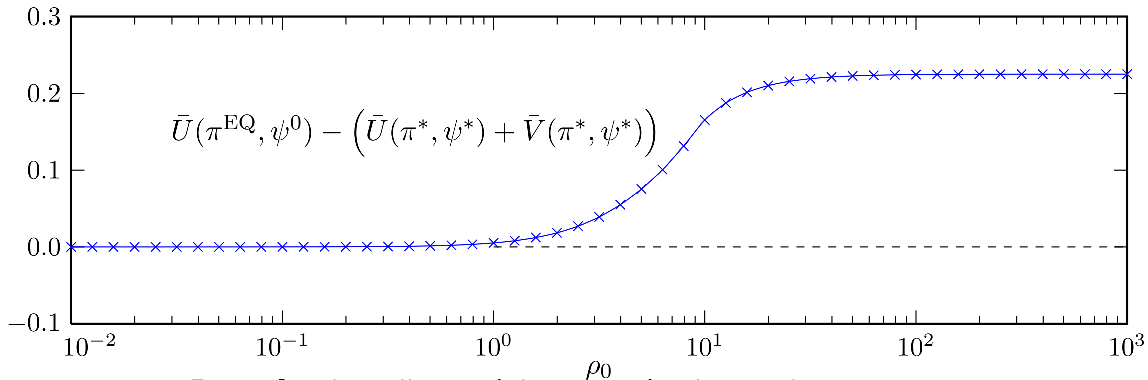


Figure 2: The spill-over of the system for the time horizon $T = 20$.

In addition to the discussion of expected profits above, we can consider the variance of the trader's profits under different policies. Given a pair of policies (π, ψ), define the *trader's normalized variance of profit* $\mathrm{Var}_U(\pi, \psi)$ as the variance under the policies (π, ψ) relative to the variance of immediate liquidation. In other words,

$$\mathrm{Var}_U(\pi, \psi) = \frac{\mathrm{Var}^{\pi,\psi}\left(\sum_{\tau=0}^{T-1} \Delta p_{\tau+1} x_\tau \;\middle|\; \phi_0\right)}{\mathrm{Var}\left(-\lambda x_0^2 + \epsilon_1 x_0 \;\middle|\; \phi_0\right)} = \frac{\mathrm{Var}^{\pi,\psi}\left(\sum_{\tau=0}^{T-1} \Delta p_{\tau+1} x_\tau \;\middle|\; \phi_0\right)}{2\lambda^2\sigma_0^4 + \sigma_\epsilon^2\sigma_0^2},$$

where, as before, the expectations are taken assuming the policies (π, ψ) are used, $y_0 = \mu_0 = 0$, and $x_0 \sim \phi_0 = N(0, \sigma_0^2)$. Similarly, it is possible to see that, for a pair of linear policies (π, ψ), the trader's normalized variance of profit depends on the model parameters $\{\sigma_0, \lambda, \sigma_\epsilon\}$ only through ρ_0.

In Figure 3, the trader's normalized variance of profit is plotted under the different policies. The lowest variance occurs when the trader equipartitions and there is no arbitrageur, this is the curve $\mathrm{Var}_U(\pi^{\mathrm{EQ}}, \psi^0)$. When the arbitrageur is present, however, the variance in equilibrium $\mathrm{Var}_U(\pi^*, \psi^*)$ is less than either when the trader equipartitions (i.e., the curve $\mathrm{Var}_U(\pi^{\mathrm{EQ}}, \psi^{\mathrm{EQ}})$) or employs the minimum revelation policy (i.e., the curve $\mathrm{Var}_U(\pi^{\mathrm{MR}}, \psi^{\mathrm{MR}})$). Figure 4 shows the entire cumulative distribution function of the trader's normalized profit under various relative volume regimes. Given the presence of the arbitrageur, the equilibrium policy has second-order dominance over equipartitioning in all relative volume regimes.

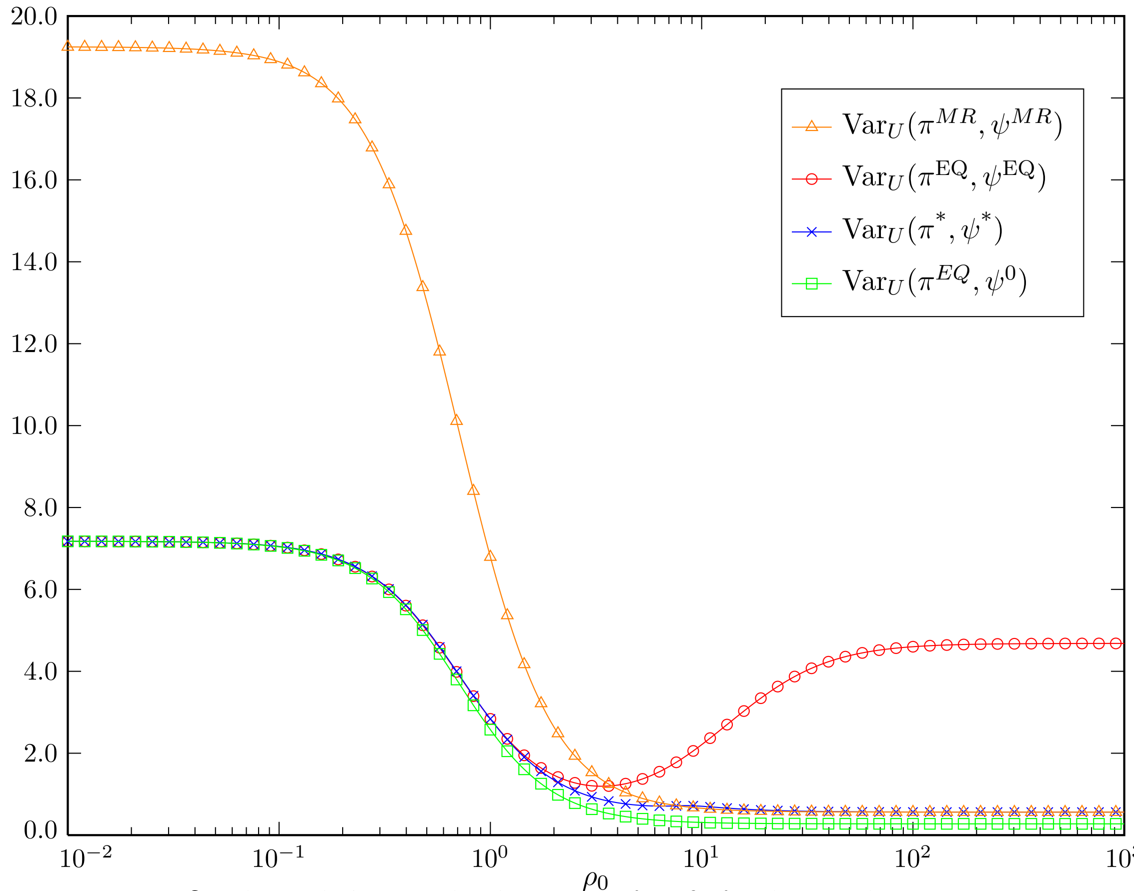
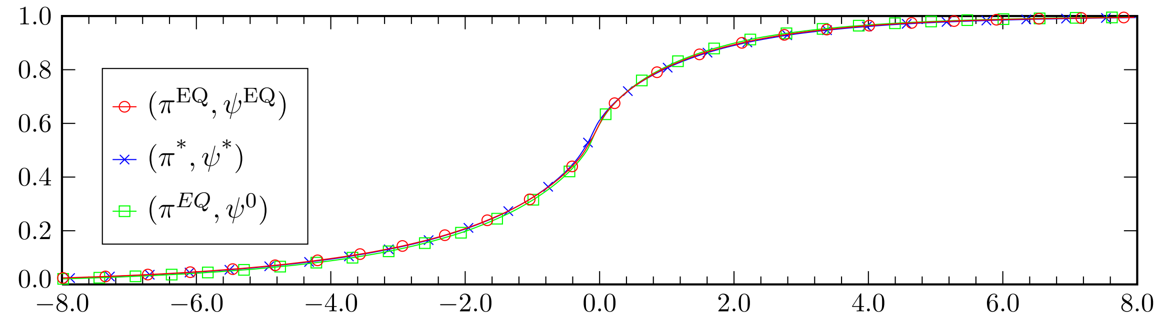


Figure 3: The trader's normalized variance of profit for the time horizon $T = 20$.

