Algorithmic Trading with Model Uncertainty

Álvaro Cartea[a], Ryan Donnelly[b], Sebastian Jaimungal[c]

[a] *Mathematical Institute, University of Oxford, UK*
Oxford-Man Institute of Quantitative Finance, Oxford, UK

[b] *Swiss Finance Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland*

[c] *Department of Statistical Sciences, University of Toronto, Canada*

Abstract

Algorithmic traders acknowledge that their models are incorrectly specified, thus we allow for ambiguity in their choices to make their models robust to misspecification in: (i) the arrival rate of market orders (MOs), (ii) the fill probability of limit orders, and (iii) the dynamics of the midprice of the asset they deal. In the context of market making, we demonstrate that market makers (MMs) adjust their quotes to reduce inventory risk and adverse selection costs. Moreover, robust market making increases the strategies Sharpe ratio and allows the MM to fine tune the tradeoff between the mean and the standard deviation of profits. We provide analytical solutions for the robust optimal strategies, show that the resulting dynamic programming equations have classical solutions and provide a proof of verification. The behavior of the ambiguity averse MM are found to generalize those of a risk averse MM, and coincide in a limiting case.

Keywords: Market Making, Ambiguity Aversion, Model Uncertainty, Algorithmic Trading, High Frequency Trading, Short Term Alpha, Adverse Selection, Robust Optimization

1. Introduction

Market makers (MMs) provide liquidity to investors who require immediacy by posting bid and offer prices to the market. In exchange, MMs expect to earn the spread between these sell and buy quotes, but bear the risk of providing liquidity at a loss when trading with better informed

Email addresses: alvaro.cartea@maths.ox.ac.uk (Álvaro Cartea), ryan.donnelly@epfl.ch (Ryan Donnelly), sebastian.jaimungal@utoronto.ca (Sebastian Jaimungal)

RD and SJ thank NSERC for partially funding this research. This paper has benefited from comments of seminar participants at the Isaac Newton Institute – Cambridge, The Fields Institute, University of Oxford, The University of Chicago, University of Michigan, and University College London. For comments and suggestions on earlier drafts, we are indebted to Erhan Bayraktar, René Carmona, Sam Cohen, Mike Ludkovski, Michael Monoyios, Christoph Reisinger.

traders and also face the risk of large losses when inventories are large and there is an unfavorable price movement. With this in mind, the goal of the (MM) is to maximize expected profits whilst managing adverse selection and inventory risks.

The standard approach to market making is to assume that the MM has perfect knowledge of the stochastic dynamics of the different state variables which are required to work out how to trade in and out of positions optimally. However, an additional source of risk is model risk. Clearly, any model is an approximation of reality and making 'optimal decisions' with the wrong framework will undoubtedly affect the profitability of market making activities. There is no such concept as the "correct" model and one way in which MMs can address this risk is to acknowledge that the model is misspecified. The main contribution of our work is to provide a framework where the MM deals securities based on a model that is robust to misspecification. The MM recognizes that she does not exactly know the probability laws of the stochastic processes required to maximize expected profits so she also considers other models when devising a profit maximizing strategy. This framework can also be applied to algorithmic trading scenarios other than market making, such as one sided optimal execution, pairs trading and other strategies that aim to profit from price predictions.

To understand how MMs account for ambiguity in their models we first consider the case when the MM is extremely confident about her choice of model and then we explicitly model misspecification and ambiguity aversion. As a starting point the MM uses a reference model where the probability law for the different state variables is known and the objective is to find the optimal market making strategy that maximizes expected profits

$$\sup_{\boldsymbol{\delta}\in\mathcal{A}} \mathbb{E}^{\mathbb{P}}\left[f(\boldsymbol{\theta})\right] , \tag{1}$$

where $\boldsymbol{\delta}$ are controls in the admissible set of strategies $\mathcal{A}$, $f(\boldsymbol{\theta})$ is the profit function which depends on the vector of state variables $\boldsymbol{\theta}$, and $\mathbb{E}^{\mathbb{P}}$ is the expectation operator under the reference measure $\mathbb{P}$ (corresponding to the reference model).

If the MM is not confident about $\mathbb{P}$ she will consider other candidate models $\mathbb{Q}$ (measures) and must specify how to choose amongst all alternatives. One method is to modify (1) by introducing a function $\mathcal{H}(\mathbb{Q}|\mathbb{P})$ which penalizes deviations from the reference model and captures the MM's degree of ambiguity aversion, to solve

$$\sup_{\boldsymbol{\delta}\in\mathcal{A}} \inf_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}\left[f(\boldsymbol{\theta}) + \mathcal{H}(\mathbb{Q}|\mathbb{P})\right] , \tag{2}$$

where $\mathcal{Q}$ is the set of all alternative models, see Uppal and Wang (2003), Maenhout (2004), Hansen et al. (2006), Hansen and Sargent (2007), Lim and Shanthikumar (2007), Hansen and Sargent (2011), Jaimungal and Sigloch (2012), and Skiadas (2013). Most of these prior works deal with portfolio optimization and/or consumption problems (Jaimungal and Sigloch (2012) is the exception where the authors investigate ambiguity in the context of credit risk) where the underlying sources of uncertainty are driven by Brownian motions (Lim and Shanthikumar (2007) and Skiadas (2013) are exceptions where uncertainty is driven by a Poisson process with ambiguity on its intensity). Here, the underlying sources of uncertainty stem not only from a Brownian motion, but also from Poisson random measures (PRMs) that drive arrival and execution price of market orders as well as the filling of posted limit orders.

