

# Optimal Microstructure Trading With a Long-Term Utility Function

E. Jerome Benveniste and Gordon Ritter

<https://ssrn.com/abstract=3057570>

Authors both affiliated with NYU Courant Institute, NYU Tandon School.

GR also teaches statistics at Rutgers and in the MFE program at Baruch, and is Senior Portfolio Manager at GSA Capital.

- ▶ Consider a trader who may have existing inventory, and who is solving a longer-term utility of wealth problem.
- ▶ The trader interacts with a market for  $n$  distinct assets, where  $n$  is allowed to be large (at least several thousand).
- ▶ Most modern exchanges and ECNs operate as continuous limit-order book markets.
- ▶ The agent can choose various ways of participating in a limit-order book market (cross the spread (aggressive), join the queue on the near side (passive), cancel all orders, etc.)  
The trader must then decide actions to take on the limit-order book markets for all of the securities.
- ▶ This is a hard problem, but it has a *smooth relaxation* which is more tractable, and which we now describe.

- ▶ Let  $q(t)$  denote the trader's continuous-time path, so  $q(t) \in \mathbb{R}^n$  denotes the portfolio holdings at time  $t$ , in numéraire units such as dollars.
- ▶ The *trading rate* (in dollars per unit time) is defined to be

$$\dot{q}(t) = \frac{dq(t)}{dt}$$

- ▶ In the smooth relaxation, the trader's problem can be formulated as follows:

$$\min_q \int_0^T L(q(t), \dot{q}(t)) dt \quad \text{s.t.} \quad q(0) = q_0 \quad (1)$$

where  $L$  is called the *Lagrangian*, and the minimum is taken over all paths  $q(t)$ .

- ▶ We emphasize that (1) is a *relaxation* of the true problem.
- ▶ Portfolio holdings actually live in a lattice in  $\mathbb{R}^n$ , since one cannot hold less than one share.
- ▶ Furthermore, certain actions available to an investor (or agent) affect the holdings  $q(t)$  in a non-deterministic way.
- ▶ An example of this non-determinism is that if the trader joins a queue using a limit order that does not get matched immediately, then whether this order will get filled is a random variable.

- ▶ A typical example of a Lagrangian which is relevant to finance is

$$L(q, v) = c(v) + \text{mean-variance utility of } q(t) \quad (2)$$

where  $c : \mathbb{R}^n \rightarrow \mathbb{R}$  is the *instantaneous trading cost function* defined as the continuous-time limit of any ordinary trading cost function.

- ▶ In the single-security case, if we trade  $\delta q$  dollars in some small time interval of length  $\delta t$ , and this costs  $\lambda \cdot \delta q / \delta t$  times traded notional for some  $\lambda > 0$ , then the total cost in dollars per unit time is

$$\lambda(\delta q / \delta t)^2 \equiv c(\delta q / \delta t)$$

where  $c(v) = \lambda v^2$ , however in general,  $c()$  need not be quadratic.

- ▶ Intuitively,  $c(v)$  can be thought of as containing permanent impact as well as the average instantaneous cost of a small order over a small time window.
- ▶ Costs for individual actions at the microstructure level will of course deviate from the average cost – we will ultimately optimize over these actions as well.

- ▶ Note that the Almgren–Chriss model for liquidation is the special case of problem (1) with Lagrangian (2) and where the mean-variance utility is only variance, and where we have a fixed endpoint  $q(T) = 0$ .

- ▶ At a stationary point, the Euler-Lagrange equation is satisfied:

$$\frac{d}{dt} \frac{\delta L}{\delta \dot{q}} - \frac{\delta L}{\delta q} = 0.$$

- ▶ A first-order system can be obtained if we introduce a new variable  $p$ , whose components are called *generalized momenta* and defined as

$$p := \frac{\delta L}{\delta v}(q, \dot{q}) \quad (3)$$

- ▶ If the conditions of the implicit function theorem are satisfied, we could in principle algebraically solve (3) for  $\dot{q}$ , obtaining

$$\dot{q} = \phi(q, p)$$

for some function  $\phi$  defined implicitly by (3).