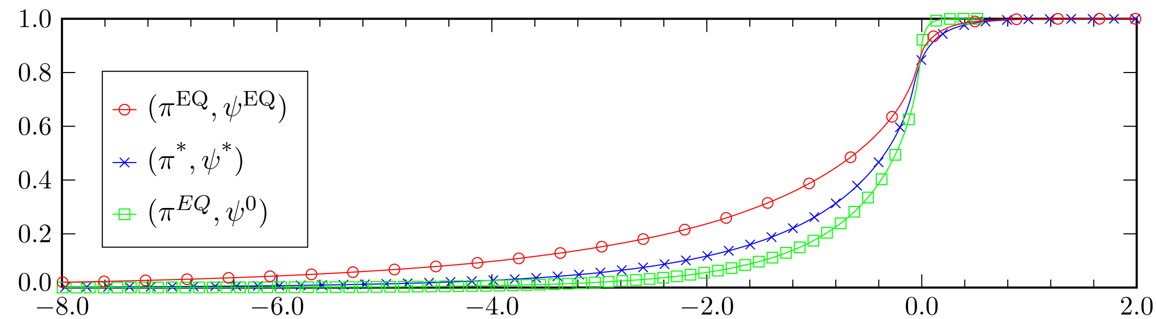
5.4. Signaling

An important aspect of the linear-Gaussian PBE policy is that it accounts for information conveyed through price movements. In order to understand this feature, define the *relative uncertainty* to be the standard deviation of the arbitrageur's belief about the trader's position at time t, relative to that of the belief at time 0; i.e., the ratio σ_t/σ_0. By considering the evolution of relative uncertainty over time for the Gaussian PBE policy versus the equipartitioning and minimum revelation policies, we can study the comparative signaling behavior.

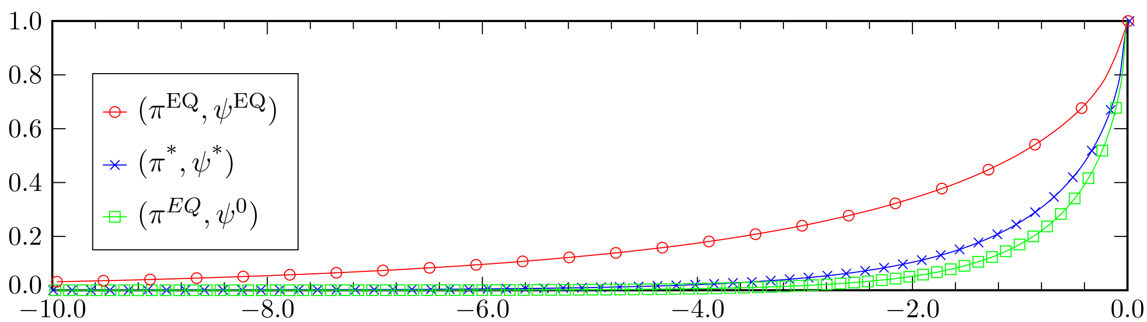
Under any linear policy, the evolution of the relative uncertainty σ_t/σ_0 over time is deterministic and depends only on the parameter ρ_0. This is because of the fact that $\sigma_t/\sigma_0 = \rho_t/\rho_0$ and the results in Section 4.1. In Figure 5, the evolution of the relative uncertainty of the PBE policy is illustrated, for different values of ρ_0, as compared to the equipartitioning and minimum revelation policies. In the low relative volume regime, the relative uncertainty of the PBE policy evolves similarly to



(a) The low relative volume regime, $\rho_0 = 1$.



(b) The moderate relative volume regime, $\rho_0 = 10$.



(c) The high relative volume regime, $\rho_0 = 100$.

Figure 4: The cumulative distribution of trader's normalized profit for the time horizon $T = 20$.

that of the equipartitioning policy. In the high relative volume regime, very little information is revealed until close to the end of trading period under the PBE policy. Indeed, the relative uncertainty between the equilibrium and the minimum revelation policies are indistinguishable on the scale of Figure 5, when $\rho_0 = 10$ or $\rho_0 = 100$. These observations are consistent with our results from Section 5.3.

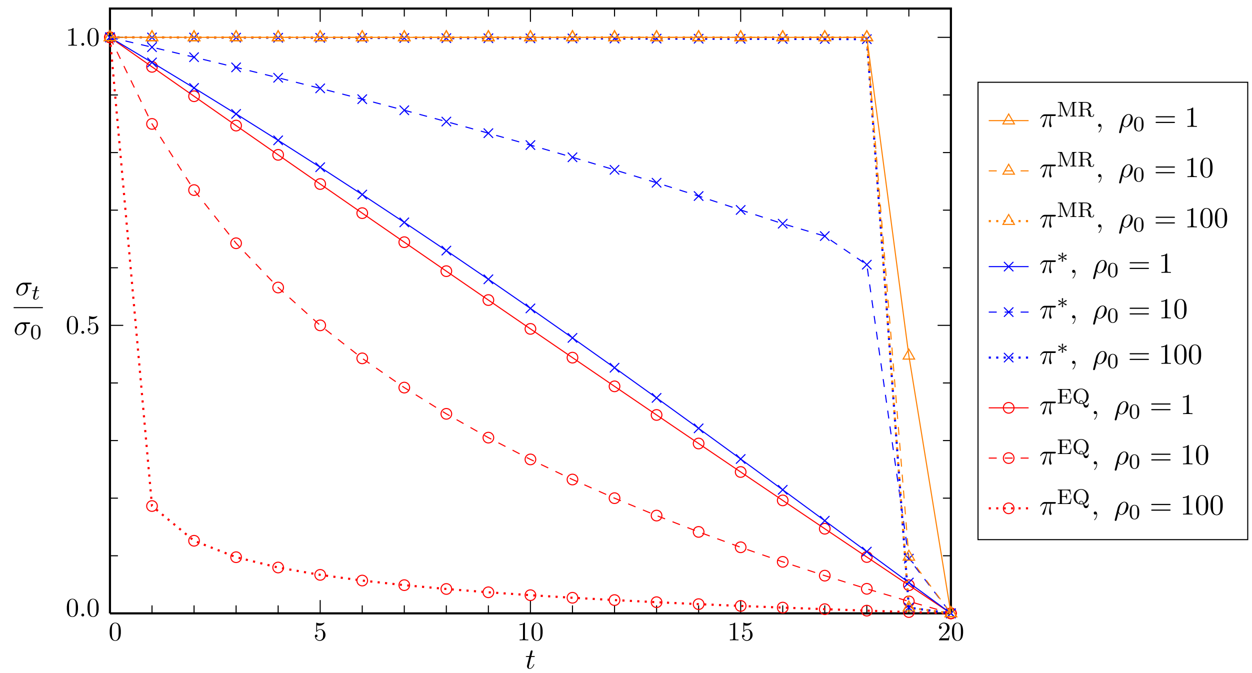


Figure 5: The evolution of relative uncertainty of the trader's position for the time horizon $T = 20$.

5.5. Adaptive Trading

One important feature of the linear-Gaussian PBE policy is that it is adaptive in the sense that the trades executed depend on the exogenous, stochastic price fluctuations. Note that these price fluctuations, which can be viewed as generated by noise traders, serve to disguise the activity of our primary trader, and that the trader does not benefit from further deception via injection of randomness beyond that which naturally arises from adapting to price fluctuations, as such a move would represent deviation from equilibrium.

The adaptive nature of the linear-Gaussian PBE policy distinguishes it from policies developed in most of the optimal execution literature. For example, the baseline equipartitioning policy of Bertsimas and Lo (1998) specifies a deterministic sequence of trades. Static policies have also been derived under more complicated models (e.g., Almgren and Chriss, 2000; Huberman and Stanzl, 2005; Obizhaeva and Wang, 2005; Alfonsi et al., 2007b). However, this behavior is in contrast to what is observed amongst institutional traders and trading algorithms that are implemented by practitioners. One justification for adaptive, price-responsive trading strategies is risk aversion. It has been observed that optimal policies for certain risk averse objectives require dynamic trading

(Hora, 2006; Almgren and Lorenz, 2006). Our model provides another justification: in the presence of asymmetric information and a strategic adversary, a trader should seek to exploit price fluctuations so as disguise trading activity.