The market making problem requires MMs to model how liquidity is provided and taken in modern electronic markets. In order driven markets all buyers and sellers provide liquidity by displaying the prices and quantities at which they wish to buy or sell a particular security. These orders are accumulated in the limit order book (LOB) until they are cancelled/amended or filled by an incoming liquidity taking market order (MO).

During a typical trading day in US markets, between 9:30 am and 4:00 pm, the LOBs for liquid stocks receive hundreds of thousands of messages with instructions to post, cancel, or amend limit orders (LOs) as well as thousands of buy and sell market orders (MOs). The shape and innovations in the LOB are important ingredients when making markets. A snapshot of the LOB allows the MM to gauge the fill probability (conditioned on an MO arriving) of LOs at different levels, and changes in the LOB convey information about how liquidity providers perceive the market.

Modelling the liquidity taking side of the market is also important. MOs may arrive in clusters, sometimes the market is one-sided (more buys than sells or vice versa) or two-sided, see Cartea et al. (2014). Along with the arrival of MOs one must also model the volumes and price impact of MOs, as well as the informational content of both LOs and MOs, see Bayraktar and Ludkovski (2011), Bayraktar and Ludkovski (2014), Gatheral et al. (2012), Schied (2013), Guéant and Lehalle (2015), and Cartea et al. (2015). Large MOs walk through the LOB. Sometimes the LOB replenishes quickly and the best bid and/or ask price reverts to the level prior to the arrival of the MO; however, at other times the best quotes do move to reflect changes in the fundamental value of the security, see Cartea and Jaimungal (2015). Thus, an MM who wishes to optimally trade in and out of positions must model the dynamics of the LOB and the arrival of buy and sell MOs.

In this paper we assume that the MM considers a reference model where: i) MOs arrive according to homogeneous Poisson processes, ii) minimum and maximum MO execution prices are exponentially distributed from the midprice, and iii) the midprice is a drifted Brownian motion. Moreover, the MM is ambiguous about her choices due to model misspecification and therefore considers other models where she incorporates a penalization which reduces to relative entropy in special cases in (2) to rank all possible alternatives. To the best of the authors' knowledge, this is the first work to account for ambiguity in optimal trading decisions in the context of algorithmic and high-frequency trading.

One of the advantages of our framework is that the MM is allowed to place different degrees of ambiguity aversion on the different building blocks of the model. Our results show that ambiguity aversion specific to the drift of the midprice changes the optimal postings in the book such that: (i) reversion to the optimal level of inventory is quicker, and (ii) total depth (buy depth plus sell depth) increases to help the MM recover the losses derived from trading with traders who possess superior information. A similar result is obtained when the MM is ambiguous specific to the rate of arrival of MOs. In this case, the MM behaves as if MOs arrive less often, a reduction that intensifies with larger inventories (long or short), and the effect is to accelerate mean reversion in inventories and to increase total depth.

Moreover, ambiguity specific to the execution price of MOs induces smaller depth on both sides of the book which has the effect of increasing the probability that LOs are filled. The intuition is that the MM is not sure that her model uses the correct distribution for the execution price of MOs and fears not being able to obtain enough business. Thus, her optimal behavior is to choose an alternative model where the probability of being filled is lower than that of the reference measure. This makes her more aggressive to attract business by posting quotes which are closer to the midprice, but also exposes the strategy to higher adverse selection costs.

So does robust market making help to improve the bottom line of MMs? To answer this question we use simulations to evaluate the performance of market making strategies when the MM is ambiguity averse. We assume a realistic model for the true dynamics where the midprice has a short-term alpha component, the shape of the book is stochastic and depends on the informational content and price impact of MOs, and the arrival rates of MOs follow mutually exciting processes to capture trade clustering, see Large (2007) and Cartea et al. (2014). Our results show that there are ranges where increasing ambiguity aversion produces a significant increase in the Sharpe ratio of the strategy. We also show that increases in the ambiguity aversion to midprice drift allows the MM to adjust the risk-return tradeoff by managing the exposure to inventory risk and adverse selection costs.

Another contribution of our work is to show that ambiguity aversion to midprice drift is equivalent to imposing a running penalty on inventories. To the best of our knowledge this result is also new in the literature. The work of Cartea and Jaimungal (2015) introduced the running penalty in an ad-hoc way and showed that the effect of this penalty is instrumental when MMs adjust their strategies to trade off risk and return, see also Guilbaud and Pham (2015).

The rest of this paper is organized as follows. In Section 2 we describe the reference model and show the MM's optimal market making strategy. In Section 3 we show how the MM makes her model robust to misspecification and prove a verification result. In Section 4 we show how the optimal strategies under the effects of ambiguity aversion compare to those of the reference model. Section 5 shows the financial benefits from robust market making. Section 6 concludes and in the Appendices we collect proofs and outline the numerical method we employ.

