The Execution Game

Ciamac C. Moallemi
Columbia University
ciamac@gsb.columbia.edu

Beomsoo Park
Stanford University
beomsoo@stanford.edu

Benjamin Van Roy
Stanford University
bvr@stanford.edu

April 28, 2008

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. 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.

Optimal execution algorithms have been developed for a number of models. In the base model of Bertsimas and Lo [1], 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 the time increments within the trading horizon. Further developments have led to optimal execution algorithms for models that incorporate price predictions [1], bid-ask spreads and resilience [2, 3], nonlinear price impact models [4, 5], and risk aversion [6, 7, 8, 9, 10].

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, allowing 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 decreases. The arbitrageur's actions amplify price impact and therefore increase execution costs.

Several recent papers study game-theoretic models of execution in the presence of strategic arbitrageurs [11, 12, 13]. 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 deliberate management of the signals he transmits by influencing prices. 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 price impact model of Bertsimas and Lo [1]. 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 surprisingly simple structure:

   (a) Trades placed by the trader are linear in the arbitrageur's estimation error and the arbitrageur's expectation of the combined position of the trader and the arbitrageur.

   (b) Trades placed by the arbitrageur are linear in the arbitrageur's expectation of the combined position of the trader and the arbitrageur.

   (c) The equilibrium policies are a function of the time horizon and a single 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 perfect Bayesian equilibrium respond to price fluctuations; the trader leverages these random outcomes to shade 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 market surplus. 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.

Beyond the immediate application to the optimal execution problem, the results in this paper also represent a contribution to the general theory of games with asymmetric information. Equilibrium in such games is notoriously difficult to compute. Typical games that have been considered are basic signaling games (see [14, Chapter 8] and the references therein), where the game has two periods and the private information takes the form of a binary-valued "type". In contrast, the game considered here has an arbitrary discrete time horizon, and the private information (the position of the trader) is a continuous value.

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. Finally, Section 6 makes some closing remarks and suggests directions for future work. Proofs of all theoretical results are presented in the appendix.

2 Problem Formulation

We consider a game that evolves over a finite horizon in discrete time steps $t = 0, \ldots, T$. 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. We denote his position at each time t by x_t. The trader requires that his final position x_T be zero. The arbitrageur begins with a position y_0. We denote his position at each time t by y_t. He requires that y_T be zero.

The price of the stock evolves according to

$$
\begin{aligned}
p_t &= p_{t-1} + \lambda(u_t + v_t) + \epsilon_t \\
&= p_{t-1} + \Delta p_t.
\end{aligned}
$$

where u_t is the quantity purchased by the trader and v_t is the quantity purchased by the arbitrageur. The sequence $\{\epsilon_t\}$ is IID 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. The positions evolve according to

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

The information structure of the game is as follows. The dynamics of the game (in particular, the parameters λ and σ_ϵ) and the common 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 changes Δp_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, we assume 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 of the form described below 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 6.

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. We consider only trader policies for which $\pi_T(x_{T-1}, y_{T-1}, \phi_{T-1}) = -x_{T-1}$; i.e., policies that result in liquidation. We denote the set of trader policies by Π. Similarly, the arbitrageur follows a policy $\psi = \{\psi_1, \ldots, \psi_T\}$. Each function ψ_{t+1} maps y_t and ϕ_t to a decision v_{t+1} made at time t. We restrict attention to arbitrageur policies for which $\psi_T(y_{T-1}, \phi_{T-1}) = -y_{T-1}$. We denote the set of arbitrageur policies by Ψ.

Note that we are restricting ourselves to policies that are Markovian in the sense that 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 that each player's action is only a function of his state. Further, we are assuming that 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

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

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.

We consider a profit to be a 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 profit generated by the trader and arbitrageur through their trades u_{t+1} and v_{t+1} are

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

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

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

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

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

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

Note that $-U_t^{\pi,(\psi,\pi)}(x_0, y_0, \phi_0)$ 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

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

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

$$U_t^{\pi,(\psi,\hat{\pi})}(x_t, y_t, \phi_t) = \max_{\pi' \in \Pi} U_t^{\pi',(\psi,\hat{\pi})}(x_t, y_t, \phi_t), \tag{2.1}$$

for all t, x_t, y_t, and ϕ_t. Further, we will refer to $\psi \in \Psi$ as a best response to $\pi \in \Pi$ if

$$V_t^{\psi,\pi}(y_t, \phi_t) = \max_{\psi' \in \Psi} V_t^{\psi',\pi}(y_t, \phi_t), \tag{2.2}$$

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 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 PBE actions only for cases where ϕ_t is Gaussian. When the initial distribution ϕ_0 is Gaussian and players employ these PBE policies, subsequent distributions ϕ_t are also Gaussian. As such, computation of PBE policies over the restricted domain 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* (2.1) *(or* (2.2)*) 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,* $\phi_1, \ldots, \phi_{T-1}$ *are Gaussian.*

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

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

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

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

for all 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 Bayesian updating 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. In addition, for each π_t and v_t, we define a dynamic programming operator $G_{v_t}^{\pi_t}$ by

$$\left(G_{v_t}^{\pi_t}V\right)(y_{t-1},\phi_{t-1}) \equiv \mathsf{E}_{v_t}^{\pi_t}\left[\lambda(u_t+v_t)y_{t-1}+V(y_t,\phi_t) \mid y_{t-1},\phi_{t-1}\right],$$

for all 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 Bayesian updating given that the arbitrageur correctly assumes the trader trades u_t.

Consider Algorithm 1. 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 $U_t^* = U_t^{\pi^*,\psi^*}$ and $V_t^* = V_t^{\psi^*,\pi^*}$. 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. This can not be done on a computer.

Algorithm 1 PBE Solver

1: Initialize the terminal value functions by setting, for all x_{T-1}, y_{T-1}, and ϕ_{T-1},

$$U_{T-1}^*(x_{T-1}, y_{T-1}, \phi_{T-1}) \leftarrow -\lambda(x_{T-1} + y_{T-1})x_{T-1}$$

$$V_{T-1}^*(y_{T-1}, \phi_{T-1}) \leftarrow -\lambda \left(\int x\phi_{T-1}(dx) + y_{T-1} \right) y_{T-1}$$

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},

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

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

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

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

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

5: **end for**

3.2 Quasilinear Policies

Given a distribution ϕ_t, define

$$\mu_t \equiv \int x\phi_t(dx), \qquad \sigma_t^2 \equiv \int (x - \mu_t)^2 \phi_t(dx), \quad \text{and} \quad \rho_t \equiv \lambda\sigma_t/\sigma_\epsilon.$$

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.

We will consider restricting attention to a class of policies that are indexed by a few parameters.

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

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

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

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

We will also refer to a policy as quasilinear if component functions associated with times $1, \ldots, T-1$ are quasilinear.

Note that quasilinear policies have a particularly intuitive structure. For the arbitrageur, at each time t, a quasilinear policy is a linear function of $y_{t-1} + \mu_{t-1}$. This quantity can be interpreted as the arbitrageur's estimate of the total market overhang at time $t-1$, that is, the number of shares outstanding which must be liquidated by time T. A quasilinear policy for the trader at time t, in addition, depends linearly on the quantity $x_{t-1} - \mu_{t-1}$. This is the error of the arbitrageur's estimate of the trader's position, that is, the private knowledge of the trader.

By restricting attention to quasilinear 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 quasilinear. 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 ϕ_{t-1} is Gaussian, $\hat{\pi}_t$ is quasilinear, and the arbitrageur assumes that the trader trades $\hat{\pi}_t(x_{t-1}, y_{t-1}, \phi_{t-1})$, then ϕ_t is Gaussian.*

It follows from this result that if π^* is quasilinear 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 quasilinear 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.