In order to understand the behavior of linear policies, it is helpful to decompose them into deterministic and stochastic components. Suppose that (π, ψ) are a pair of linear policies, and that $y_0 = \mu_0 = 0$. Given Definition 3 and Theorem 1, it is easy to see that, for each $1 \leq t \leq T$, there exist vectors $\alpha_{\epsilon,t}, \beta_{\epsilon,t}, \gamma_{\epsilon,t} \in \mathbb{R}^t$ and scalars $\alpha_{x_0,t}, \beta_{x_0,t}, \gamma_{x_0,t} \in \mathbb{R}$, each of which depend on the parameters $\{\sigma_0, \lambda, \sigma_\epsilon\}$ only through the ρ_0, such that

$$x_t = \alpha_{x_0,t} x_0 + \frac{1}{\lambda} \alpha_{\epsilon,t}^\top \epsilon^t, \quad y_t = \beta_{x_0,t} x_0 + \frac{1}{\lambda} \beta_{\epsilon,t}^\top \epsilon^t, \quad \mu_t = \gamma_{x_0,t} x_0 + \frac{1}{\lambda} \gamma_{\epsilon,t}^\top \epsilon^t. \tag{18}$$

Here, $\epsilon^t = (\epsilon_1, \ldots, \epsilon_t)$ is the vector of exogenous disturbances up to time t. The first terms in (18) represent deterministic components of the policy and the second terms represent zero-mean stochastic components that depend on market price fluctuations. For the equipartitioning and minimum revelation policies, the stochastic components are zero. On the other hand, the Gaussian PBE policy does have non-zero stochastic components.

Figure 6 shows the deterministic component of the linear-Gaussian PBE versus those of the equipartitioning and minimum revelation policies. As the relative volume vanishes (i.e., $\rho_0 \to 0$), the trader ignores the presence of the arbitrageur and the PBE policy approaches the equipartitioning policy. At the other extreme, as the relative volume grows (i.e., $\rho_0 \to \infty$), in equilibrium the trader seeks to conceal his activity as much as possible, and hence the PBE policy approaches the minimum revelation policy.

Figure 7 illustrates sample paths of the trader's position under the linear-Gaussian PBE policy. Along each path, the trader deviates from the deterministic schedule based on the random fluctuations of the market. It is through this response to random fluctuations that the trader optimizes his signaling and thus his influence on the arbitrageur's beliefs. In general, if the arbitrageur's estimate of the trader's position becomes more accurate, the trader accelerates his selling to avoid front-running. On the other hand, if the arbitrageur is misled as to the trader's position, the trader delays his selling relative to the deterministic schedule to minimize his signal.

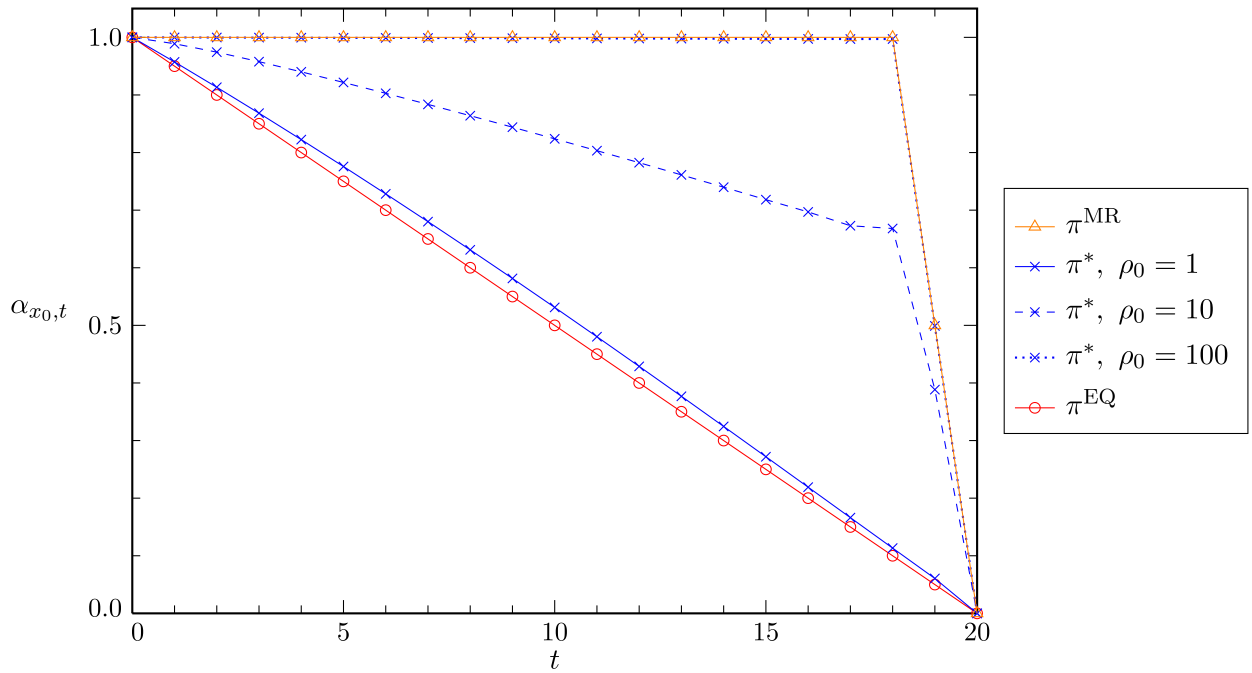


Figure 6: The deterministic components of trading strategies for the time horizon $T = 20$.

6. Extensions

In this section, we revisit some of the assumptions in the problem formulation of Section 2. At a high level, the main feature of our model that enables tractability is that, in equilibrium, each agent solves a linear-quadratic Gaussian control problem. This requires that the evolution of the model over time be described by a linear system and that the objectives of the trader and arbitrageur be quadratic functions that decompose additively over time. As we shall see shortly, there are a number of extensions of the model one may consider, incorporating important phenomena such as risk aversion and transient price impact, that maintain this structure. Such extensions remain tractable and can be addressed using straightforward adaptations of the techniques we have developed.

6.1. Time Horizon

Our model assumes that the trader begins his liquidation at time 1 and completes it by time T, and that this time interval is common knowledge. In some instances, public knowledge of the beginning and end of the liquidation interval might be reasonable since, for example, this interval will often correspond to a single trading day. More generally, however, it may be desirable to impose

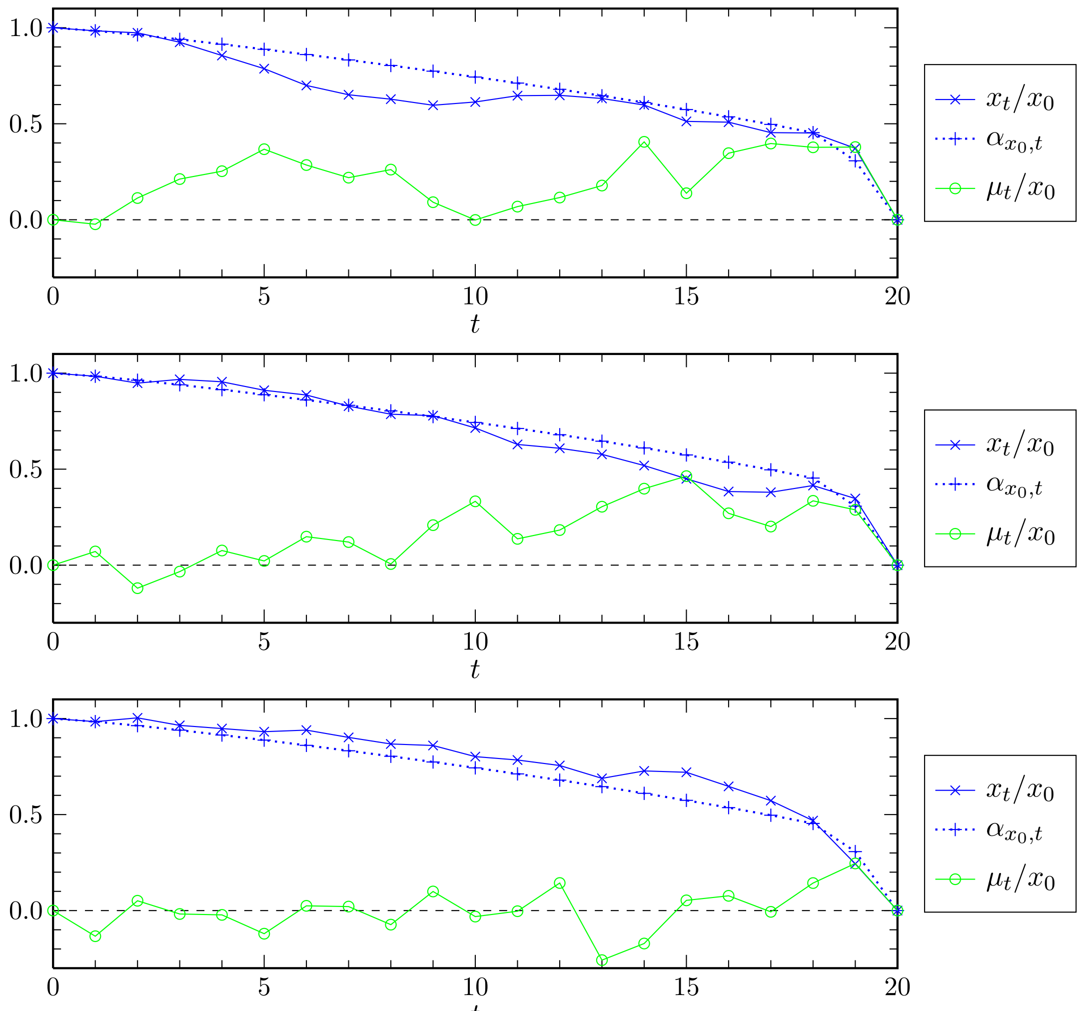


Figure 7: Sample paths of the evolution of the trader's actual and expected positions, and the arbitrageur's mean belief, when $T = 20$, $x_0 = \sigma_0 = 10^5$, $\mu_0 = y_0 = 0$, $\sigma_\epsilon = 0.125$, $\lambda = 10^{-5}$.

uncertainty on the part of arbitrageur as to the beginning and end of the liquidation. Unfortunately, it is not clear how to allow for this in a tractable fashion in our current framework.