2. Reference Model

The profit maximization problem that the MM solves consists of deciding the level at which she sends limit buy and limit sell orders to the LOB, see for instance Avellaneda and Stoikov (2008), Guilbaud and Pham (2013), Guéant et al. (2012), and Fodra and Labadie (2012).[2] These liquidity providing quotes rest in the LOB until cancelled or filled by an incoming MO. Thus, to pose the market making problem the MM must specify the reference model for: A) midquote dynamics, B) MO arrival dynamics, and C) the interaction between incoming MOs and posted LOs. To this end, we work on a completed filtered probability space $(\Omega, \mathbb{F}, \mathcal{F} = \{\mathcal{F}_t\}_{0\le t\le T}, \mathbb{P})$ where $\mathcal{F}$ is the natural filtration generated by the midprice $(S_t)_{0\le t\le T}$ and the jump processes $(P_t^{\pm})_{0\le t\le T}$ (both of which are defined below).

A: **Reference Model: Midprice.** The midprice S_t satisfies

$$dS_t = \alpha\, dt + \sigma\, dW_t\,, \tag{3}$$

where $\alpha \in \mathbb{R}$ and $\sigma > 0$ are constants and $(W_t)_{0\le t\le T}$ is a standard Brownian motion.

B: **Reference Model: LOB.** On the supply side of the market, liquidity providing participants send quotes to the LOB where they specify the volume of shares they are willing to buy or sell and how far away from the midquote S_t these are posted. The MM's reference model is that the distance from the midprice to the maximum (for buys) or minimum (for sells) execution price of an incoming MO is exponentially distributed with parameters κ^- and κ^+, for the buy and sell side respectively. An LO is filled by an MO if the maximum (for market buys) or minimum (for market sells) execution price exceeds the price of the LO. Then, upon the arrival of an MO the probability that it fills an LO of one share which is y dollars away from the midquote $(S_t \pm y)$ is

$$p^{\pm}(y) = e^{-\kappa^{\pm} y}\,. \tag{4}$$

Consequently, the deeper the MM posts, the less likely it is that the LOs will be filled.

C: **Reference Model: MO Arrival.** On the demand side, liquidity taking participants send buy and sell MOs according to homogeneous Poisson processes M_t^+ and M_t^- with intensity parameters λ^+ and λ^- for the buy and sell side respectively. Note that a sell MO hits the buy side of the LOB and a buy MO lifts the offer on the sell side of the LOB.

To formalize the MM's problem we need to introduce more notation. Let $N_t^{\pm} = \int_0^t \int_{\delta_s^{\pm}}^{+\infty} \mu^{\pm}(dy, ds)$ denote the counting processes for filled LOs where $\mu^{\pm}(dy, ds)$ are independent PRMs (mutually

[2]For other trading algorithms that are designed to find the best execution prices for investors who wish to minimize the price impact of large buy or sell orders see Almgren (2003), Kharroubi and Pham (2010), Cartea and Jaimungal (2013), Bayraktar and Ludkovski (2014), Jaimungal and Kinzebulatov (2014), and Cartea et al. (2016).

independent of W) with compensators $\nu^{\pm}(dy, ds) = \lambda^{\pm}\, F^{\pm}(dy)\, ds$ where $F^{\pm}(dy) = \kappa^{\pm} e^{-\kappa^{\pm} y} dy$, and $\delta_t^{\pm}$ are predictable processes representing the depth relative to midprice of orders that the MM posts. The y variable corresponds to the distance from the midprice to the minimum or maximum execution price of incoming MOs, while the time dependence models their arrivals. Further, the number of MOs is given by $M_t^{\pm} = \int_0^t \int_0^{+\infty} \mu^{\pm}(dy, ds)$. Lastly, let $P_t^{\pm} = \int_0^t \int_0^{+\infty} y\mu^{\pm}(dy, ds)$.

As mentioned above, we work on the natural filtration generated by $(S_t)_{0 \le t \le T}$ and $(P_t^{\pm})_{0 \le t \le T}$. Observing the history of the processes $(P_t^{\pm})_{0 \le t \le T}$ means the agent knows the time and execution price of each market order and allows her to reconstruct the counting processes $(N_t^{\pm})_{0 \le t \le T}$ and $(M_t^{\pm})_{0 \le t \le T}$, which count the number of filled sell/buy LOs that the trader posts and the number of MOs that arrive in the market respectively.

Upon a buy or sell order being filled, the MM pays $S_t - \delta_t^-$ or receives $S_t + \delta_t^+$ and therefore the MM's wealth process $(X_t)_{0 \le t \le T}$ satisfies the SDE

$$dX_t = (S_t + \delta_t^+) dN_t^+ - (S_t - \delta_t^-) dN_t^- \,.$$

For instance, a buy MO arrives with intensity λ^+ and hits the sell side of the book and N_t^+ increases by one if the maximum execution price of the MO is greater than $S_t + \delta_t^+$. Therefore, the rate of execution of an LO which is posted $\delta^{\pm}$ away from the midprice is $\Lambda_t^{\pm} = \lambda^{\pm}\, e^{-\kappa^{\pm} \delta_t^{\pm}}$. Finally, the total inventory of the MM is given by $q_t = N_t^- - N_t^+$ and so $q_t \in \mathbb{Z}$ for each t.