Algorithm 2 Quasilinear-Gaussian PBE Solver

1: Initialize the terminal value functions by setting, for all x_{T-1}, y_{T-1}, and Gaussian ϕ_{T-1},

$$U_{T-1}^*(x_{T-1}, y_{T-1}, \phi_{T-1}) \leftarrow -\lambda(x_{T-1} + y_{T-1})x_{T-1}$$

$$V_{T-1}^*(y_{T-1}, \phi_{T-1}) \leftarrow -\lambda \left(\int x \phi_{T-1}(dx) + y_{T-1} \right) y_{T-1}$$

2: **for** $t = T-1, T-2, \ldots, 1$ **do**

3: Compute quasilinear (π_t^*, ψ_t^*) such that for all x_{t-1}, y_{t-1}, and Gaussian ϕ_{t-1},

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

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

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

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

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

5: **end for**

For the remainder of this paper, we will focus on computation of quasilinear-Gaussian PBE, and as such, we will restrict attention to Gaussian beliefs, with all policies and value functions defined only over this restricted domain.

3.3 Value Function Decomposition

Value functions computed by Algorithm 2 exhibit special structure that simplifies their representation. We now define the form of this special structure.

Definition 4. *A function U_t is* **trader-quadratic-decomposable** *(TQD) if there are coefficients $c_{xx,t}^{\rho_t}$, $c_{yy,t}^{\rho_t}$, $c_{xy,t}^{\rho_t}$, and $c_{0,t}^{\rho_t}$, which are functions of ρ_t, such that*

$$U_t(x_t, y_t, \phi_t) = \lambda\Bigg(\tfrac{1}{2}(y_t^2 - \mu_t^2) + \tfrac{1}{2}(x_t - \mu_t)(y_t - \mu_t) - \tfrac{1}{2}c_{xx,t}^{\rho_t}(x_t - \mu_t)^2 \\ - \tfrac{1}{2}c_{yy,t}^{\rho_t}(y_t + \mu_t)^2 - c_{xy,t}^{\rho_t}(x_t - \mu_t)(y_t + \mu_t) + \frac{\sigma_\epsilon^2}{\lambda^2}c_{0,t}^{\rho_t}\Bigg),$$

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

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

for all y_t and ϕ_t.

It is clear that U_{T-1}^* and V_{T-1}^* are TQD/AQD. The following theorem captures how TQD and AQD structure are retained through the recursion of Algorithm 2.

Theorem 2. *If U_t^* is TQD and V_t^* is AQD, and Step 3 of Algorithm 2 produces a quasilinear pair (π_t^*, ψ_t^*), then U_{t-1}^* and V_{t-1}^*, defined by Step 4 of Algorithm 2 are TQD and AQD.*

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 Existence

Algorithm 2 relies for each t on existence of a pair (π_t^*, ψ_t^*) of quasilinear functions that satisfy single-stage equilibrium conditions. In Section 5, for a range of problem instances, we compute quasilinear functions that satisfy such equilibrium conditions. However, whether such equilibria exist for all cases remains an open issue. Here, we support plausibility by presenting results on best responses to quasilinear policies. The first asserts that if ψ_t and $\hat{\pi}_t$ are quasilinear then there is a quasilinear best-response π_t in the single-stage game.

Theorem 3. *If U_t is TQD, ψ_t is quasilinear, and $\hat{\pi}_t$ is quasilinear, then there exists a quasilinear 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 π_t is quasilinear then there is a quasilinear best-response ψ_t in the single-stage game.

Theorem 4. *If V_t is AQD and π_t is quasilinear then there exists a quasilinear 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 quasilinear policy then it suffices for the trader (arbitrageur) to restrict himself to quasilinear policies. Though not a proof of existence, this observation that the set of quasilinear policies is closed under the operation of best response motivates an aim to compute quasilinear-Gaussian PBE.

3.5 Dependence on Problem Data

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

$$\left\{ a_{x,t+1}^{\rho_t}, a_{y,t+1}^{\rho_t}, b_{y,t+1}^{\rho_t}, c_{xx,t}^{\rho_t}, c_{yy,t}^{\rho_t}, c_{xy,t}^{\rho_t}, c_{0,t}^{\rho_t}, d_{yy,t}^{\rho_t}, d_{0,t}^{\rho_t} \right\},$$

for every $\rho_t > 0$ and $t = 0, \ldots, T-2$. These coefficients parameterize quasilinear-Gaussian PBE policies and corresponding value functions. Note that the output depends on λ and

σ_ϵ only through ρ_t. Hence, given any λ and σ_ϵ, 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.

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

Consider a quasilinear-Gaussian PBE (π^*, ψ^*). Since π_t^* and ψ_t^* are quasilinear, they can be written as

$$\begin{aligned}
\pi_t^*(x_{t-1}, y_{t-1}, \phi_{t-1}) &= a_{x,t}^{\rho_{t-1}}(x_{t-1} - \mu_{t-1}) + a_{y,t}^{\rho_{t-1}}(y_{t-1} + \mu_{t-1}), \\
\psi_t^*(y_{t-1}, \phi_{t-1}) &= b_{y,t}^{\rho_{t-1}}(y_{t-1} + \mu_{t-1}).
\end{aligned}$$

for all x_{t-1}, y_{t-1}, and ϕ_{t-1}. Here, the coefficients are deterministic functions of ρ_{t-1}. For a fixed value of ρ_{t-1}, the coefficients can be stored as three numerical values. However, it is not feasible to simultaneously store coefficients associated with all possible values of ρ_{t-1}. Fortunately, as established in the following result, the trader's policy π^* and the initial value ρ_0 determine subsequent values of ρ_t.

Theorem 5. *If ϕ_{t-1} is Gaussian, and the arbitrageur assumes that the trader's policy $\hat{\pi}_t$ is quasilinear with*

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

then 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}.

It follows that for a fixed value of ρ_0, over the relevant portion of its domain, a quasilinear-Gaussian PBE can be encoded in terms of $3(T-1)$ numerical values. We will design an algorithm that aims to compute these $3(T-1)$ parameters, which we will denote by $a_{x,t}$, $a_{y,t}$ and $b_{y,t}$, for $t = 1, \dots, T-1$. Note that 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 $a_{x,t}$, $a_{y,t}$, and $b_{y,t}$ characterize quasilinear-Gaussian PBE policies restricted to the sequence $\rho_0, \dots, \rho_{T-1}$ generated in the quasilinear-Gaussian PBE. We do not know in advance what this sequence will be, and as such, our algorithm will simultaneously compute this sequence alongside the policy parameters.

Algorithm 3 searches for ρ_{T-1}. For each candidate $\hat{\rho}_{T-1}$, a recursion computes preceding values $\hat{\rho}_{T-2}, \dots, \hat{\rho}_0$ along with policy parameters for times $T-1, \dots, 1$. Assuming that single-stage equilibria are successfully computed along the way, the resulting policies form a quasilinear-Gaussian PBE, restricted to the sequence $\hat{\rho}_0, \dots, \hat{\rho}_{T-1}$ that they would generate if $\rho_0 = \hat{\rho}_0$. The search algorithm seeks a value of $\hat{\rho}_{T-1}$ such that the resulting $\hat{\rho}_0$ is indeed equal to ρ_0. Since information accumulates, it is natural to conjecture that in a quasilinear-Gaussian PBE, each ρ_t is monotonically increasing in ρ_{t-1}, and therefore, ρ_{T-1} is monotonically increasing in ρ_0. This motivates the bisection search: if a choice of $\hat{\rho}_{T-1}$ leads to a value $\hat{\rho}_0 > \rho_0$, the value should be reduced, and vice versa. The search begins with upper and lower bounds of 0 and $\min(\rho_0, 1)$; it is not hard to establish that ρ_{T-1} is within these bounds. This search procedure is reminiscent of the work of Kyle [15], in a different context.

Note that Step 7 of the algorithm treats $\hat{\rho}_{t-1}$ as a free variable that is solved alongside the policy parameters $a_{x,t}$, $a_{y,t}$, and $b_{y,t}$. These variables can be computed through solving a cubic equation, as discussed in Appendix B. Algorithm 3 is implementable and we use it in computational studies presented in the next section.