The model further assumes that the arbitrageur must liquidate his position by time $T + 1$. Then, the value function of the arbitrageur at time T with position y_T, is given by $V_T^*(y_T) = -\lambda y_T^2$. This was used in (5)–(6) to determine the value functions U_{T-1}^* and V_{T-1}^*, which form the base case of the backward induction. This assumption can easily be relaxed. For example, suppose that the arbitrageur has T_a additional trading periods. It is easy to see that, after time T, the arbitrageur will optimally equipartition over the remaining T_a periods. Therefore the value of a position y_T at time T will take the form $V_T^*(y_T) = -\lambda \frac{T_a+1}{2T_a} y_T^2$, following the analysis in Bertsimas and Lo (1998).

So long as V_T^* is a quadratic function, our discussion in Sections 3 and 4 carries through, with a different choice of terminal value functions.

6.2. Risk Aversion

Our model assumes that both the trader and arbitrageur are risk-neutral. One way to account for risk aversion is to follow the approach suggested by Hora (2006). In particular, we could assume that, for example, the trader seeks to optimize the objective function

$$\mathsf{E}\left[\sum_{\tau=0}^{T-1}\left\{\Delta p_{\tau+1}x_\tau - \frac{\eta}{2}\big(\Delta p_{\tau+1}x_\tau - \mathsf{E}[\Delta p_{\tau+1}x_\tau \mid x_\tau, y_\tau, \phi_\tau]\big)^2 - \zeta x_\tau^2\right\} \,\middle|\, x_0, y_0, \phi_0\right],$$

The second term in the sum penalizes for variance in revenue in each time period, with $\eta \geq 0$ capturing the degree of risk aversion. This final term represents a per stage holding cost, with the parameter $\zeta \geq 0$ expressing the degree to which the trader would prefer to execute sooner rather than later. The risk neutral case previously considered corresponds to the choice of $\eta = \zeta = 0$. For any nonnegative parameter choices, the objective remains a time separable positive definite quadratic function. Hence, the methods of Sections 3 and 4 can be suitably adapted.

6.3. Price Impact & Price Dynamics

Our model assumed permanent and linear price impact. Empirically, it has been observed that transient price impact is a significant component of price dynamics, and it is important to account for this in the design of execution strategies.

Also, our model assumes a price impact coefficient that is constant. Empirically, the non-stationarity of price impact may be a significant phenomena, varying, for example, according to the time-of-day. Further, one theoretical justification for a permanent, linear price impact is the model of Kyle (1985). For that model, however, the price sensitivity is time dependent.

More generally, our analysis applies when there is some collection of state variables (for example, $\{x_t, y_t, \mu_t\}$) that evolve as a linear dynamical system with Gaussian disturbances, and where changes in price are linear in the state variables. In order to incorporate transient price impact and non-stationarity, assume that prices evolve according to

$$p_t = p_0 + \underbrace{\sum_{\tau=1}^{t} \lambda_\tau (u_\tau + v_\tau + z_\tau)}_{\text{permanent price impact}} + \underbrace{\sum_{\tau=1}^{t} \alpha^{t-\tau} \gamma_\tau (u_\tau + v_\tau + z_\tau)}_{\text{transient price impact}}. \tag{19}$$

Here, u_τ and v_τ are the trades of the trader and arbitrageur, respectively, as time τ. In place of the exogenous noise term in the original price dynamics (1), z_τ is an IID $N(0, \sigma_z^2)$ random variable representing the quantity of noise trades at time τ. The second term in (19) captures a permanent, linear price impact that is non-stationary with sensitivity $\lambda_\tau \geq 0$ at time τ. The final term represents a transient, linear price impact that is non-stationary with sensitivity $\gamma_\tau \geq 0$ at time τ and recovery rate $\alpha \in [0, 1)$.

These price dynamics can be rewritten as

$$p_t = p_{t-1} + (\lambda_t + \gamma_t)(u_t + v_t + z_t) - (1 - \alpha) s_{t-1},$$

where s_t is defined to be geometrically weighted total order flow

$$s_t \triangleq \sum_{\tau=0}^{t} \alpha^{t-\tau} \gamma_\tau (u_\tau + v_\tau + z_\tau) = \alpha s_{t-1} + \gamma_t (u_t + v_t + z_t).$$

Now, suppose that the trader's decision u_t is a linear function of $\{x_{t-1}, y_{t-1}, \mu_{t-1}, s_{t-1}\}$, and the arbitrageur's decision v_t is a linear function of $\{y_{t-1}, \mu_{t-1}, s_{t-1}\}$. Then, it will be the case that $\{x_t, y_t, \mu_t, s_t\}$ evolve as a linear dynamical system, and that the price changes are linear in these state variables. Therefore, the analysis in Sections 3 and 4 can be suitably modified and repeated, with an augmented state space. Note that, since s_t is a function of only of the *total* quantities traded at times up to t, it is reasonable to assume that this is public knowledge known to both the trader and arbitrageur.

Other aspects of more complicated price dynamics can also be incorporated via such state augmentation. For example, one may consider linear factor models or other otherwise add exogenous explanatory variables to the evolution of prices, so long as the dependencies are linear. Similarly, models that incorporate drift in the price process, such as short term momentum or mean reversion, can be considered.

6.4. Parameterized Policies

Beyond solving specific classes of models, results from the optimal execution literature offer useful guidance on how to structure parameterized execution policies that can be effective even if modeling assumptions are not entirely valid. In this vein, concepts we have developed can enhance parameterized policies that one might design based on prior literature.

For example, consider designing an execution system which begins the trading day with a position that must be liquidated by the end of that trading day. A number of models previously considered in the literature result in deterministic linear policies (see, e.g., Bertsimas and Lo, 1998; Obizhaeva and Wang, 2005; Alfonsi et al., 2007a). In particular, for each tth time period during the course of the day, there is a parameter a_t that indicates the fraction of the position to sell during that time period. These parameters $a_0, \ldots, a_{T-1}$ depend on asset-specific characteristics such as volatility and market impact model parameters.

Modeling assumptions often do not match reality. As such, it is useful to add flexibility by parametrizing the execution policy. For example, we might employ a policy that sells a fraction $\theta_t a_t$ of the position during each tth time period, where $\theta_0, \ldots, \theta_{T-1}$ are asset-independent parameters. Then, these parameters can be tuned based on experience from trading all assets. It is important that the number of parameters does not scale with the number of assets, because we would then be unlikely to have a sufficient amount of data to tune parameters. In this regard, the way $a_0, \ldots, a_{T-1}$ capture variations across assets is critical to the design of an effective parametrization.

Our work motivates a generalized class of parameterized policies that adapt trades as price movements are observed. Our model is optimized by an execution strategy with three sequences of coefficients: $\{a_{x,t}, a_{y,t}, a_{\mu,t} \mid t = 0, \ldots, T-1\}$. By simulating arbitrageur activity over the course of the day and applying these coefficients appropriately, we produce a sequence of trades that adapt to price fluctuations. Similarly with the case of a deterministic policy, we can introduce parameters $\{\theta_{x,t}, \theta_{y,t}, \theta_{\mu,t} \mid t = 0, \ldots, T-1\}$ that scale the policy coefficients, and tune these parameters based on experience. Once again, these parameters are asset-independent while the coefficients $\{a_{x,t}, a_{y,t}, a_{\mu,t} \mid t = 0, \ldots, T-1\}$ capture dependence of the policy on asset-specific characteristics such as volatility and market impact model parameters.