We assume that the MM is risk-neutral but capital constrained so she cannot build large, long or short, inventory positions. Thus, the MM restricts her inventory so that $-\infty < \underline{q} \le q_t \le \overline{q} < +\infty$ for all $t \le T$. The MM seeks the strategy $(\delta_t^{\pm})_{0 \le t \le T}$ which maximizes expected terminal wealth

$$H(t, x, q, S) = \sup_{(\delta_s^{\pm})_{t \le s \le T} \in \mathcal{A}} \mathbb{E}^{\mathbb{P}}_{t,x,q,S} \left[X_T + q_T \left(S_T - \ell(q_T) \right) \right] \,, \tag{5}$$

where the terminal date of the strategy is $T > t$, $\mathbb{E}^{\mathbb{P}}_{t,x,q,S}[\,\cdot\,]$ denotes $\mathbb{P}$ expectation conditional on $X_{t^-} = x$, $q_{t^-} = q$ and $S_t = S$, and $\mathcal{A}$ denotes the set of admissible strategies which are non-negative $\mathcal{F}_t$-predictable processes such that inventories are bounded above by $\overline{q} > 0$ and below by $\underline{q} < 0$ both finite. If $q_t = \overline{q}$, then the MM places only limit sell orders (i.e., she sets $\delta^- = +\infty$) and if $q_t = \underline{q}$, then the MM places only limit buy orders (i.e., she sets $\delta^+ = +\infty$). Moreover, the function ℓ, with $\ell(0) = 0$ and $\ell(q)$ increasing in q, is a liquidation penalty which consists of fees and market impact costs when the MM sends an MO to unwind terminal inventory. For example $\ell(q) = \theta\, q$ represents a linear impact when liquidating q shares.

To solve the optimal control problem described above, we consider the associated Hamilton-Jacobi-Bellman (HJB) equation (see Fleming and Soner (2006) and Pham (2009)):

$$\partial_t H + \alpha \partial_S H + \tfrac{1}{2}\sigma^2 \partial_{SS} H + \sup_{\delta^+ \geq 0} \left\{ \lambda^+ e^{-\kappa^+ \delta^+} \, \Delta^+ H \right\} + \sup_{\delta^- \geq 0} \left\{ \lambda^- e^{-\kappa^- \delta^-} \, \Delta^- H \right\} = 0 \,, \tag{6}$$

where the operator $\Delta^{\pm}$ acts as follows

$$\Delta^{\pm} H(t,x,q,S) = H(t, x \pm (S \pm \delta^{\pm}), q \mp 1, S) - H(t,x,q,S), \tag{7}$$

and subject to the terminal condition

$$H(T,x,q,S) = x + q \Big(S - \ell(q) \Big) \,, \tag{8}$$

where x, S, and q are the quantities at t^- (and not t). The terminal condition is inherited from the formulation of the control problem (5). The boundary conditions in the x and S dimensions are that the value function is linear in each – indeed as we show below, the value function admits an ansatz which is explicitly linear in x and S in the entire domain. For the q direction, recall that the problem is set on the bounded domain $[\underline{q}, \overline{q}]$, the agent can only trade to move away from these boundary, and q is discrete. Hence, there is no need to specify the boundary conditions in q as the agent's optimization problem couples the value function along the boundary to its value in the interior.

Moreover, recall that the set of admissible strategies imposes bounds on q_t, this means that when $q_t = \overline{q}$ ($\underline{q}$) the MM posts one-sided LOs which are obtained by solving (6) with the term proportional to λ^- (λ^+) absent. Alternatively, one can view these boundary cases as imposing $\delta^- = +\infty$ ($\delta^+ = +\infty$) when $q = \overline{q}$ ($\underline{q}$).

Intuitively, the various terms in the HJB equation represent the arrival of MOs that may be filled by LOs together with the diffusion and drift of the asset price through the terms $\frac{1}{2}\sigma^2 \partial_{SS} H$ and $\alpha \partial_S H$. The supremum over δ^+ contain the terms due to the arrival of a market buy order, which is filled by a limit sell order with fill rate $\lambda^+ e^{-\kappa^+ \delta^+}$, and the change in the value function H. Similarly, the supremum over δ^- contain the analogous terms for sell MOs which are filled by buy LOs.

To solve the HJB equation we use the terminal condition (8) to make an ansatz for H. In particular, write

$$H(t,x,q,S) = x + q\, S + h_q(t) \,, \tag{9}$$

and upon substitution in (6) we obtain

$$\partial_t h_q + \alpha q + \sup_{\delta^+, \delta^- \geq 0} \left\{ \lambda^+ e^{-\kappa^+ \delta^+} \left(\delta^+ + h_{q-1} - h_q \right) + \lambda^- e^{-\kappa^- \delta^-} \left(\delta^- + h_{q+1} - h_q \right) \right\} = 0 \,, \tag{10}$$

with terminal condition $h_q(T) = -q\,\ell(q)$. This allows us to solve for the optimal feedback controls, in terms of $h_q(t)$, as shown in the proposition below. Here the subscript q in h_q denotes dependence on q. We delay a proof of existence and uniqueness to this equation until we have developed the control problem that incorporates ambiguity aversion. The solution to equation (10) is a special case of the solution to equation (23) which appears in Proposition 3. A verification that solutions to the HJB equation (6) yield the value function defined in (5) is also postponed to the more general setting when we include model ambiguity (see Theorem 4).