Algorithm 3 Quasilinear-Gaussian PBE Solver with Variance Search

$\underline{\rho}_{T-1} \leftarrow 0$
$\overline{\rho}_{T-1} \leftarrow \min(\rho_0, 1)$
while $\overline{\rho}_{T-1} - \underline{\rho}_{T-1} > \delta$ **do**
  $\hat{\rho}_{T-1} \leftarrow (\overline{\rho}_{T-1} + \underline{\rho}_{T-1})/2$
  Initialize the terminal value functions by setting, for all x_{T-1}, y_{T-1}, and Gaussian ϕ_{T-1} with variance $(\hat{\rho}_{T-1}\sigma_\epsilon/\lambda)^2$,

$$U^*_{T-1}(x_{T-1}, y_{T-1}, \phi_{T-1}) \leftarrow -\lambda(x_{T-1} + y_{T-1})x_{T-1}$$

$$V^*_{T-1}(y_{T-1}, \phi_{T-1}) \leftarrow -\lambda(\mu_{T-1} + y_{T-1})y_{T-1}$$

  for $t = T-1, T-2, \ldots, 1$ **do**
    Compute $\hat{\rho}_{t-1}$ and quasilinear (π^*_t, ψ^*_t) such that for all x_{t-1}, y_{t-1}, and Gaussian ϕ_{t-1} with variance $(\hat{\rho}_{t-1}\sigma_\epsilon/\lambda)^2$,

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

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

$$\rho_t = \hat{\rho}_t$$

    Compute the value functions at the previous time step by setting, for all x_{t-1}, y_{t-1}, and Gaussian ϕ_{t-1} with variance $(\hat{\rho}_{t-1}\sigma_\epsilon/\lambda)^2$,

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

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

  end for
  if $\hat{\rho}_0 \leq \rho_0$ **then**
    $\underline{\rho}_{T-1} \leftarrow \hat{\rho}_{T-1}$
  else
    $\overline{\rho}_{T-1} \leftarrow \hat{\rho}_{T-1}$
  end if
end while

5 Computational Results

In this section, we present computational results generated using Algorithm 3. In Section 5.1, we introduce two alternative policies, the equipartitioning policy and the minimum revelation policy. These are intuitive polices which will serve as a basis of comparison to the quasilinear-Gaussian PBE policy. In Section 5.2, we discuss the importance of the parameter $\rho_0 \equiv \lambda\sigma_0/\sigma_\epsilon$ in the qualitative behavior of the PBE policy and interpret ρ_0^2 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 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 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 quasilinear-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 in [1]. We call the resulting policy π^{EQ} an *equipartitioning* policy. It is defined by

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

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. Clearly, trading during the final time period T reveals no relevant information to the arbitrageur. It is further true that trading during the penultimate time period $T-1$ reveals no useful information to the arbitrageur. This is because the arbitrageur is constrained to liquidate his remaining holdings at time T, hence the arbitrageur's decision at time T is not influenced by his belief ϕ_{T-1}. We define the *minimum revelation* policy π^{MR} to be a policy that efficiently exploits these facts by liquidating the trader's position evenly across only the last two time periods. That is,

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

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

5.2 Relative Volume

As we observed in Section 4.1, quasilinear-Gaussian PBE policies are determined as a function of the composite parameter $\rho_0 \equiv \lambda\sigma_0/\sigma_\epsilon$. 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, we interpreted ϵ_t as the exogenous, random component of price changes. Alternatively, we 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,

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

In other words, we can interpret ρ_0 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 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

In this section, we will compare how various policies for the trader perform.

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

$$\mathsf{E}\left[U_0^{\pi,\psi}(x_0, 0, \phi_0) \;\middle|\; x_0 \sim \phi_0 \right].$$

Alternatively, if the trader liquidates his entire position immediately, the expected profit becomes

$$\mathsf{E}\left[-\lambda x_0^2 \;\middle|\; x_0 \sim \phi_0 \right] = -\lambda \sigma_0^2.$$

We define the *trader's normalized profit* $\bar{U}(\pi, \psi)$ to be the ratio

$$\bar{U}(\pi, \psi) \equiv \frac{\mathsf{E}\left[U_0^{\pi,\psi}(x_0, 0, \phi_0) \,\middle|\, x_0 \sim \phi_0 \right]}{\lambda \sigma_0^2}.$$

Similarly, the *arbitrageur's normalized profit* $\bar{V}(\pi, \psi)$ is defined to be

$$\bar{V}(\pi, \psi) \equiv \frac{\mathsf{E}\left[V_0^{\pi,\psi}(x_0, 0, \phi_0) \,\middle|\, x_0 \sim \phi_0 \right]}{\lambda \sigma_0^2}.$$

Given a quasilinear-Gaussian PBE (π^*, ψ^*), since the value function $U_0^{\pi^*,\psi^*}$ is TQD, we have

$$\bar{U}(\pi^*, \psi^*) = \frac{\lambda \left(-\frac{1}{2} c_{xx,0}^{\rho_0} \sigma_0^2 + \frac{\sigma_\epsilon^2}{\lambda^2} c_{0,0}^{\rho_0} \right)}{\lambda \sigma_0^2} = -\frac{1}{2} c_{xx,0}^{\rho_0} + \frac{1}{\rho_0^2} c_{0,0}^{\rho_0},$$

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

$$\bar{V}(\pi^*, \psi^*) = \frac{\lambda \left(\frac{\sigma_\epsilon^2}{\lambda^2} d_{0,0}^{\rho_0} \right)}{\lambda \sigma_0^2} = \frac{1}{\rho_0^2} d_{0,0}^{\rho_0},$$

where $d_{0,0}^{\rho_0}$ is the arbitrageur's appropriate value function coefficients at time $t = 0$. Thus, the normalized profits of the PBE policy depends on the parameters $(\sigma_0, \lambda, \sigma_\epsilon)$ only through the quantity $\rho_0 \equiv \lambda \sigma_0 / \sigma_\epsilon$.

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 (2.2), via dynamic programming. Using a similar argument as above, it is easy to see that $\bar{U}(\pi^{\mathrm{EQ}}, \psi^{\mathrm{EQ}})$ and $\bar{V}(\pi^{\mathrm{EQ}}, \psi^{\mathrm{EQ}})$ 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,

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

In Figure 1, the relative profit of the 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 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 fairs particularly badly from the perspective of the trader, performing a up to a factor of 2 worse than the PBE policy. 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 $\psi_t^0 \equiv 0$ 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.

Examining Figure 1, it is clear that, for any given pair of policies, the magnitude of the normalized loss of the trader exceeds the normalized profit of the arbitrageur. The difference in these two quantities can be interpreted as a benefit to the other participants

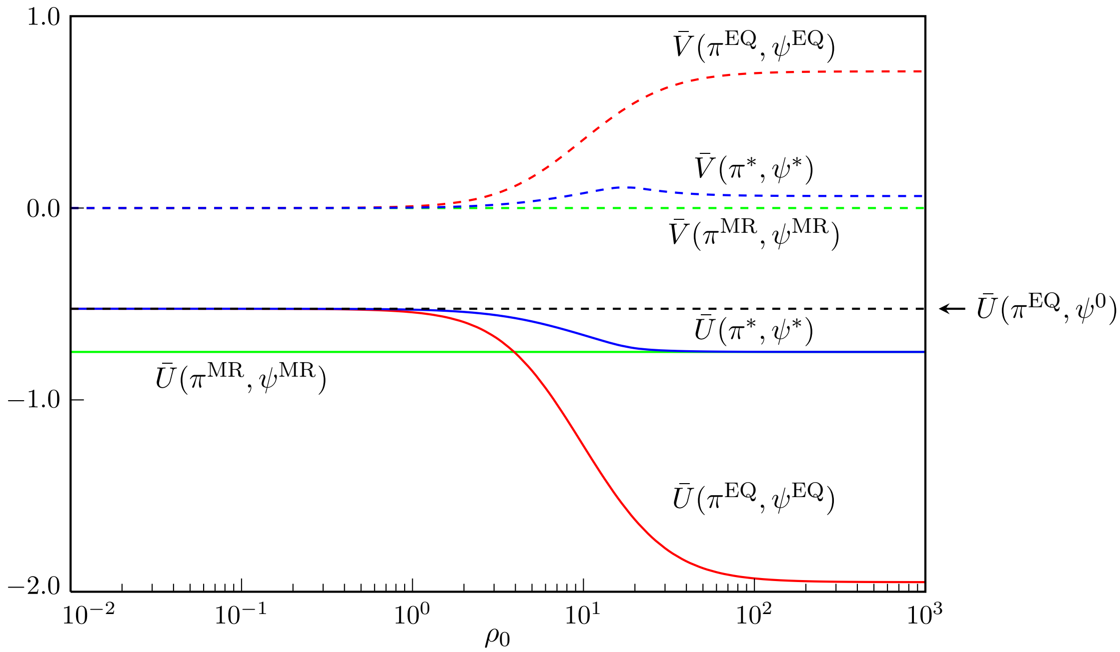


Figure 1: The relative profit of trading strategies for the time horizon $T = 20$.