7. Conclusion

Our model captures strategic interactions between a trader aiming to liquidate a position and an arbitrageur trying to detect and profit from the trader's activity. The algorithm we have developed computes Gaussian perfect Bayesian equilibrium behavior. It is interesting that the resulting trader policy takes on such a simple form: the number of shares to liquidate at time t is linear in the trader's position x_{t-1}, the arbitrageur's position y_{t-1} and the arbitrageur's estimate μ_{t-1} of x_{t-1}. The coefficients of the policy depend only on the relative volume parameter ρ_0, which quantifies the magnitude of the trader's position relative to the typical market activity, and the time horizon T. This policy offers useful guidance beyond what has been derived in models that do not account for arbitrageur behavior. In the absence of an arbitrageur, it is optimal to trade equal amounts over each time period, which corresponds to a policy that is linear in x_{t-1}. The difference in the PBE policy stems from its accounting of the arbitrageur's inference process. In particular, the policy reduces information revealed to the arbitrageur by delaying trades and takes advantage of situations where the arbitrageur has been misled by unusual market activity.

Our model represents a starting point for the study of game theoretic behavior in trade execution. It has an admittedly simple structure, and this allows for a tractable analysis that highlights the importance of information signaling. There are a number of extensions to this model that are possible, however, and that warrant further discussion:

1. **(Flexible Time Horizon)** We assume finite time horizons T and $T+1$ for the trader and arbitrageur, respectively. The choice of time horizon has an impact on the resulting equilibrium policies, and there are clearly end-of-horizon effects in the policies computed in Section 5. To some extent it seems artificial to impose a fixed time horizon as an exogenous restriction on behavior. Fixed horizon models preclude the trader from delaying liquidation beyond the horizon even if this can yield significant benefits, for example. A better model would be to consider an infinite horizon game, where risk aversion provides the motivation for liquidating a position sooner rather than later.

2. **(Uncertain Trader)** In our model, we assume that the arbitrageur is uncertain of the trader's position, but that the trader knows everything. A more realistic model would allow for uncertainty on the part of the trader as well, and would allow for the arbitrageur to mislead the trader.

3. **(Multi-player Games)** Our model restricts to a single trader and arbitrageur. A natural extension would be to consider multiple traders and arbitrageurs that are uncertain about each others' positions and must compete in the marketplace as they unwind. Such a generalized model could be useful for analysis of important liquidity issues such as those arising from the credit crunch of 2007.

Also of interest are the potential empirical implications of the model. If we make the assumption that the trade execution horizon is a single day, the observations in Section 5 suggest particular patterns for intraday volume. For example, if ρ_0 is large, the volume traded should be much higher near the end of the day then at other times. Similarly, the structure of the equilibrium trading policies for the trader and arbitrageur will generate specific, time-varying auto-correlation in the increments of the price process. Formulating tests of such empirical predictions in any interesting area for future research.

Finally, beyond the immediate context of our model, there are many directions worth exploring. One important avenue is to factor data beyond price into the execution strategy. For example, volume data may play a significant role in the arbitrageur's inference, in which case it should also influence execution decisions. Limit order book data may also be relevant. Developing tractable models that account for such data remains a challenge. One initiative to incorporate limit order book data into the decision process is presented by Nevmyvaka et al. (2006).

Acknowledgments

The authors are grateful to the anonymous reviewers for helpful comments and corrections. The first author wishes to thank Mark Broadie and Gabriel Weintraub for useful discussions. The second author wishes to thank the Samsung Scholarship Foundation for financial support. This research was conducted while the third author was visiting the Faculty of Commerce and Accountancy at Chulalongkorn University and supported by the Chin Sophonpanich Foundation Fund.

References

A. Alfonsi, A. Schied, and A. Schulz. Constrained portfolio liquidation in a limit order book model. Working paper, 2007a.

A. Alfonsi, A. Schied, and A. Schulz. Optimal execution strategies in limit order books with general shape functions. Working paper, 2007b.

R. Almgren. Optimal execution with nonlinear impact functions and trading-enhanced risk. *Applied Mathematical Finance*, 10:1–18, 2003.

R. Almgren and N. Chriss. Optimal control of portfolio transactions. *Journal of Risk*, 3:5–39, 2000.

R. Almgren and J. Lorenz. Adaptive arrival price. Working paper, April 2006.

K. Back. Insider trading in continuous time. *Review of Financial Studies*, 5(3):387–409, 1992.

K. Back and S. Baruch. Information in securities markets: Kyle meets Glosten and Milgrom. *Econometrica*, 72(2):433–465, March 2004.

D. Bertsimas and A. W. Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1:1–50, 1998.

M. K. Brunnermeier and L. H. Pedersen. Predatory trading. *Journal of Finance*, 60(4):1825–1863, 2005.

H. H. Cao, M. D. Evans, and R. K. Lyons. Inventory information. *Journal of Business*, 79(1):325–363, January 2006.

B. I. Carlin, M. S. Lobo, and S. Viswanathan. Episodic liquidity crises: Cooperative and predatory trading. *Journal of Finance*, 62(5):2235–2274, 2007.

P. M. DeMarzo and B. Urošević. Ownership dynamics and asset pricing with a large shareholder. *Journal of Political Economy*, 114(4):774–815, August 2006.

R. Dubil. Optimal liquidation of large security holdings in thin markets. In Y. Shachmurove, editor, *Research in International Entrepreneurial Finance and Business Ventures at the Turn of the Third Millennium*. Academy of Entrepreneurial Finance, 2002.

C. Duhigg. Stock traders find speed pays, in milliseconds. *The New York Times*, page A1, July 24, 2009.

R. Engle and R. Ferstenberg. Execution risk. Working Paper 12165, NBER, April 2006.

F. D. Foster and S. Viswanathan. Strategic trading with asymmetrically informed traders and long-lived information. *Journal of Financial and Quantitative Analysis*, 29(4):499–518, December 1994.

F. D. Foster and S. Viswanathan. Strategic trading when agents forecast the forecasts of others. *Journal of Finance*, 51(4):1437–1478, September 1996.

D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, Cambridge, MA, 1991.

M. Guo and A. S. Kyle. Dynamic strategic informed trading with risk-averse market makers. Working paper, 2005.

C. W. Holden and A. Subrahmanyam. Long-lived private information and imperfect competition. *Journal of Finance*, 47(1):247–270, March 1992.

C. W. Holden and A. Subrahmanyam. Risk aversion, imperfect competition, and long-lived information. *Economics Letters*, 44:181–190, 1994.

M. Hora. Tactical liquidity trading and intraday volume. Working paper, available at `http://ssrn.com/paper=931667`, September 2006.

G. Huberman and W. Stanzl. Optimal liquidity trading. *Review of Finance*, 9:165–200, 2005.

A. S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335, November 1985.

J. Lorenz. *Optimal Trading Algorithms: Portfolio Transactions, Multiperiod Portfolio Selection, and Competitive Online Search*. PhD thesis, ETH Zürich, 2008.

A. Madhavan and M. Cheng. In search of liquidity: Block trades in the upstairs and downstairs markets. *The Review of Financial Studies*, 10(1):175–203, 1997.

Y. Nevmyvaka, Y. Feng, and M. Kearns. Reinforcement learning for optimized trade execution. In *Proceedings of the 23rd International Conference on Machine Learning*. Association for Computing Machinery, 2006.

A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. Working paper, 2005.

M. Oehmke. Liquidating illiquid collateral. Working paper, 2010.

A. Schied and T. Schönenborn. Optimal portfolio liquidation for CARA investors. Working paper, 2007.

T. Schönenborn and A. Schied. Liquidation in the face of adversity: stealth vs. sunshine trading, predatory trading vs. liquidity provision. Working paper, 2007.

A. Subramanian and R. A. Jarrow. The liquidity discount. *Mathematical Finance*, 11(4):447–474, 2001.

D. Vayanos. Strategic trading and welfare in a dynamic market. *Review of Economic Studies*, 66(2):219–254, 1999.

D. Vayanos. Strategic trading in a dynamic noisy market. *Journal of Finance*, 56:131–171, 2001.

A. Proofs

Theorem 1. *If the belief distribution ϕ_{t-1} at time is Gaussian, and the arbitrageur assumes that the trader's policy $\hat{\pi}_t$ is linear with $\hat{\pi}_t(x_{t-1}, y_{t-1}, \phi_{t-1}) = \hat{a}_{x,t}^{\rho_{t-1}} x_{t-1} + \hat{a}_{y,t}^{\rho_{t-1}} y_{t-1} + \hat{a}_{\mu,t}^{\rho_{t-1}} \mu_{t-1}$, then the belief distribution ϕ_t is also Gaussian. The mean μ_t is a linear function of y_{t-1}, μ_{t-1}, and the observed price change Δp_t, with coefficients that are deterministic functions of the scaled variance ρ_{t-1}. The scaled variance ρ_t evolves according to*

$$\rho_t^2 = \left(1 + \hat{a}_{x,t}^{\rho_{t-1}}\right)^2 \left(\frac{1}{\rho_{t-1}^2} + (\hat{a}_{x,t}^{\rho_{t-1}})^2\right)^{-1}.$$

In particular, ρ_t is a deterministic function of ρ_{t-1}.