Proposition 1 (Optimal Feedback Controls). *The optimal feedback controls of the HJB equation are given by*

$$\begin{aligned}
\delta_q^{+*}(t) &= \left(\tfrac{1}{\kappa^+} - h_{q-1}(t) + h_q(t)\right)_+ , \quad q \neq \underline{q}\,, \\
\delta_q^{-*}(t) &= \left(\tfrac{1}{\kappa^-} - h_{q+1}(t) + h_q(t)\right)_+ , \quad q \neq \overline{q}\,.
\end{aligned}$$

where $(\cdot)_+ = \max(\cdot, 0)$.

Proof. Apply first order conditions to each supremum term in equation (10). If the resulting critical value is negative, it is an easy task to check that the maximum is achieved at $\delta^\pm = 0$. □

To understand the intuition behind the feedback controls we first note that when $\delta^{\pm*} > 0$ it can be decomposed into two terms. The first component, $1/\kappa^\pm$, is the optimal strategy that a risk-neutral MM would employ in the absence of both terminal date restrictions; i.e. $T = \infty$, and no inventory constraints. To see this, note that the expected gains from buying (selling) the asset at the midprice S_t, followed by selling (buying) it using an LO at $S_{t-} + \delta^+$ ($S_{t-} - \delta^-$), is given by $\delta^\pm e^{-\kappa^\pm \delta^\pm}$, which is maximized if the LO is posted $1/\kappa^\pm$ away from the midprice. The second term $-h_{q\pm1}(t) + h_q(t)$, controls for inventories through time. As expected, if inventories are long, then the strategy consists of posting LOs that increase the probability of sell orders being hit by posting closer to the midprice.

Figure 1 shows the optimal sell depth when $\kappa^\pm = 15$, $\lambda^\pm = 2$, $\sigma = 0.01$, $\alpha = 0$, $\ell(q) = \theta\, q$, $\theta = 0.01$, $\overline{q} = -\underline{q} = 3$ and $T = 10$ seconds when (10) is solved numerically after substituting the feedback controls in Proposition 1. We observe that when inventories increase (decrease), the sell depth decreases (increases) because the MM's strategy is to make round-trip trades to earn their total depth. The optimal buy depth is not shown due to the symmetry between the optimal buy and sell depths which is caused by $\lambda^+ = \lambda^-$, $\kappa^+ = \kappa^-$, and $\alpha = 0$ (this will be shown later in a more general setting, Proposition 6).

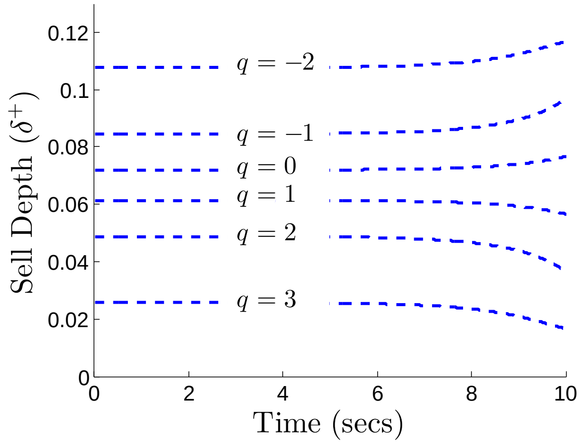


Figure 1: Optimal sell depth for an ambiguity neutral MM. Parameter values are $\kappa^{\pm} = 15$, $\lambda^{\pm} = 2$, $\sigma = 0.01$, $\alpha = 0$, $\ell(q) = \theta\, q$, $\theta = 0.01$, $\overline{q} = -\underline{q} = 3$ and $T = 10$ seconds.

3. Robust Modelling to Model Misspecification

Because the MM knows that her model is misspecified she considers alternative models of the midprice dynamics, fill probabilities, and MO arrival – specified under a candidate measure $\mathbb{Q}$ equivalent to the reference measure $\mathbb{P}$. The MM ranks the alternatives by evaluating all admissible strategies over a set of equivalent measures $\mathcal{Q}$ to choose the one which makes her model robust to misspecification (below we define the class of candidate measures $\mathcal{Q}$ more precisely). To this end, the MM introduces a penalty function in the optimization problem which measures the 'cost' of rejecting the reference measure $\mathbb{P}$ and accepting a candidate model $\mathbb{Q}$. For instance, if the MM is very confident about the reference model, any 'small' deviation from the measure $\mathbb{P}$ is heavily penalized, i.e. it is very costly to choose an alternative. On the other hand, if the MM is extremely ambiguous about her choice of the reference measure, considering other models will only result in a very small penalty.