in the market. Define the *market surplus* to be the quantity

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

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

5.4 Signaling

An important aspect of the PBE policy is that it accounts for information conveyed through price movements. In order to understand this feature, we define the relative uncertainty to be the standard deviation of the arbitrageur's belief of the trader's decision 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 PBE policy versus the equipartitioning and minimum

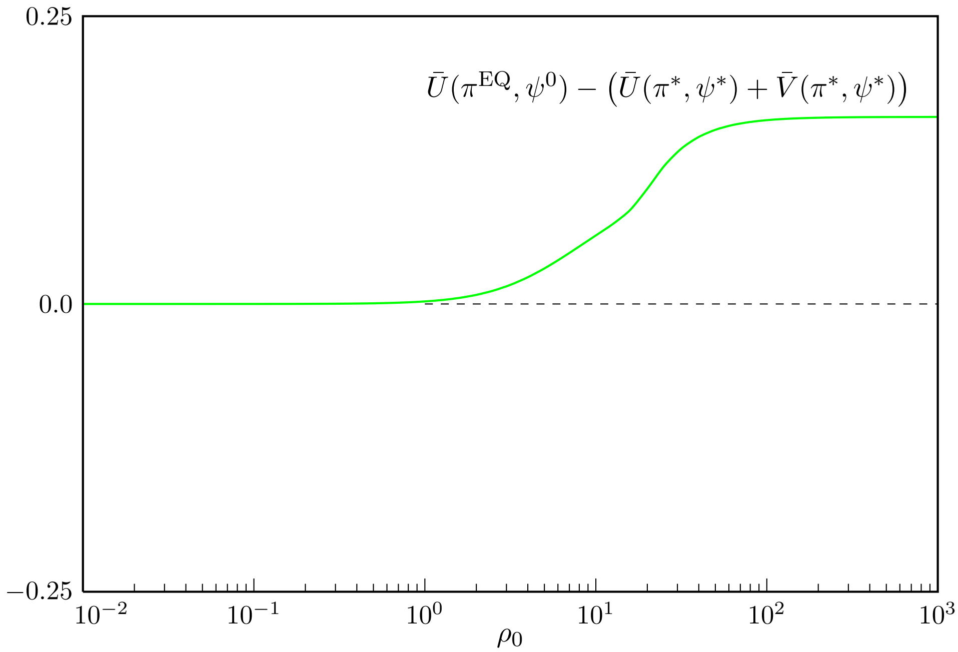


Figure 2: The market surplus of the system for the time horizon $T = 20$.

revelation policies, we can study the comparative signaling behavior.

Relative uncertainty has another interpretation. If we assume, as in Section 5.3, that the arbitrageur has initial position $y_0 = 0$ and initial belief $\phi_0 = N(0, \sigma_0^2)$, and that the trader's initial position is sampled from the distribution ϕ_0 (and thus is consistent with the arbitrageur's belief), then

$$\frac{\sigma_t}{\sigma_0} = \sqrt{\frac{\mathsf{E}\left[x_t^2 \mid x_0 \sim \phi_0\right]}{\mathsf{E}\left[x_0^2 \mid x_0 \sim \phi_0\right]}}.$$

Thus, relative uncertainty at time t gives a measure of the size of the trader's outstanding position at that time, in a root-mean-squared sense.

Under the PBE 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. Under the equipartitioning policy, the relative uncertainty decreases linearly, according to

$$\frac{\sigma_t}{\sigma_0} = \frac{T-t}{T}.$$

Under the minimum revelation policy, the relative uncertainty decays only over the final two time steps, according to

$$\frac{\sigma_t}{\sigma_0} = \begin{cases} 1 & \text{if } t < T-1, \\ \frac{1}{2} & \text{if } t = T-1, \\ 0 & \text{if } t = T. \end{cases}$$

In Figure 3, we can see the evolution of the relative uncertainty of the PBE policy, 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 very similarly to that of the equipartitioning policy, decaying almost linearly. In the high relative volume regime, almost very little information is revealed until close to the end of the trading period. These observations are consistent with our results from Section 5.3.

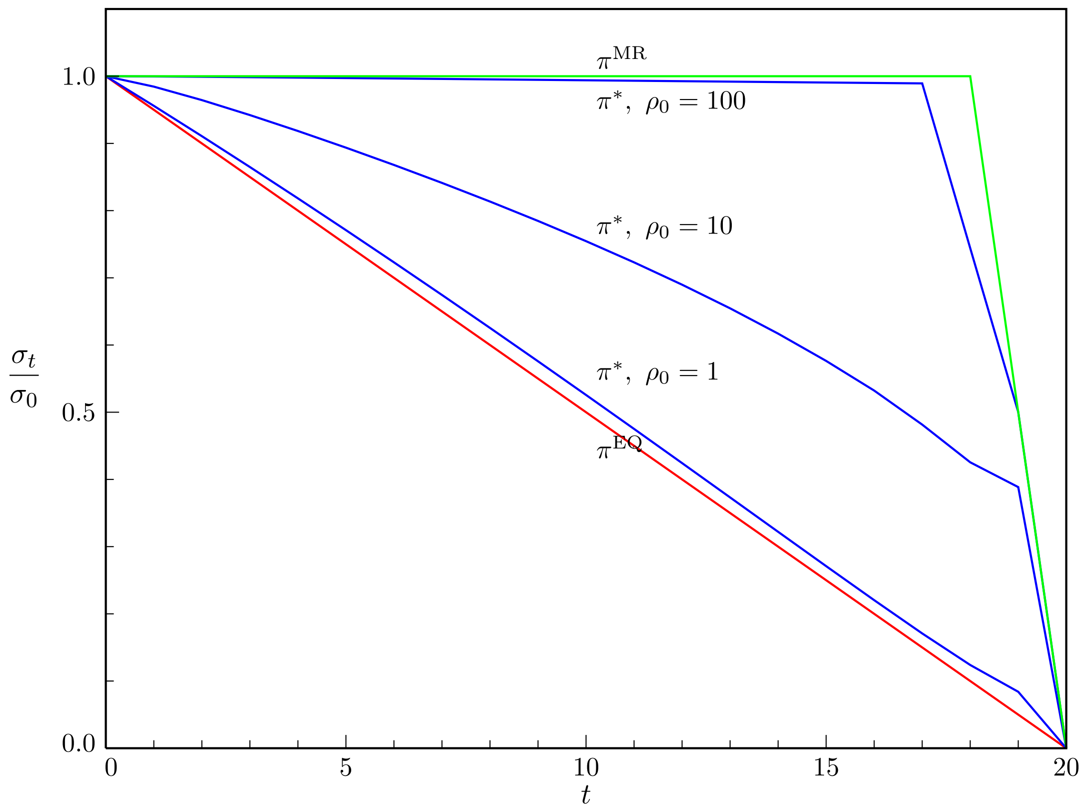


Figure 3: The evolution over time of the relative uncertainty σ_t/σ_0 of the trader's position for the time horizon $T = 20$.