Proof. Set $\{K_{t-1}, h_{t-1}\}$ to be the information form parameters for the Gaussian distribution ϕ_{t-1}, so that

$$K_{t-1} \triangleq 1/\sigma_{t-1}^2, \qquad \text{and} \qquad h_{t-1} \triangleq \mu_{t-1}/\sigma_{t-1}^2.$$

Define ϕ_{t-1}^+ to be the distribution of x_{t-1} conditioned on all information seen by the arbitrageur at times up to and including t. That is,

$$\phi_{t-1}^+(S) \triangleq \Pr\left(x_{t-1} \in S \mid \phi_{t-1}, y_{t-1}, \lambda(\hat{\pi}_t(x_{t-1}, y_{t-1}, \phi_{t-1}) + v_t) + \epsilon_t = \Delta p_t\right),$$

where Δp_t is the price change observed at time t. By Bayes' rule, this distribution has density

$$\begin{aligned}
\phi_{t-1}^+(dx) &\propto \phi_{t-1}(dx)\, \exp\left(-\frac{\left(\Delta p_t - \lambda(\pi_t(x, y_{t-1}, \phi_{t-1}) + \psi_t(y_{t-1}, \phi_{t-1}))\right)^2}{2\sigma_\epsilon^2}\right) \\
&\propto \exp\Bigg(-\tfrac{1}{2}\left(K_{t-1} + \frac{\lambda^2 (\hat{a}_{x,t}^{\rho_{t-1}})^2}{\sigma_\epsilon^2}\right) x^2 \\
&\qquad\qquad + \left(h_{t-1} + \frac{\lambda(\Delta p_t - \lambda(\hat{a}_{y,t}^{\rho_{t-1}} y_{t-1} + \hat{a}_{\mu,t}^{\rho_{t-1}} \mu_{t-1} + \psi_t))\hat{a}_{x,t}}{\sigma_\epsilon^2}\right) x\Bigg)\, dx.
\end{aligned}$$

Thus, ϕ_{t-1}^+ is a Gaussian distribution, with variance

$$\left(K_{t-1} + \frac{\lambda^2 (\hat{a}_{x,t}^{\rho_{t-1}})^2}{\sigma_\epsilon^2}\right)^{-1},$$

and mean

$$\left(K_{t-1} + \frac{\lambda^2 (\hat{a}_{x,t}^{\rho_{t-1}})^2}{\sigma_\epsilon^2}\right)^{-1} \left(h_{t-1} + \frac{\lambda(\Delta p_t - \lambda(\hat{a}_{y,t}^{\rho_{t-1}} y_{t-1} + \hat{a}_{\mu,t}^{\rho_{t-1}} \mu_{t-1} + \psi_t))\hat{a}_{x,t}}{\sigma_\epsilon^2}\right).$$

Now, note that

$$x_t = x_{t-1} + \hat{\pi}_t(x_{t-1}, y_{t-1}, \phi_{t-1}) = (1 + \hat{a}_{x,t}^{\rho_{t-1}}) x_{t-1} + \hat{a}_{y,t}^{\rho_{t-1}} y_{t-1} + \hat{a}_{\mu,t}^{\rho_{t-1}} \mu_{t-1}.$$

Then, ϕ_t is also a Gaussian distribution, with variance

$$\sigma_t^2 = (1 + \hat{a}_{x,t}^{\rho_{t-1}})^2 \left(K_{t-1} + \frac{\lambda^2 (\hat{a}_{x,t}^{\rho_{t-1}})^2}{\sigma_\epsilon^2}\right)^{-1} = (1 + \hat{a}_{x,t})^2 \left(\frac{1}{\sigma_{t-1}^2} + \frac{\lambda^2 (\hat{a}_{x,t}^{\rho_{t-1}})^2}{\sigma_\epsilon^2}\right)^{-1}, \tag{20}$$

and mean

$$\begin{aligned}\mu_t &= \hat{a}_{y,t}^{\rho_{t-1}} y_{t-1} + \hat{a}_{\mu,t}^{\rho_{t-1}} \mu_{t-1} \\ &\quad + (1 + \hat{a}_{x,t}^{\rho_{t-1}}) \frac{\mu_{t-1}/\rho_{t-1}^2 + (\Delta p_t/\lambda - \hat{a}_{y,t}^{\rho_{t-1}} y_{t-1} - \hat{a}_{\mu,t}^{\rho_{t-1}} \mu_{t-1} - \psi_t)\hat{a}_{x,t}}{1/\rho_{t-1}^2 + (\hat{a}_{x,t}^{\rho_{t-1}})^2}.\end{aligned} \tag{21}$$

The conclusions of the theorem immediately follow. ■

In order to prove Theorems 2–4, it is necessary to explicitly evaluate the operator $F_{u_t}^{(\psi_t,\pi_t)}$ applied to quadratic functions of $\{x_t, y_t, \mu_t\}$ and the operator $G_{v_t}^{\pi_t}$ applied to quadratic functions of $\{y_t, \mu_t\}$. The following lemma is helpful for this purpose, as it provides expressions for the expectation of μ_t and μ_t^2 under various distributions.

Lemma 1. *Assume that the the policies ψ_t and π_t are linear with*

$$\pi_t(x_{t-1}, y_{t-1}, \phi_{t-1}) = a_{x,t}^{\rho_{t-1}} x_{t-1} + a_{y,t}^{\rho_{t-1}} y_{t-1} + a_{\mu,t}^{\rho_{t-1}} \mu_{t-1},$$

$$\psi_t(y_{t-1}, \phi_{t-1}) = b_{y,t}^{\rho_{t-1}} y_{t-1} + b_{\mu,t}^{\rho_{t-1}} \mu_{t-1}.$$

Define

$$\gamma_t^{\rho_{t-1}} \triangleq \frac{1 + a_{x,t}^{\rho_{t-1}}}{1/\rho_{t-1}^2 + (a_{x,t}^{\rho_{t-1}})^2}.$$

Then,

$$\mathsf{E}^{(\psi_t,\pi_t)}_{u_t}\left[\mu_t \mid x_{t-1}, y_{t-1}, \phi_{t-1}\right] = a^{\rho_{t-1}}_{y,t} y_{t-1} + a^{\rho_{t-1}}_{\mu,t}\mu_{t-1} + \gamma^{\rho_{t-1}}_t \mu_{t-1}/\rho^2_{t-1} + \gamma^{\rho_{t-1}}_t a^{\rho_{t-1}}_{x,t}\left(u_t - a^{\rho_{t-1}}_{y,t} y_{t-1} - a^{\rho_{t-1}}_{\mu,t}\mu_{t-1}\right), \tag{22a}$$

$$\mathrm{Var}^{(\psi_t,\pi_t)}_{u_t}\left[\mu_t \mid x_{t-1}, y_{t-1}, \phi_{t-1}\right] = \left(\gamma^{\rho_{t-1}}_t a^{\rho_{t-1}}_{x,t}\sigma_\epsilon/\lambda\right)^2, \tag{22b}$$

$$\mathsf{E}^{(\psi_t,\pi_t)}_{u_t}\left[\mu_t^2 \,\middle|\, x_{t-1}, y_{t-1}, \phi_{t-1}\right] = \mathrm{Var}^{(\psi_t,\pi_t)}_{u_t}\left[\mu_t \mid x_{t-1}, y_{t-1}, \phi_{t-1}\right] + \left(\mathsf{E}^{(\psi_t,\pi_t)}_{u_t}\left[\mu_t \mid x_{t-1}, y_{t-1}, \phi_{t-1}\right]\right)^2, \tag{22c}$$

$$\mathsf{E}^{\pi_t}_{v_t}\left[\mu_t \mid y_{t-1}, \phi_{t-1}\right] = a^{\rho_{t-1}}_{y,t} y_{t-1} + (1 + a^{\rho_{t-1}}_{x,t} + a^{\rho_{t-1}}_{\mu,t})\mu_{t-1}, \tag{22d}$$

$$\mathrm{Var}^{\pi_t}_{v_t}\left[\mu_t \mid y_{t-1}, \phi_{t-1}\right] = \left(\gamma^{\rho_{t-1}}_t a^{\rho_{t-1}}_{x,t}\sigma_\epsilon/\lambda\right)^2 \left(1 + \left(a^{\rho_{t-1}}_{x,t}\right)^2 \rho^2_{t-1}\right), \tag{22e}$$

$$\mathsf{E}^{\pi_t}_{v_t}\left[\mu_t^2 \,\middle|\, y_{t-1}, \phi_{t-1}\right] = \mathrm{Var}^{\pi_t}_{v_t}\left[\mu_t \mid y_{t-1}, \phi_{t-1}\right] + \left(\mathsf{E}^{\pi_t}_{v_t}\left[\mu_t \mid y_{t-1}, \phi_{t-1}\right]\right)^2. \tag{22f}$$