In this way the MM chooses strategies which are robust to model misspecification by augmenting the optimization problem to maximize expected penalized terminal wealth over all admissible market making strategies, while minimizing expected penalized terminal wealth over a set of alternative models. Therefore, the optimization problem in (5) becomes

$$H(t,x,q,S) = \sup_{(\delta_s^{\pm})_{t\le s\le T}\in\mathcal{A}} \inf_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}_{t,x,q,S}\left[X_T + q_T\Big(S_T - \ell(q_T)\Big) + \mathcal{H}_{t,T}(\mathbb{Q}|\mathbb{P})\right], \tag{11}$$

where $\mathbb{E}^{\mathbb{Q}}_{t,x,q,S}[\cdot]$ denotes $\mathbb{Q}$ expectation conditional on $X_{t^-} = x$, $q_{t^-} = q$ and $S_t = S$, $\mathcal{H}_{t,T}(\mathbb{Q}|\mathbb{P})$ is the penalization function introduced by the MM, and the class of measures $\mathcal{Q}$ reflects all alternative models that the MM considers.

A popular choice for the penalty function is the entropic penalty function

$$\mathcal{H}_{t,T}(\mathbb{Q}|\mathbb{P}) = \frac{1}{\psi} \log \left\{ \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right)_T \Big/ \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right)_t \right\} , \tag{12}$$

where $\psi > 0$ is a constant that shows how confident the MM is about her reference model. If the MM is extremely confident about the reference model then ψ is very small. In the limiting case $\psi \to 0$ the MM is ambiguity neutral and therefore rejects any alternative model. If the MM is extremely ambiguous about the reference model then ψ is very large – in the extreme case $\psi \to \infty$ the MM considers the worst case scenario when making markets.

Using the above entropic penalty function forces the MM to have the same level of ambiguity aversion towards each aspect of the model. But it is possible that the MM has different levels of confidence towards different aspects of the model, and we introduce a more general penalty function motivated by entropic penalization which allows the MM to have different levels of ambiguity aversion stemming from uncertainty in the model with respect to three factors: (i) drift, (ii) fill probabilities / market order execution price, and (iii) arrival rate of market orders. Note that the idea of splitting ambiguity into components and combining them with varying weights is similar in spirit to Uppal and Wang (2003) who combine ambiguity from several sub-sets of asset classes in a diffusive setting. It is also similar to source-dependent risk aversion in which different risk aversion parameters are associated with each source of risk, such as in Hugonnier et al. (2013). Here, however, a direct ambiguity aversion decomposition will be carried out on PRMs.

3.1. Measure Class and Decomposition

This section will introduce the full class of measures $\mathcal{Q}$ and an approach for decomposing the measure change to account for varying levels of ambiguity on different aspects of the model. Define two Radon-Nikodym derivative processes by

$$\frac{d\mathbb{Q}^{\alpha}(\eta)}{d\mathbb{P}} = \exp\left\{ -\frac{1}{2} \int_0^T \left(\frac{\alpha - \eta_t}{\sigma} \right)^2 dt - \int_0^T \frac{\alpha - \eta_t}{\sigma}\, dW_t \right\} , \tag{13}$$

and

$$\frac{d\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)}{d\mathbb{Q}^{\alpha}(\eta)} = \prod_{i=\pm} \exp\left\{ -\int_0^T \int_0^\infty (e^{g_t^i(y)} - 1)\, \nu^i(dy, dt) + \int_0^T \int_0^\infty g_t^i(y)\, \mu^i(dy, dt) \right\} . \tag{14}$$

Next, define a Radon-Nikodym derivative by their product

$$\frac{d\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)}{d\mathbb{P}} = \frac{d\mathbb{Q}^{\alpha}(\eta)}{d\mathbb{P}} \frac{d\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)}{d\mathbb{Q}^{\alpha}(\eta)} . \tag{15}$$

The superscript α, λ, κ is to indicate that the drift, arrival rate, and fill probability all become different in changing from the reference measure $\mathbb{P}$ to the new measure $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)$. The full class of candidate measures to be considered by the MM is

$$\mathcal{Q}^{\alpha,\lambda,\kappa} = \Bigg\{ \mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g) \,:\, \eta,\, g \text{ are } \mathcal{F}\text{-predictable, } \mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}^{\alpha}(\eta)}{d\mathbb{P}}\right] = 1,$$
$$\mathbb{E}^{\mathbb{Q}^{\alpha}(\eta)}\left[\frac{d\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)}{d\mathbb{Q}^{\alpha}(\eta)}\right] = 1, \text{and } \mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\int_0^T \int_0^\infty y^2 e^{g_t^{\pm}(y)} \nu^{\pm}(dy, dt)\right] < \infty \Bigg\}. \tag{16}$$

The constraints imposed on the first two expectations above guarantee that the Radon-Nikodym derivatives as defined in (13) and (14) yield probability measures. The constraint on the $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)$ expectation is to ensure that the variance of the profits earned by the MM through filled orders is finite in any candidate measure. In the measure $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)$ the drift of the midprice is no longer the constant α, but is changed to η_t. Similarly, the compensator of $\mu^{\pm}(dy, dt)$ is $\nu_{\mathbb{Q}}^{\pm}(dy, dt) = e^{g_t^{\pm}(y)} \nu^{\pm}(dy, dt) = e^{g_t^{\pm}(y)} \lambda^{\pm} F^{\pm}(dy) dt$, see Jacod and Shiryaev (1987), Chapter III.3c, Theorem 3.17. Quantities that will also be of interest are the intensity of MO arrivals and the fill probabilities of LOs in the candidate measure. The intensity of MO arrivals is given by

$$\lambda_t^{\pm\mathbb{Q}} = \lambda^{\pm} \int_0^\infty e^{g_t^{\pm}(y)} F^{\pm}(dy) \,, \tag{17}$$

and the probability that an MO fills an LO posted at a price $S_t \pm y$ under the candidate measure is

$$p_t^{\pm}(y) = \frac{\int_y^\infty e^{g_t^{\pm}(y')}\, F^{\pm}(dy')}{\int_0^\infty e^{g_t^{\pm}(y')} F^{\pm}(dy')} \,. \tag{18}$$