5.5 Dynamic Trading

One important feature of the PBE policy is that it is dynamic and exhibits complex behavior that is market dependent. The quantities traded depend on the random exogenous fluctuations of the market. Indeed, the trader may seek to exploit these fluctuations so as to minimize execution costs. This is in contrast to the equipartitioning and minimum revelation policies, which are deterministic.

We can observe this dynamic behavior as follows: define the random variable

$$\Delta \equiv \sum_{t=1}^{T} \epsilon_t.$$

The variable Δ is the cumulative exogenous movement of the market over the trading horizon. Define

$$\bar{x}_t \equiv \mathsf{E}[x_t \mid \Delta], \quad \bar{y}_t \equiv \mathsf{E}[y_t \mid \Delta], \quad \bar{\mu}_t \equiv \mathsf{E}[\mu_t \mid \Delta].$$

These quantities are, respectively, the expectation of the trader's position, the arbitrageur's position, and the arbitrageur's mean belief, conditioned on a particular level of cumulative market movement. By conditioning on the variable Δ, we can explore the most likely behavior of the system under various market scenarios.

Figure 4 plots the evolution of $(\bar{x}_t, \bar{y}_t, \bar{\mu}_t)$ under such several scenarios, given the parameters

$$x_0 = \sigma_0 = 10^5, \quad \mu_0 = 0, \quad \lambda = 5 \times 10^{-5}, \quad \sigma_\epsilon = 0.125, \quad T = 20.$$

(Here, we use values for λ, σ_ϵ, and T suggested in [1].) Note that, in this instance, $x_0 \neq \mu_0$. That is, the arbitrageur's initial mean estimate is incorrect.

In Figure 4(a), we see a neutral market scenario, where $\Delta = 0$. Note that, since $\rho_0 = 40$, the system is in a high relative volume regime. Hence, the trader attempts to conceal his true position and trades only minimally prior to the end of the time horizon.

In Figure 4(b), we see a 2 standard deviation up market scenario, where $\Delta = 2\sigma_\epsilon\sqrt{T}$. Here, the exogenous upward movement of the market leads the arbitrageur to believe that the trader is short the stock, when, in fact, the trader is long. The trader then anticipates buying on the part of the arbitrageur, and seeks to exploit this by increasing his position.

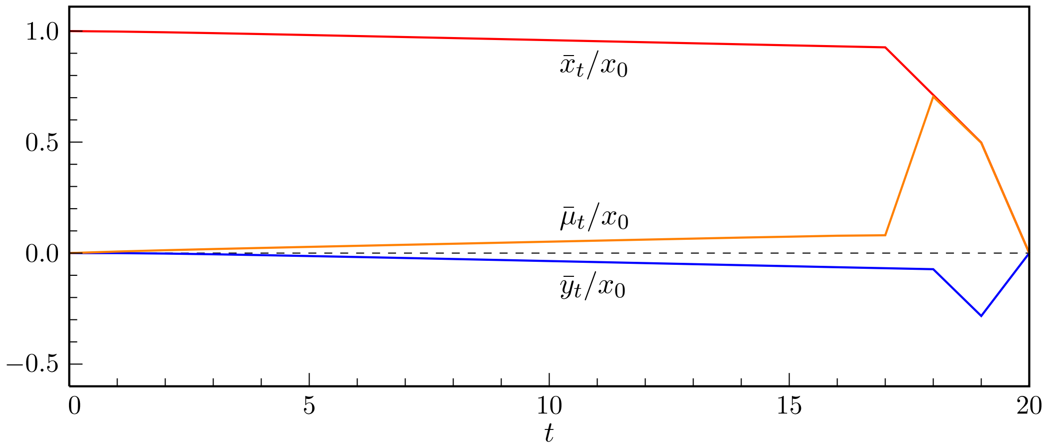
In Figure 4(c), we see a 2 standard deviation down market scenario, where $\Delta = -2\sigma_\epsilon \sqrt{T}$. In this case, the arbitrageur assumes that the downward movement of the market is due to selling on the part of the trader, and attempts to front-run future selling. The trader is thus forced to liquidate his position faster than in the other scenarios.

6 Conclusion

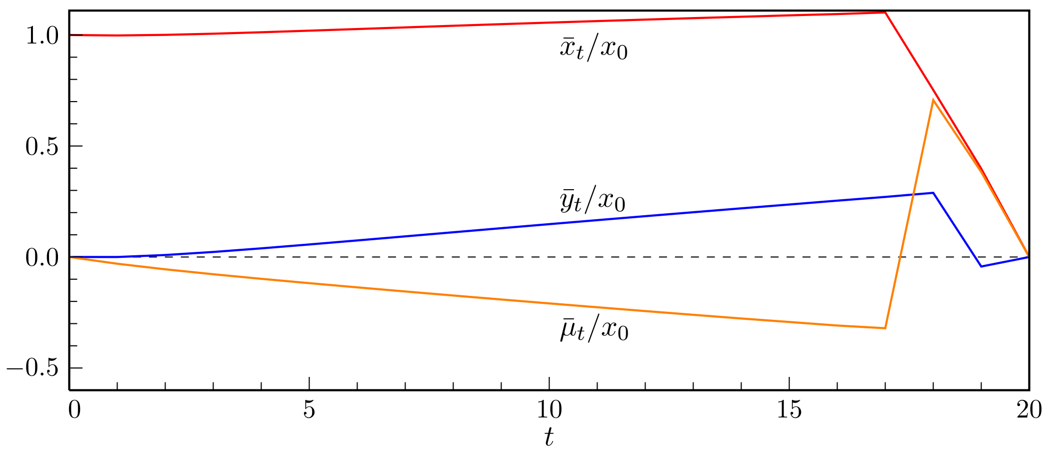
Our model captures strategic interactions between a trader aiming to liquidate a position and an arbitrageur trying to detect and profit from the trader's activity. The algorithm we have developed computes perfect Bayesian equilibrium behavior. It is interesting that the resulting trader policy takes on such a simple form: the number of shares to liquidate at time t is linear in the difference $x_{t-1} - \mu_{t-1}$ between the trader's position and the arbitrageur's estimate and the sum $y_{t-1}+\mu_{t-1}$ of the arbitrageur's position and his estimate of the trader's position. 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, takes advantage of situations where the arbitrageur has been misled by unusual market activity, and occasionally places trades intended to mislead the arbitrageur.

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:

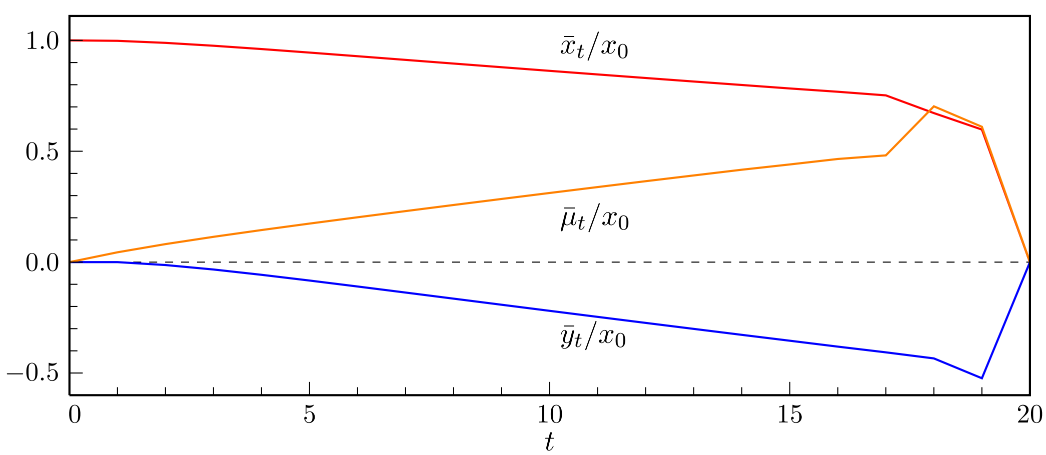
1. **(Risk Aversion)** We assume that both the trader and arbitrageur are risk-neutral.