Proof. The lemma follows directly from taking expectations of the mean update equation (21). ■

Theorem 2. *If U^*_t is TQD and V^*_t is AQD, and Step 3 of Algorithm 2 produces a linear pair (π^*_t, ψ^*_t), then U^*_{t-1} and V^*_{t-1}, defined by Step 4 of Algorithm 2 are TQD and AQD, respectively.*

Proof. Suppose that

$$V^*_t(y_t, \phi_t) = -\lambda\left(\tfrac{1}{2} d^{\rho_t}_{yy,t} y_t^2 + \tfrac{1}{2} d^{\rho_t}_{\mu\mu,t}\mu_t^2 + d^{\rho_t}_{y\mu,t} y_t \mu_t - \frac{\sigma^2_\epsilon}{\lambda^2} d^{\rho_t}_{0,t}\right),$$

$$\pi^*_t(x_{t-1}, y_{t-1}, \phi_{t-1}) = a^{\rho_{t-1}}_{x,t} x_{t-1} + a^{\rho_{t-1}}_{y,t} y_{t-1} + a^{\rho_{t-1}}_{\mu,t}\mu_{t-1},$$

$$\psi^*_t(y_{t-1}, \phi_{t-1}) = b^{\rho_{t-1}}_{y,t} y_{t-1} + b^{\rho_{t-1}}_{\mu,t}\mu_{t-1}.$$

If the trader uses the policy π_t^* and the arbitrageur uses the policy ψ_t^*, we have

$$u_t = a_{x,t}^{\rho_{t-1}} x_{t-1} + a_{y,t}^{\rho_{t-1}} y_{t-1} + a_{\mu,t}^{\rho_{t-1}} \mu_{t-1},$$
$$v_t = b_{y,t}^{\rho_{t-1}} y_{t-1} + b_{\mu,t}^{\rho_{t-1}} \mu_{t-1},$$
$$y_t = y_{t-1} + b_{y,t}^{\rho_{t-1}} y_{t-1} + b_{\mu,t}^{\rho_{t-1}} \mu_{t-1}.$$

Using these facts, Theorem 1, and (22d)–(22f) from Lemma 1, we can explicitly compute

$$\begin{aligned} V_{t-1}^*(y_{t-1}, \phi_{t-1}) &= \left(G_{\psi_t^*}^{\pi_t^*} V\right)(y_{t-1}, \phi_{t-1}) \\ &= \mathsf{E}_{\psi_t^*}^{\pi_t^*}\left[\lambda(u_t + v_t) y_{t-1} + V_t^*(y_t, \phi_t) \;\middle|\; y_{t-1}, \phi_{t-1}\right] \\ &= -\lambda\left(\tfrac{1}{2} d_{yy,t-1}^{\rho_{t-1}} y_t^2 + \tfrac{1}{2} d_{\mu\mu,t-1}^{\rho_{t-1}} \mu_t^2 + d_{y\mu,t-1}^{\rho_{t-1}} y_t \mu_t - \frac{\sigma_\epsilon^2}{\lambda^2} d_{0,t-1}^{\rho_{t-1}}\right), \end{aligned}$$

where

$$\begin{aligned} \rho_t^2 &= \left(1 + \hat{a}_{x,t}^{\rho_{t-1}}\right)^2 \left(\frac{1}{\rho_{t-1}^2} + (\hat{a}_{x,t}^{\rho_{t-1}})^2\right)^{-1}, \\ d_{yy,t-1}^{\rho_{t-1}} &= \left(d_{yy,t}^{\rho_t} - \frac{(d_{y\mu,t}^{\rho_t})^2}{d_{yy,t}^{\rho_t}}\right)(a_{y,t}^{\rho_t})^2 + 2\left(\frac{d_{y\mu,t}^{\rho_t}}{d_{yy,t}^{\rho_t}} - 1\right) a_{y,t}^{\rho_t} - \frac{1}{d_{yy,t}^{\rho_t}} + 2, \\ d_{y\mu,t-1}^{\rho_{t-1}} &= -a_{\mu,t}^{\rho_t} - a_{x,t}^{\rho_t} + \left(\frac{d_{y\mu,t}^{\rho_t}}{d_{yy,t}^{\rho_t}} + \left(d_{\mu\mu,t}^{\rho_t} - \frac{(d_{y\mu,t}^{\rho_t})^2}{d_{yy,t}^{\rho_t}}\right) a_{y,t}^{\rho_t}\right)(1 + a_{x,t}^{\rho_t} + a_{y,t}^{\rho_t}), \\ d_{\mu\mu,t-1}^{\rho_{t-1}} &= \left(d_{\mu\mu,t}^{\rho_t} - \frac{(d_{y\mu,t}^{\rho_t})^2}{d_{yy,t}^{\rho_t}}\right)(1 + a_{x,t}^{\rho_t} + a_{y,t}^{\rho_t})^2, \\ d_{0,t-1}^{\rho_{t-1}} &= d_{0,t}^{\rho_t} + \frac{d_{\mu\mu,t}^{\rho_t}}{2}\left(a_{x,t}^{\rho_t} \gamma_t^{\rho_{t-1}} \frac{\sigma_\epsilon}{\lambda}\right)^2 \left(1 + (\rho_{t-1} a_{x,t}^{\rho_t})^2\right). \end{aligned}$$

Therefore, V_{t-1}^* is AQD. Similarly, we can check that U_{t-1}^* is TQD. ■

Theorem 3. *Suppose that U_t^* and V_t^* and TQD/AQD value functions specified by (10)–(11), and (π_t^*, ψ_t^*) are linear policies specified by (7)–(8). Assume that, for all ρ_{t-1}, the policy coefficients* *satisfy the first order conditions*

$$0 = \left(\rho_t^2 c^{\rho_t}_{\mu\mu,t} + 2\rho_t c^{\rho_t}_{x\mu,t} + c^{\rho_t}_{xx,t}\right)\left(a^{\rho_{t-1}}_{x,t}\right)^3 + \left(3c^{\rho_t}_{xx,t} + 3\rho_t c^{\rho_t}_{x\mu,t} - 1\right)\left(a^{\rho_{t-1}}_{x,t}\right)^2 + \left(3c^{\rho_t}_{xx,t} + \rho_t c^{\rho_t}_{x\mu,t} - 2\right)a^{\rho_{t-1}}_{x,t} + c^{\rho_t}_{xx,t} - 1, \tag{23}$$

$$a^{\rho_{t-1}}_{y,t} = -\frac{\left(b^{\rho_{t-1}}_{y,t} + 1\right)\left(c^{\rho_t}_{xy,t} + \alpha_t c^{\rho_t}_{y\mu,t}\right)}{c^{\rho_t}_{xx,t} + (\alpha_t + 1)c^{\rho_t}_{x\mu,t} + \alpha_t c^{\rho_t}_{\mu\mu,t}}, \tag{24}$$

$$a^{\rho_{t-1}}_{\mu,t} = -\frac{a^{\rho_{t-1}}_{x,t} b^{\rho_{t-1}}_{\mu,t}\left(c^{\rho_t}_{xy,t} + \alpha_t c^{\rho_t}_{y\mu,t}\right) + \alpha_t\left(c^{\rho_t}_{x\mu,t} + \alpha_t c^{\rho_t}_{\mu\mu,t}\right)/\rho_{t-1}^2}{a^{\rho_{t-1}}_{x,t}\left(c^{\rho_t}_{xx,t} + (\alpha_t + 1)c^{\rho_t}_{x\mu,t} + \alpha_t c^{\rho_t}_{\mu\mu,t}\right)}, \tag{25}$$

$$b^{\rho_{t-1}}_{y,t} = \frac{1 - d^{\rho_t}_{y\mu,t} a^{\rho_{t-1}}_{y,t}}{d^{\rho_t}_{yy,t}} - 1, \qquad b^{\rho_{t-1}}_{\mu,t} = -\frac{\left(1 + a^{\rho_{t-1}}_{\mu,t} + a^{\rho_{t-1}}_{x,t}\right) d^{\rho_t}_{y\mu,t}}{d^{\rho_t}_{yy,t}}, \tag{26}$$

and the second order conditions

$$c^{\rho_t}_{xx,t} + (\alpha_t + 1)c^{\rho_t}_{x\mu,t} + \alpha_t c^{\rho_t}_{\mu\mu,t} > 0, \qquad d^{\rho_t}_{yy,t} > 0, \tag{27}$$

where the quantities α_t *and* ρ_t *satisfy*

$$\alpha_t = \frac{a^{\rho_{t-1}}_{x,t}\left(1 + a^{\rho_{t-1}}_{x,t}\right)}{1/\rho_{t-1}^2 + \left(a^{\rho_{t-1}}_{x,t}\right)^2}, \qquad \rho_t^2 = \left(1 + a^{\rho_{t-1}}_{x,t}\right)^2 \left(\frac{1}{\rho_{t-1}^2} + \left(a^{\rho_{t-1}}_{x,t}\right)^2\right)^{-1}. \tag{28}$$