- ▶ The Euler equation then takes the form

$$\dot{p} = \frac{\delta L}{\delta q}(q, \dot{q}) = \frac{\delta L}{\delta q}(q, \phi(q, p)) \equiv \psi(q, p)$$

- ▶ As the functions  $\phi, \psi$  are algebraic (not involving derivatives), we have a system of  $2n$  first-order ODE given by

$$\dot{q} = \phi(q, p), \quad \dot{p} = \psi(q, p) \quad (4)$$

- ▶ These equations can be expressed more symmetrically by introducing the *Hamiltonian*

$$H(q, p) = p\phi(q, p) - L(q, \phi(q, p))$$

- ▶ Eqns. (4) are equivalently written in a form known as *Hamilton's equations*:

$$\dot{q} = \frac{\delta H}{\delta p}, \quad \dot{p} = -\frac{\delta H}{\delta q} \quad (5)$$

- ▶ In many cases, Hamilton's equations can be solved explicitly; to illustrate this, we will consider a classic example, which is (2) with a fixed endpoint  $q(T) = 0$ .

- ▶ Suppose  $c(v) = \frac{1}{2}v'\Lambda v$  where  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  is a diagonal matrix with  $\lambda_i > 0$  for all  $i$ .
- ▶ From (3), the generalized momenta are  $p = \Lambda\dot{q}$  and hence algebraically solving, one has  $\phi(q, p) = \Lambda^{-1}p$ .
- ▶ The Hamiltonian is then

$$H(q, p) = \frac{1}{2}p\Lambda^{-1}p - \frac{1}{2}\kappa q'\Sigma q.$$

- ▶ Hamilton's equations then become:

$$\dot{q} = \Lambda^{-1} p, \quad \dot{p} = \kappa \Sigma q \quad (6)$$

or more simply  $dx/dt = Ax$  where  $x : \mathbb{R} \rightarrow \mathbb{R}^{2n}$  and  $A \in M(2n; \mathbb{R})$  are given by

$$x = \begin{pmatrix} q \\ p \end{pmatrix}, \quad A = \begin{pmatrix} 0 & \Lambda^{-1} \\ \kappa \Sigma & 0 \end{pmatrix}. \quad (7)$$

- ▶ Duhamel's formula yields  $x(t) = e^{tA} x_0$ , which can be computed by using the Jordan canonical form of  $A$ , or simply diagonalizing  $A$  if possible.
- ▶ Note that this diagonalization only needs to be recomputed when the model underlying the agent's utility function (risk model, alphas, etc) changes; it does not need to be recomputed in response to high frequency microstructure data.

- ▶ The *value function* is defined by

$$V(t, x) = - \min_q \int_t^T L(q(s), \dot{q}(s)) ds \quad \text{s.t.} \quad q(t) = x \quad (8)$$

- ▶ Thus  $V(t, x)$  is the remaining utility gain from time  $t$  obtained from following the best policy, when the current state at time  $t$  is  $x$ .
- ▶ We use the negative of the integral so that a higher value is better.

- ▶ In our paper it has been shown that along an optimal trajectory  $q$ , we have

$$\nabla V = \frac{\partial L}{\partial v}(q, \dot{q}) = p$$

where the second equality is the definition of  $p$  from eq. (3).

- ▶ The fact that  $p = \nabla V$  really helps the intuition: modifying trading rates to align with  $p$  is equivalent to pointing the trader in the steepest ascent direction for the value function.

- ▶ We now restrict attention to Lagrangians that take the form

$$L(q, v) = c(v) + f(q)$$

which includes the mean-variance example discussed before.

- ▶ Using the convex duality between the Lagrangian and Hamiltonian, we note that the optimal instantaneous trade  $\dot{q}(t)$  at each time  $t$  is the argument  $v$  which solves

$$\dot{q}^* = \underset{v}{\operatorname{argmax}} [p \cdot v - c(v)]. \quad (9)$$

- ▶ In the absence of costs, the objective function in (9) is maximized when the vector of instantaneous trading rates points in the same direction as  $p$ .
- ▶ Turning this around gives intuition about the meaning of  $p$ : it's the direction in  $\mathbb{R}^n$  which, ideally, you'd like the vector of trading rates to point in.