(a) Neutral market, $\Delta = 0$.



(b) Up market, $\Delta = 2\sigma_\epsilon\sqrt{T}$.



(c) Down market, $\Delta = -2\sigma_\epsilon\sqrt{T}$.

Figure 4: The expected evolution over time of the trader's position, the arbitrageur's position, and the arbitrageur's mean belief, conditioned on the exogenous market movement.

Risk aversion is clearly an important facet of investor behavior, and should be included in the trade execution model.

2. **(Flexible Time Horizon)** We assume a finite time horizon T for the trader and arbitrageur. 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.

3. **(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.

4. **(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.

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 in [16].

Acknowledgments

The first author wishes to thank Mark Broadie and Gabriel Weintraub for helpful comments. The second author is grateful to 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

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

[2] A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. preprint, 2005.

[3] A. Alfonsi, A. Schied, and A. Schulz. Constrained portfolio liquidation in a limit order book model. preprint, 2007.

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

[5] A. Alfonsi, A. Schied, and A. Schulz. Optimal execution strategies in limit order books with general shape functions. preprint, 2007.

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

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

[8] 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.

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

[10] A. Schied and T. Schönenborn. Optimal portfolio liquidation for CARA investors. preprint, 2007.

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

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

[13] T. Schönenborn and A. Schied. Liquidation in the face of adversity: stealth vs. sunshine trading, predatory trading vs. liquidity provision. preprint, 2007.

[14] D. Fudenberg and J. Tirole. *Game Theroy*. MIT Press, Cambridge, MA, 1991.

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

[16] 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 Proofs

Theorem 1. *If ϕ_{t-1} is Gaussian, $\hat{\pi}_t$ is quasilinear, and the arbitrageur assumes that the trader trades $\hat{\pi}_t(x_{t-1}, y_{t-1}, \phi_{t-1})$, then ϕ_t is Gaussian.*

Proof. Suppose that

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

where $\hat{a}_{\mu,t}^{\rho_{t-1}} \equiv \hat{a}_{y,t}^{\rho_{t-1}} - \hat{a}_{x,t}^{\rho_{t-1}}$. Set (K_{t-1}, h_{t-1}) to be the information form parameters for the Gaussian distribution ϕ_{t-1}, so that

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

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,

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

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

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

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

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

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

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

and mean

$$
\begin{aligned}
\mu_t &= \hat{a}^{\rho_{t-1}}_{y,t} y_{t-1} + \hat{a}^{\rho_{t-1}}_{\mu,t}\mu_{t-1} \\
&\quad + (1+\hat{a}^{\rho_{t-1}}_{x,t}) \frac{h_{t-1} + \frac{\lambda\left(\Delta p_t - \lambda(\hat{a}^{\rho_{t-1}}_{y,t} y_{t-1} + \hat{a}^{\rho_{t-1}}_{\mu,t}\mu_{t-1} + \psi_t)\right)\hat{a}_{x,t}}{\sigma^2_\epsilon}}{K_{t-1} + \frac{\lambda^2(\hat{a}^{\rho_{t-1}}_{x,t})^2}{\sigma^2_\epsilon}} \\
&= \hat{a}^{\rho_{t-1}}_{y,t}(y_{t-1}+\mu_{t-1}) - \hat{a}^{\rho_{t-1}}_{x,t}\mu_{t-1} \\
&\quad + (1+\hat{a}^{\rho_{t-1}}_{x,t}) \frac{\frac{\mu_{t-1}}{\sigma^2_{t-1}} + \frac{\lambda\left(\Delta p_t - \lambda(\hat{a}^{\rho_{t-1}}_{y,t}(y_{t-1}+\mu_{t-1}) - \hat{a}^{\rho_{t-1}}_{x,t}\mu_{t-1} + \psi_t)\right)\hat{a}_{x,t}}{\sigma^2_\epsilon}}{\frac{1}{\sigma^2_{t-1}} + \frac{\lambda^2(\hat{a}^{\rho_{t-1}}_{x,t})^2}{\sigma^2_\epsilon}} \qquad \text{(A.2)} \\
&= \hat{a}^{\rho_{t-1}}_{y,t}(y_{t-1}+\mu_{t-1}) - \hat{a}^{\rho_{t-1}}_{x,t}\mu_{t-1} \\
&\quad + (1+\hat{a}^{\rho_{t-1}}_{x,t}) \frac{\mu_{t-1}/\rho^2_{t-1} + \left(\Delta p_t/\lambda - \hat{a}^{\rho_{t-1}}_{y,t}(y_{t-1}+\mu_{t-1}) + \hat{a}^{\rho_{t-1}}_{x,t}\mu_{t-1} - \psi_t\right)\hat{a}_{x,t}}{1/\rho^2_{t-1} + (\hat{a}^{\rho_{t-1}}_{x,t})^2}.
\end{aligned}
$$

■

In order to prove Theorems 2–4, it is necessary to explicitly evaluate the operator $F^{(\psi_t,\pi_t)}_{u_t}$ applied to quadratic functions of (x_t, y_t, μ_t) and the operator $G^{\pi_t}_{v_t}$ applied to quadratic functions of (y_t, μ_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 quasilinear policies ψ_t and π_t are defined so that*

$$
\begin{aligned}
\pi_t(x_{t-1}, y_{t-1}, \phi_{t-1}) &= a^{\rho_{t-1}}_{x,t}(x_{t-1}-\mu_{t-1}) + a^{\rho_{t-1}}_{y,t}(y_{t-1}+\mu_{t-1}), \\
\psi_t(y_{t-1}, \phi_{t-1}) &= b^{\rho_{t-1}}_{y,t}(y_{t-1}+\mu_{t-1}).
\end{aligned}
$$

Define

$$
\gamma^{\rho_{t-1}}_t \equiv \frac{1+a^{\rho_{t-1}}_{x,t}}{1/\rho^2_{t-1} + (a^{\rho_{t-1}}_{x,t})^2}.
$$

Then,

$$
\text{(A.3a)} \qquad \begin{aligned} \mathsf{E}_{u_t}^{(\psi_t,\pi_t)}\left[\mu_t \mid x_{t-1}, y_{t-1}, \phi_{t-1}\right] = a_{y,t}^{\rho_{t-1}}(y_{t-1}+\mu_{t-1}) - a_{x,t}^{\rho_{t-1}}\mu_{t-1} \\ + \gamma_t^{\rho_{t-1}}/\rho_{t-1}^2 \\ + \gamma_t^{\rho_{t-1}} a_{x,t}^{\rho_{t-1}}\left(u_t + a_{x,t}^{\rho_{t-1}}\mu_{t-1} - a_{y,t}^{\rho_{t-1}}(y_{t-1}+\mu_{t-1})\right), \end{aligned}
$$

$$
\text{(A.3b)} \qquad \mathrm{Var}_{u_t}^{(\psi_t,\pi_t)}\left[\mu_t \mid x_{t-1}, y_{t-1}, \phi_{t-1}\right] = \left(\gamma_t^{\rho_{t-1}} a_{x,t}^{\rho_{t-1}} \sigma_\epsilon/\lambda\right)^2,
$$

$$
\text{(A.3c)} \qquad \begin{aligned} \mathsf{E}_{u_t}^{(\psi_t,\pi_t)}\left[\mu_t^2 \mid x_{t-1}, y_{t-1}, \phi_{t-1}\right] = \mathrm{Var}_{u_t}^{(\psi_t,\pi_t)}\left[\mu_t \mid x_{t-1}, y_{t-1}, \phi_{t-1}\right], \\ + \left(\mathsf{E}_{u_t}^{(\psi_t,\pi_t)}\left[\mu_t \mid x_{t-1}, y_{t-1}, \phi_{t-1}\right]\right)^2, \end{aligned}
$$

$$
\text{(A.3d)} \qquad \mathsf{E}_{v_t}^{\pi_t}\left[\mu_t \mid y_{t-1}, \phi_{t-1}\right] = a_{y,t}^{\rho_{t-1}}(y_{t-1}+\mu_{t-1}) + \mu_{t-1},
$$

$$
\text{(A.3e)} \qquad \mathrm{Var}_{v_t}^{\pi_t}\left[\mu_t \mid y_{t-1}, \phi_{t-1}\right] = \left(\gamma_t^{\rho_{t-1}} a_{x,t}^{\rho_{t-1}} \sigma_\epsilon/\lambda\right)^2 \left(1 + \left(a_{x,t}^{\rho_{t-1}}\right)^2 \rho_{t-1}^2\right),
$$

$$
\text{(A.3f)} \qquad \mathsf{E}_{v_t}^{\pi_t}\left[\mu_t^2 \mid y_{t-1}, \phi_{t-1}\right] = \mathrm{Var}_{v_t}^{\pi_t}\left[\mu_t \mid y_{t-1}, \phi_{t-1}\right] + \left(\mathsf{E}_{v_t}^{\pi_t}\left[\mu_t \mid y_{t-1}, \phi_{t-1}\right]\right)^2.
$$