Note that if $g_t^{\pm} = 0$ then the MO and LO dynamics of the reference model are retrieved. Also, if $g_t^{\pm}(y)$ does not depend on y, then only the MO intensity is changed while fill probabilities remain as they are in the reference model. And lastly, if $g_t^{\pm}(y)$ is chosen so that

$$\int_0^\infty e^{g_t^{\pm}(y)} F^{\pm}(dy) = 1, \tag{19}$$

for all $t \in [0, T]$, then the MO intensity in the candidate measure will remain constant at $\lambda^{\pm}$ while the fill probabilities change to those in (18).

The form of the penalty function we introduce relies on decomposing the measure change induced by the Radon-Nikodym derivative in (14) into two separate measure changes. To this end, given the random field $g_t^{\pm}(y)$, introduce two new random fields defined by

$$g_t^{\pm\lambda} = \log\left(\int_0^\infty e^{g_t^{\pm}(y)}\, F^{\pm}(dy) \right) \quad \text{and} \quad g_t^{\pm\kappa}(y) = g_t^{\pm}(y) - g_t^{\pm\lambda} \,. \tag{20}$$

The first random field does not depend on y, and it is an easy task to check that $g_t^{\pm\kappa}(y)$ satisfies the condition in (19). Thus, any measure $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g) \in \mathcal{Q}^{\alpha,\lambda,\kappa}$ allows us to uniquely define two measures $\mathbb{Q}^{\alpha,\lambda}(\eta, g)$ and $\mathbb{Q}^{\alpha,\kappa}(\eta, g)$ via the Radon-Nikodym derivatives:

$$\frac{d\mathbb{Q}^{\alpha,\lambda}(\eta,g)}{d\mathbb{Q}^{\alpha}(\eta)} = \prod_{i=\pm} \exp\left\{ -\int_0^T \int_0^\infty (e^{g_t^{i\lambda}} - 1)\, \nu^i(dy,dt) + \int_0^T \int_0^\infty g_t^{i\lambda}\, \mu^i(dy,dt) \right\} ,$$

$$\frac{d\mathbb{Q}^{\alpha,\kappa}(\eta,g)}{d\mathbb{Q}^{\alpha}(\eta)} = \prod_{i=\pm} \exp\left\{ -\int_0^T \int_0^\infty (e^{g_t^{i\kappa}(y)} - 1)\, \nu^i(dy,dt) + \int_0^T \int_0^\infty g_t^{i\kappa}(y)\, \mu^i(dy,dt) \right\} .$$

These measure changes represent decomposing the full change from $\mathbb{Q}^\alpha(\eta)$ to $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)$ into a sequence of two changes, one in which only MO intensity is changed, and the other in which only fill probability is changed. This can be represented graphically as in Figure 2.

$$\begin{tikzcd}
 & & \mathbb{Q}^{\alpha,\lambda} \arrow[dr, "g^{\pm\kappa}"] & \\
\mathbb{P} \arrow[r, "\eta"] & \mathbb{Q}^{\alpha} \arrow[ur, "g^{\pm\lambda}"] \arrow[dr, "g^{\pm\kappa}"'] \arrow[rr, dotted, "g^{\pm}"] & & \mathbb{Q}^{\alpha,\lambda,\kappa} \\
 & & \mathbb{Q}^{\alpha,\kappa} \arrow[ur, "g^{\pm\lambda}"'] &
\end{tikzcd}$$

Figure 2: The three natural alternative routes from the reference measure $\mathbb{P}$ to a candidate measure $\mathbb{Q}^{\alpha,\lambda,\kappa}$ in which midprice drift, MO intensity, and execution price distribution of MO have been altered.

Henceforth it will be convenient to think of $g_t^{\pm\lambda}$ and $g_t^{\pm\kappa}(y)$ as separately defined objects. If $g_t^{\pm\lambda}$ does not depend on y and $g_t^{\pm\kappa}(y)$ satisfies the integral condition (19), then by defining the random field $g_t^{\pm}(y) = g_t^{\pm\lambda} + g_t^{\pm\kappa}(y)$, it is easily checked that the relations in (20) hold and there is a one to one correspondence between arbitrary random fields $g_t^{\pm}(y)$ and pairs of random fields $(g_t^{\pm\lambda}, g_t^{\pm\kappa}(y))$. This equivalent viewpoint will assist in the computation of the optimal candidate measure in (11).