- ▶ Solving (9) falls short of a full set of instructions for the trader who has to interact with a continuous limit order book market.
- ▶ In reality The trader must choose the action  $a \in \mathcal{A}$  from a discrete menu  $\mathcal{A}$  of actions called the action space.
- ▶ In fact we will always consider *finite* action spaces  $|\mathcal{A}| < \infty$ .
- ▶ The discrete choice of action must nonetheless take into account the long-term utility as expressed by the value function.

- ▶ In principle, finding the optimal discrete choice involves solving the Bellman equation:

$$\operatorname{argmax}_{a \in \mathcal{A}} \int d^n x p(\delta, x | a) \mathbb{E}[\delta + V(t, x)]$$

where  $\delta$  denotes the immediate reward associated to the action.

- ▶ However, this would require calculating the value function for all possible portfolios  $x \in \mathbb{R}^n$ , and doing a high-dimensional integral (when  $n$  is large).
- ▶ We searched instead for a way to use only *local* information, (gradient of the value function), instead of *global* information (integral of the value function over all space).

- ▶ The local version is as follows.
- ▶ Note that finding paths satisfying (9) is another way of solving the Bellman equation.
- ▶ Re-write (9) as

$$\operatorname{argmax}_v \mathbb{E}[\tau \cdot v - c(v)] \quad (10)$$

where  $\tau$  is a random variable having the property  $\mathbb{E}[\tau] = p$ .

- ▶ Eqn. (10) can be interpreted as maximizing expected profit net of costs in a model where the gradient of the value function,  $p$ , plays the role of microstructure alphas:

$$\operatorname{argmax}_v \mathbb{E}[\text{wealth}(v)], \quad \text{where } \mathbb{E}[\tau] = p. \quad (11)$$

- ▶ A hypothetical myopic risk-neutral agent who simply trades to maximize expected profit of the next trade will actually exhibit fully optimal behavior as long as this agent uses the generalized momenta as microstructure alphas.

- ▶ Recall that we cannot choose  $v$ ; rather, we choose an action from the (finite) menu of actions  $\mathcal{A}$  to try to implement the trading at rate  $v$ .
- ▶ Let us replace (11) with an implementable version:

$$\operatorname{argmax}_{a \in \mathcal{A}} \mathbb{E}[\text{wealth} \mid a], \quad \text{where} \quad \mathbb{E}[\tau] = \rho. \quad (12)$$

- ▶ Eq. (12) enjoys a very special role; we view it as the fundamental key insight which allows solutions of this entire general class of continuous Lagrangian-based problems with discrete action spaces.

- ▶ We now embark upon an extended example to illustrate the procedure of applying Eq. (12), Consider an agent interacting with a continuous limit-order book market.
- ▶ The simplest (but still realistic) action space we could imagine is the following one, which essentially distills the complex modus operandi of algorithmic execution into three possible actions: wait, aggressive, and passive, which we describe in turn.
- ▶ The wait action is simply to cancel all orders and wait.
- ▶ Then in this example,  $|\mathcal{A}| = 3$ .

- ▶ Aggressive execution refers to trading primarily by means of market orders (or marketable limit orders) and ensures that a child order will be filled completely and with some immediacy, but also pays a real (and perhaps very high) cost for the convenience of immediacy.

- ▶ Passive execution corresponds roughly to trading via limit order placement strategies which aim to constantly remain in the queue on the near side of the spread until the order is filled, presumably by an adversary who chooses to aggress.
- ▶ Passive execution avoids both spread pay and, arguably, avoids market impact as well.
- ▶ There is still impact in the sense that some liquidity on the other side is removed from the market, but the price at which passive orders are filled – if they are – is guaranteed to be better than any level on the other side of the book.

- ▶ Let  $a_i$  be the aggression level in the  $i$ -th security.
- ▶ The fill probability is conditional on the aggression level.
- ▶ Define  $f_i$  to be the fill probability conditional on the *passive* level of aggression:

$$p_f | a_i = \begin{cases} 0 & a_i = \mathfrak{w} := \text{wait} \\ f_i & a_i = \mathfrak{p} := \text{passive} \\ 1 & a_i = \mathfrak{a} := \text{aggressive} \end{cases} \quad (13)$$

- ▶ In this generalized situation, we assume that (10) remains valid, but the agent must choose one of the three levels of aggression, denoted  $a$ , and the expectation in (10) is computed over a larger probability space which includes the randomness of getting filled if the agent should happen to choose  $a = p$ , the passive action.
- ▶ Let us write

$$R_{i,t}(v, a)$$

for the (random) profit or loss from an order of quantity  $v$  on stock  $i$  using action  $a$  over a short interval  $[t, t + \epsilon]$ .