Proof. The lemma follows directly from taking expectations of the mean update equation (A.2). ■

Theorem 2. *If U_t^* is TQD and V_t^* is AQD, and Step 3 of Algorithm 2 produces a quasilinear pair (π_t^*, ψ_t^*), then U_{t-1}^* and V_{t-1}^*, defined by Step 4 of Algorithm 2 are TQD and AQD.*

Proof. Suppose that

$$
\begin{aligned} V_t^*(y_t, \phi_t) &= \lambda\left(-\tfrac{1}{2}(y_t^2 - \mu_t^2) - \tfrac{1}{2}d_{yy,t}^{\rho_t}(y_t+\mu_t)^2 + \frac{\sigma_\epsilon^2}{\lambda^2}d_{0,t}^{\rho_t}\right), \\ \pi_t^*(x_{t-1}, y_{t-1}, \phi_{t-1}) &= a_{x,t}^{\rho_{t-1}}(x_{t-1}-\mu_{t-1}) + a_{y,t}^{\rho_{t-1}}(y_{t-1}+\mu_{t-1}), \\ \psi_t^*(y_{t-1}, \phi_{t-1}) &= b_{y,t}^{\rho_{t-1}}(y_{t-1}+\mu_{t-1}). \end{aligned}
$$

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

$$
\begin{aligned}
u_t &= a_{x,t}^{\rho_{t-1}}(x_{t-1} - \mu_{t-1}) + a_{y,t}^{\rho_{t-1}}(y_{t-1} + \mu_{t-1}), \\
v_t &= b_{y,t}^{\rho_{t-1}}(y_{t-1} + \mu_{t-1}), \\
y_t &= y_{t-1} + b_{y,t}^{\rho_{t-1}}(y_{t-1} + \mu_{t-1})
\end{aligned}
$$

Using these facts, and (A.3d)–(A.3f) from Lemma 1, we can explicitly compute

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

where

$$
\begin{aligned}
d_{yy,t-1}^{\rho_{t-1}} &= \left(b_{y,t}^{\rho_{t-1}}\right)^2 - 2a_{y,t}^{\rho_{t-1}} - \left(a_{y,t}^{\rho_{t-1}}\right)^2 + \left(1 + b_{y,t}^{\rho_{t-1}} + a_{y,t}^{\rho_{t-1}}\right)^2 d_{yy,t}^{\rho_t}, \\
d_{0,t-1}^{\rho_{t-1}} &= d_{0,t}^{\rho_t} + \tfrac{1}{2}\left(\gamma_t^{\rho_{t-1}} a_{x,t}^{\rho_{t-1}}\right)^2 (1 - d_{yy,t}^{\rho_t})\left(1 + \left(a_{x,t}^{\rho_{t-1}}\rho_{t-1}\right)^2\right).
\end{aligned}
$$

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

Theorem 3. *If U_t is TQD, ψ_t is quasilinear, and $\hat{\pi}_t$ is quasilinear, then there exists a quasilinear 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.

Proof. Suppose that

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

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

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

Using these facts, and (A.3a)–(A.3c) from Lemma 1, we can explicitly compute

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

It is easily checked that $\big(F_{u_t}^{(\psi_t,\hat{\pi}_t)} U_t\big)(x_{t-1}, y_{t-1}, \phi_{t-1})$ 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 linear in $(x_{t-1}, y_{t-1}, \mu_{t-1})$. The value of u_t^* can be explicitly computed as

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

where

$$
\begin{aligned}
\rho_t^2 &= \big(1 + \hat{a}_{x,t}^{\rho_{t-1}}\big)^2 \left(\frac{1}{\rho_{t-1}^2} + (\hat{a}_{x,t}^{\rho_{t-1}})^2\right)^{-1}, \\
a_{x,t}^{\rho_{t-1}} &= \frac{1}{2Z_t^{\rho_t}}\big(-2c_{xx,t}^{\rho_t} + \gamma_t^{\rho_{t-1}}\hat{a}_{x,t}^{\rho_{t-1}}(2c_{xx,t}^{\rho_t} - 2c_{xy,t}^{\rho_t} - 1) + 2\big), \\
a_{y,t}^{\rho_{t-1}} &= \frac{1}{2Z_t^{\rho_t}}\Big((\gamma_t^{\rho_{t-1}}\hat{a}_{x,t}^{\rho_{t-1}} - 1)Y_{x,t}^{\rho_{t-1}} \\
&\qquad + 2\gamma_t^{\rho_{t-1}}\hat{a}_{x,t}^{\rho_{t-1}}(-b_{y,t}^{\rho_{t-1}} + \gamma_t^{\rho_{t-1}}\hat{a}_{x,t}^{\rho_{t-1}}\hat{a}_{y,t}^{\rho_{t-1}} - \hat{a}_{y,t}^{\rho_{t-1}} - 1)c_{yy,t}^{\rho_t}\Big), \\
Y_{x,t}^{\rho_{t-1}} &= (b_{y,t}^{\rho_{t-1}} + 1)(2c_{xy,t}^{\rho_t} - 1) \\
&\qquad + \hat{a}_{y,t}^{\rho_{t-1}}\big(2(\gamma_t^{\rho_{t-1}}\hat{a}_{x,t}^{\rho_{t-1}} - 1)c_{xx,t}^{\rho_t} + (2 - 4\gamma_t^{\rho_{t-1}}\hat{a}_{x,t}^{\rho_{t-1}})c_{xy,t}^{\rho_t} + 1\big), \\
Z_t^{\rho_t} &= c_{xx,t}^{\rho_t} + \gamma_t^{\rho_{t-1}}\hat{a}_{x,t}^{\rho_{t-1}}\left(\left(\gamma_t^{\rho_{t-1}}\hat{a}_{x,t}^{\rho_{t-1}} - 2\right)c_{xx,t}^{\rho_t} + 2c_{xy,t}^{\rho_t} + \gamma_t^{\rho_{t-1}}\hat{a}_{x,t}^{\rho_{t-1}}\left(c_{yy,t}^{\rho_t} - 2c_{xy,t}^{\rho_t}\right) + 1\right).
\end{aligned}
$$

Clearly, u_t^* is quasilinear. ■

Theorem 4. *If V_t is AQD and π_t is quasilinear then there exists a quasilinear 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

$$\begin{aligned} V_t(y_t, \phi_t) &= \lambda \left(-\tfrac{1}{2}(y_t^2 - \mu_t^2) - \tfrac{1}{2} d_{yy,t}^{\rho_t} (y_t + \mu_t)^2 + \frac{\sigma_\epsilon^2}{\lambda^2} d_{0,t}^{\rho_t} \right), \\ \pi_t(x_{t-1}, y_{t-1}, \phi_{t-1}) &= a_{x,t}^{\rho_{t-1}} (x_{t-1} - \mu_{t-1}) + a_{y,t}^{\rho_{t-1}} (y_{t-1} + \mu_{t-1}). \end{aligned}$$