Then, (π_t^*, ψ_t^*) *satisfy the single-stage equilibrium conditions*

$$\pi_t^*(x_{t-1}, y_{t-1}, \phi_{t-1}) \in \operatorname*{argmax}_{u_t} \left(F^{(\psi_t^*, \pi_t^*)}_{u_t} U_t^*\right)(x_{t-1}, y_{t-1}, \phi_{t-1}), \tag{29}$$

$$\psi_t^*(y_{t-1}, \phi_{t-1}) \in \operatorname*{argmax}_{v_t} \left(G^{\pi_t^*}_{v_t} V_t^*\right)(y_{t-1}, \phi_{t-1}), \tag{30}$$

for all x_{t-1}*,* y_{t-1}*, and Gaussian* ϕ_{t-1}*.*

Proof. As we will discuss in the proof of Theorem 4, the optimizing value u_t^* in (29) is a linear function of x_{t-1}, y_{t-1} and z_{t-1}, whose coefficients depend on $\{a^{\rho_{t-1}}_{x,t}, a^{\rho_{t-1}}_{y,t}, a^{\rho_{t-1}}_{\mu,t}, b^{\rho_{t-1}}_{y,t}, b^{\rho_{t-1}}_{\mu,t}\}$. By equating the coefficients of $\{x_{t-1}, y_{t-1}, z_{t-1}\}$ with $\{a^{\rho_{t-1}}_{x,t}, a^{\rho_{t-1}}_{y,t}, a^{\rho_{t-1}}_{\mu,t}\}$, respectively, we can obtain (23), (24) and 25. (26) can be derived by considering (30) in the same way. (27) corresponds to the second order conditions for the two maximization problems. ∎

Theorem 4. *If U_t is TQD, ψ_t is linear, and $\hat{\pi}_t$ is linear, then there exists a linear π_t such that*

$$\pi_t(x_{t-1}, y_{t-1}, \phi_{t-1}) \in \operatorname*{argmax}_{u_t} \left(F_{u_t}^{(\psi_t, \hat{\pi}_t)} U_t \right) (x_{t-1}, y_{t-1}, \phi_{t-1}),$$

for all x_{t-1}, y_{t-1}, and Gaussian ϕ_{t-1}, so long as the optimization problem is bounded. Similarly, if V_t is AQD and π_t is linear then there exists a linear ψ_t such that

$$\psi_t(y_{t-1}, \phi_{t-1}) \in \operatorname*{argmax}_{v_t} \left(G_{v_t}^{\pi_t} V_t \right) (y_{t-1}, \phi_{t-1}),$$

for all y_{t-1} and Gaussian ϕ_{t-1}, so long as the optimization problem is bounded.

Proof. Suppose that

$$\begin{aligned}
U_t(x_t, y_t, \phi_t) = -\lambda\Bigg(& \tfrac{1}{2} c_{xx,t}^{\rho_t} x_t^2 + \tfrac{1}{2} c_{yy,t}^{\rho_t} y_t^2 + \tfrac{1}{2} c_{\mu\mu,t}^{\rho_t} \mu_t^2 \\
& + c_{xy,t}^{\rho_t} x_t y_t + c_{x\mu,t}^{\rho_t} x_t \mu_t + c_{y\mu,t}^{\rho_t} y_t \mu_t - \frac{\sigma_\epsilon^2}{\lambda^2} c_{0,t}^{\rho_t} \Bigg), \\
\hat{\pi}_t(x_{t-1}, y_{t-1}, \phi_{t-1}) = {} & \hat{a}_{x,t}^{\rho_{t-1}} x_{t-1} + \hat{a}_{y,t}^{\rho_{t-1}} y_{t-1} + \hat{a}_{\mu,t}^{\rho_{t-1}} \mu_{t-1}, \\
\psi_t(y_{t-1}, \phi_{t-1}) = {} & b_{y,t}^{\rho_{t-1}} y_{t-1} + b_{\mu,t}^{\rho_{t-1}} \mu_{t-1}.
\end{aligned}$$

If the trader takes the action u_t, while the arbitrageur uses the policy ψ_t^* and assumes that the trader uses the policy $\hat{\pi}_t$, we have

$$\begin{aligned}
v_t &= b_{y,t}^{\rho_{t-1}} y_{t-1} + b_{\mu,t}^{\rho_{t-1}} \mu_{t-1}, \\
x_t &= x_{t-1} + u_t, \\
y_t &= y_{t-1} + b_{y,t}^{\rho_{t-1}} y_{t-1} + b_{\mu,t}^{\rho_{t-1}} \mu_{t-1}.
\end{aligned}$$

Using these facts, Theorem 1, and (22a)–(22c) from Lemma 1, we can explicitly compute

$$\left(F_{u_t}^{(\psi_t, \hat{\pi}_t)} U_t \right) (x_{t-1}, y_{t-1}, \phi_{t-1}) = \mathsf{E}_{u_t}^{(\psi_t, \hat{\pi}_t)} \left[\lambda(u_t + v_t) x_{t-1} + U_t(x_t, y_t, \phi_t) \;\middle|\; x_{t-1}, y_{t-1}, \phi_{t-1} \right].$$

It is easy to see that $\left(F_{u_t}^{(\psi_t, \hat{\pi}_t)} U_t \right) (x_{t-1}, y_{t-1}, \phi_{t-1})$ is quadratic in u_t. Moreover, the coefficient of u_t^2 is independent of $\{x_{t-1}, y_{t-1}, \mu_{t-1}\}$ while the coefficient of u_t is linear in $\{x_{t-1}, y_{t-1}, \mu_{t-1}\}$.

Therefore, the optimizing u_t^* is a linear function of $\{x_{t-1}, y_{t-1}, \mu_{t-1}\}$, whose coefficients can be computed by substitution and rearrangement of the resulting terms.

Similarly, suppose that

$$V_t(y_t, \phi_t) = -\lambda \left(\tfrac{1}{2} d^{\rho_t}_{yy,t} y_t^2 + \tfrac{1}{2} d^{\rho_t}_{\mu\mu,t} \mu_t^2 + d^{\rho_t}_{y\mu,t} y_t \mu_t - \frac{\sigma_\epsilon^2}{\lambda^2} d^{\rho_t}_{0,t} \right),$$

$$\pi_t(x_{t-1}, y_{t-1}, \phi_{t-1}) = a^{\rho_{t-1}}_{x,t} x_{t-1} + a^{\rho_{t-1}}_{y,t} y_{t-1} + a^{\rho_{t-1}}_{\mu,t} \mu_{t-1}.$$

If the arbitrageur takes the action v_t and assumes that the trader uses the policy π_t, we have

$$\begin{aligned} u_t &= a^{\rho_{t-1}}_{x,t} x_{t-1} + a^{\rho_{t-1}}_{y,t} y_{t-1} + a^{\rho_{t-1}}_{\mu,t} \mu_{t-1}, \\ y_t &= y_{t-1} + v_t. \end{aligned}$$

Using these facts, Theorem 1, and (22d)–(22f) from Lemma 1, we can explicitly compute

$$\left(G^{\pi_t}_{v_t} V_t\right)(y_{t-1}, \phi_{t-1}) = \mathsf{E}^{\pi_t}_{v_t} \left[\lambda(\pi_t + v_t) y_{t-1} + V_t(y_t, \phi_t) \mid y_{t-1}, \phi_{t-1} \right].$$

It is easily checked that $\left(G^{\pi_t}_{v_t} V_t\right)(y_{t-1}, \phi_{t-1})$ is quadratic in v_t. Moreover, the coefficient of v_t^2 is independent of $\{y_{t-1}, \mu_{t-1}\}$ while the coefficient of v_t is linear in $\{y_{t-1}, \mu_{t-1}\}$. Therefore, the optimizing v_t^* is a linear of $\{y_{t-1}, \mu_{t-1}\}$, whose coefficients can be computed by substitution and rearrangement of the resulting terms. ■