- ▶ Then

$$R_{i,t}(v, a) = \mathbf{1}_{\text{fill}}(r_i v_i - \text{cost}(v_i, a))$$

where  $r_i = \text{mid}_i(t + \epsilon) / \text{mid}_i(t) - 1$  is the midpoint return over the interval,  $\mathbf{1}_{\text{fill}}$  is the indicator variable for the event of the order being filled (for simplicity, we assume there are no partial fills), and  $\text{cost}(v_i, a)$  is the cost of filling the quantity  $v_i$  by using the action  $a$ .

- ▶ This cost function will in general be different from the cost function  $c(v)$  appearing in our objective, which represents an average cost of trading, without reference to particular execution choices.
- ▶ For example, for  $a = \alpha$ ,  $\text{cost}(v, a)$  will include the spread; for  $a = \beta$ , it may include an adverse selection cost.
- ▶ In our setting, the myopic risk-neutral agent's beliefs are such that  $\mathbb{E}[r_i] = \rho_i$ .

- ▶ For the sake of deriving a concrete trading rule, assume that  $r_i$  and  $\mathbf{1}_{\text{fill}}$  are independent, that passive executions incur the average cost, and that aggressive ones, in addition, incur a spread cost.
- ▶ Then we have

$$\mathbb{E}[R(v_i, a_i) | a_i] = \begin{cases} 0 & a_i = \mathfrak{w} \\ f_i(p_i v_i - c(v_i)) & a_i = \mathfrak{p} \\ p_i v_i - c(v_i) - s_i v_i & a_i = \mathfrak{a} \end{cases} \quad (14)$$

where  $s_i$  denotes one-half the bid-offer spread.

- ▶ We assume that, since the aggressive action involves executing at the far side of the limit order book, the agent incurs a cost of one half-spread relative to the mid price if the agent chooses  $a_i = \mathfrak{a}$ , ie. chooses to aggress.

- ▶ Under the assumptions used to derive (14), the action  $a$  which optimizes

$$\mathbb{E}[R(v_i, a_i) | a]$$

is determined as follows.

- ▶ If

$$p_i v_i - c(v_i) - s_i v_i > f_i(p_i v_i - c(v_i)) > 0$$

then the agent chooses  $a = \alpha$ , to aggress.

- ▶ If  $p_i v_i > c(v_i)$ , the agent chooses the passive mode,  $a = \beta$ .
- ▶ Finally if  $p_i v_i < c(v_i)$  the agent chooses to wait,  $a = \omega$  (intuitively because further trading would go against the gradient of the value function).

- ▶ Consider a hypothetical execution desk with orders to execute in  $n$  distinct assets, where  $n$  is large (many thousands).
- ▶ If the desk attempts to complete this task by only trading passively, they will encounter serious problems: namely, they will fall behind on orders which are moving in the direction of the trade, and they will build up unwanted factor exposure (and associated risk) at the portfolio level.
- ▶ This can be catastrophic, as the examples in our paper show.

- ▶ Suppose the desk's portfolio has gotten “out of balance.”
- ▶ Then one may ask: “is there a vector of microstructure alphas which point the way back to optimality?”
- ▶ We answer this in the affirmative: the necessary microstructure alphas are given by the generalized momenta, or equivalently the gradient of the value function.
- ▶ A hypothetical myopic risk-neutral agent who simply trades to maximize expected profit of the next trade will actually exhibit fully optimal behavior as long as this agent uses the generalized momenta as microstructure alphas.
- ▶ This agent solves:

$$\operatorname{argmax}_{a \in \mathcal{A}} \mathbb{E}[\text{wealth} \mid a], \quad \text{where} \quad \mathbb{E}[\tau] = p.$$

- ▶ To be practical, the myopic expected-profit maximizations would need to be performed very quickly, because today's electronic limit order book markets can change rapidly, and experience high message throughput.
- ▶ The only “hard” part is computing  $p$ , which changes each time the portfolio changes (ie. each time an order is filled).
- ▶ Fortunately, this can be done in real time; our method is computationally tractable.