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

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

Using these facts, and (A.3d)–(A.3f) from Lemma 1, we can explicitly compute

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

It is easily checked that $\left(G_{v_t}^{\pi_t} V_t \right)(y_{t-1}, \phi_{t-1})$ is quadratic in v_t. Moreover, the coefficient of v_t^2 is independent of (y_{t-1}, μ_{t-1}) while the coefficient of v_t is linear in (y_{t-1}, μ_{t-1}). Therefore, the optimizing v_t^* is linear in (y_{t-1}, μ_{t-1}). The value of v_t^* can be explicitly computed as

$$v_t^* = - \frac{\left(a_{y,t}^{\rho_{t-1}} + 1 \right) d_{yy,t}^{\rho_t}}{d_{yy,t}^{\rho_t} + 1} (y_{t-1} + \mu_{t-1}),$$

where

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

Clearly, v_t^* is quasilinear. ■

Theorem 5. *If ϕ_{t-1} is Gaussian, and the arbitrageur assumes that the trader's policy $\hat{\pi}_t$ is* *quasilinear with*

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

then 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. The result follows directly from (A.1) in the proof of Theorem 1. ■

B Single-Stage Quasilinear Equilibrium Computation

In Step 7 of Algorithm 3, we solve for the single-stage quasilinear equilibrium policy parameters $(a_{x,t}, a_{y,t}, b_{y,t})$ and the scaled uncertainty parameter $\hat{\rho}_{t-1}$. In this section, we describe how this is accomplished as follows.

First, for every $\rho > 0$, define $\mathcal{N}(\rho)$ to be the set of Gaussian distributions with variance $(\rho\sigma_\epsilon/\lambda)^2$. By hypothesis, the value of $\hat{\rho}_t$ is fixed. Thus, we can assume that $\phi_t \in \mathcal{N}(\hat{\rho}_t)$. Now, suppose that $\phi_{t-1} \in \mathcal{N}(\hat{\rho}_{t-1})$, for some $\hat{\rho}_{t-1} > 0$ (which we will solve for shortly).

Since V_t^* is AQD and U_t^* is TQD, they can be parameterized for $\phi_t \in \mathcal{N}(\hat{\rho}_t)$ as

$$\begin{aligned} V_t^*(y_t, \mu_t) &= \lambda \left(-\tfrac{1}{2}(y_t^2 - \mu_t^2) - \tfrac{1}{2} d_{yy,t}(y_t + \mu_t)^2 + \frac{\sigma_\epsilon^2}{\lambda^2} d_{0,t} \right), \\ U_t^*(x_t, y_t, \mu_t) &= \lambda \Big(\tfrac{1}{2}(y_t^2 - \mu_t^2) + \tfrac{1}{2}(x_t - \mu_t)(y_t - \mu_t) - \tfrac{1}{2} c_{xx,t}(x_t - \mu_t)^2 \\ &\qquad - \tfrac{1}{2} c_{yy,t}(y_t + \mu_t)^2 - c_{xy,t}(x_t - \mu_t)(y_t + \mu_t) + \frac{\sigma_\epsilon^2}{\lambda^2} c_{0,t} \Big). \end{aligned}$$

Now, suppose that that the arbitrageur believes the trader is employing a policy

$$\hat{\pi}_t(x_{t-1}, y_{t-1}, \mu_{t-1}) = \hat{a}_{x,t}(x_{t-1} - \mu_{t-1}) + \hat{a}_{y,t}(y_{t-1} + \mu_{t-1}),$$

for all $\phi_{t-1} \in \mathcal{N}(\hat{\rho}_{t-1})$. From Theorem 5, in order to guarantee that $\phi_t \in \mathcal{N}(\hat{\rho}_t)$, it must be the case either $\hat{a}_{x,t} = -1$ (in which case $\hat{\rho}_{t-1}$ is indeterminate) or that

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

By the same arguments as in Theorem 4, the arbitrageur's optimal response problem

$$\psi_t^*(y_{t-1}, \mu_{t-1}) \in \operatorname*{argmax}_{v_t} \left(G_{v_t}^{\hat{\pi}_t} V_t^* \right) (y_{t-1}, \mu_{t-1})$$

is bounded if

$$1 + d_{yy,t} > 0, \tag{B.1}$$

in which case it has a unique optimal solution of the form

$$\psi_t^*(y_{t-1}, \mu_{t-1}) = b_{yy,t}(y_{t-1} - \mu_{t-1}),$$

where

$$b_{yy,t} = -\frac{(\hat{a}_{y,t} + 1)\, d_{yy,t}}{d_{yy,t} + 1}.$$

Similarly, by the same arguments as in Theorem 3, the trader's optimal response problem

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

is bounded if

$$c_{xx,t} + \gamma_t a_{x,t} \left((\gamma_t a_{x,t} - 2)\, c_{xx,t} + 2c_{xy,t} + \gamma_t a_{x,t} \left(c_{yy,t} - 2c_{xy,t} \right) + 1 \right) > 0, \tag{B.2}$$

where

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

In this case, it has a unique optimal solution of the form

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

In order for the PBE to exist, $\hat{\pi}_t = \pi_t^*$ must hold. Explicit solution of the optimal response problems for the trader and arbitrageur reveal that sufficient conditions for this are:

1. $a_{x,t}$ solves the cubic polynomial

$$\begin{aligned} 0 = {} & 2\left(c_{xx,t}\left(1/\rho_t^2 - 1\right)^2 + 2c_{xy,t}\left(1/\rho_t^2 - 1\right) + c_{yy,t} + 1/\rho_t^2\right)(a_{x,t})^3 \\ & + 1/\rho_t^2\left(6c_{xy,t} + 6c_{xx,t}\left(1/\rho_t^2 - 1\right) - 2/\rho_t^2 + 3\right)(a_{x,t})^2 \\ & + 1/\rho_t^2\left(2c_{xy,t} - c_{xx,t}\left(2 - 6/\rho_t^2\right) - 4/\rho_t^2 + 1\right)a_{x,t} \\ & + 2\left(c_{xx,t} - 1\right)/\rho_t^4. \end{aligned} \tag{B.3}$$

2. The trader and arbitrageur's optimal response problems are bounded, that is, (B.1) and (B.2) hold.

3. The variance of the distributions at time $t-1$ is well-defined and unique, that is,

$$a_{x,t} \neq -1,$$

   and

$$\hat{\rho}_{t-1}^2 = \left(\left(\frac{1 + a_{x,t}}{\hat{\rho}_t}\right)^2 - (\hat{a}_{x,t})^2\right)^{-1} > 0.$$

4. $a_{y,t}$ and $b_{y,t}$ are determined according to

$$\begin{aligned} a_{y,t} &= \frac{2\left(d_{yy,t} + 1\right)}{2d_{yy,t} + 2c_{xy,t} + \gamma_t a_{x,t}\left(-2c_{xy,t} + 2c_{yy,t} + 1\right) + 1} - 1, \\ b_{y,t} &= -\frac{2d_{yy,t}}{2d_{yy,t} + 2c_{xy,t} + \gamma_t a_{x,t}\left(-2c_{xy,t} + 2c_{yy,t} + 1\right) + 1}. \end{aligned}$$

In practice, we first solve for the roots of (B.3) for putative values of $a_{x,t}$. For each root, we attempt to verify the remainder of the conditions.

It may be the case that it is impossible to satisfy all of the conditions. In this case, we assume that the value of $\hat{\rho}_t$ was set too high, and that there does not exist an equilibrium with variance $(\hat{\rho}_t \sigma_\epsilon / \lambda)^2$ at time t. Therefore, we escape from the loop immediately and lower the guess of ρ_{T-1}. Equivalently, we set

$$\overline{\rho}_{T-1} \leftarrow (\underline{\rho}_{T-1} + \overline{\rho}_{T-1})/2.$$

We resume the loop with this new upper bound.