3.2. Penalty Function

Using relative entropy as the penalty function implies a graphical representation shown in Figure 3.

$$\mathbb{P} \xrightarrow{\quad\quad\psi\quad\quad} \mathbb{Q}^{\alpha,\lambda,\kappa}$$

Figure 3: The penalization implied by relative entropy with ambiguity aversion parameter ψ.

$$\begin{tikzcd}
 & & \mathbb{Q}^{\alpha,\lambda} \arrow[dr, "\varphi_\kappa"] & \\
\mathbb{P} \arrow[r, "\varphi_\alpha"] & \mathbb{Q}^{\alpha} \arrow[ur, "\varphi_\lambda"] \arrow[dr, "\varphi_\kappa"'] & & \mathbb{Q}^{\alpha,\lambda,\kappa} \\
 & & \mathbb{Q}^{\alpha,\kappa} \arrow[ur, "\varphi_\lambda"'] &
\end{tikzcd}$$

Figure 4: Ambiguity weights associated with each sequential step of the full measure change.

We propose instead a penalty function which corresponds to the graphical representation shown in Figure 4. The interpretation of Figure 4 is that the MM has different levels of ambiguity towards misspecification of the factor of the model associated with each measure change. Thus, φ_α represents the level of aversion to misspecification of the drift, φ_λ represents aversion to misspecification of MO arrivals, and φ_κ represents aversion to misspecification of MO execution price or fill probability. Furthermore, the penalty function will be defined in such a way that the following properties hold:

1. The expectation of the penalization term is non-negative.
2. If $\varphi_\alpha = \varphi_\lambda = \varphi_\kappa = \varphi$, then the penalization term is equivalent to relative entropy with an ambiguity aversion level of φ.
3. Let x be one of the labels α, λ, or κ. When φ_x is equal to zero, then the optimal candidate measure $\mathbb{Q}^{*\alpha,\lambda,\kappa}$ which results from the optimization in (11) has dynamics associated with x that are identical to those of the reference measure $\mathbb{P}$. This holds for any combination of φ_x's that are equal to zero.[3]

The penalty function we choose is given by

$$\mathcal{H}\left(\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)\,|\,\mathbb{P}\right) = \begin{cases} \frac{1}{\varphi_\alpha}\log\left(\frac{d\mathbb{Q}^{\alpha}(\eta)}{d\mathbb{P}}\right) + \frac{1}{\varphi_\lambda}\log\left(\frac{d\mathbb{Q}^{\alpha,\lambda}(\eta,g)}{d\mathbb{Q}^{\alpha}(\eta)}\right) + \frac{1}{\varphi_\kappa}\log\left(\frac{d\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}{d\mathbb{Q}^{\alpha,\lambda}(\eta,g)}\right), & \text{if } \varphi_\lambda \geq \varphi_\kappa\,, \\ \frac{1}{\varphi_\alpha}\log\left(\frac{d\mathbb{Q}^{\alpha}(\eta)}{d\mathbb{P}}\right) + \frac{1}{\varphi_\kappa}\log\left(\frac{d\mathbb{Q}^{\alpha,\kappa}(\eta,g)}{d\mathbb{Q}^{\alpha}(\eta)}\right) + \frac{1}{\varphi_\lambda}\log\left(\frac{d\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}{d\mathbb{Q}^{\alpha,\kappa}(\eta,g)}\right), & \text{if } \varphi_\lambda \leq \varphi_\kappa\,. \end{cases} \tag{21}$$

We allow the ambiguity parameters to be equal to zero by using the convention $0 \cdot \infty = 0$. Note that when $\varphi_\lambda = \varphi_\kappa$, both of the expressions in (21) are equal. The form of (21) implicitly defines which route is taken from $\mathbb{P}$ to $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)$ in Figure 2. If $\varphi_\lambda > \varphi_\kappa$ then the sequence of measure

[3]For example, if $\varphi_\alpha = 0$, then the optimal measure will have the same drift as the reference model. If $\varphi_\alpha = \varphi_\kappa = 0$, then the optimal measure will have the same drift and fill probabilities as the reference model.

changes is performed such that the MO intensity changes first, and then the fill probability changes second. This makes $\mathbb{Q}^{\alpha,\lambda}(\eta, g)$ the intermediate measure that is visited. The changes are performed in the opposite order if $\varphi_\kappa > \varphi_\lambda$. If $\varphi_\lambda = \varphi_\kappa$, the two routes will contribute the same penalty so either one can be considered, or the direct route from $\mathbb{Q}^{\alpha}(\eta)$ to $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)$ can be taken.

Furthermore, note that the $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)$ expectation of the penalty function can be written as:

$$
\begin{aligned}
\mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\mathcal{H}\left(\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)\,|\,\mathbb{P}\right)\right] = \quad & \frac{1}{\varphi_\alpha}\mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\frac{1}{2}\int_0^T\left(\frac{\alpha-\eta_t}{\sigma}\right)^2 dt\right] \\
& + \sum_{i=\pm}\left\{\mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\int_0^T \mathcal{K}_i^{\varphi_\lambda,\varphi_\kappa}(g_t^{i\lambda}, g_t^{i\kappa})\,dt\right]\mathbb{1}_{\varphi_\lambda\geq\varphi_\kappa}\right. \\
& \left. + \,\mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\int_0^T \mathcal{K}_i^{\varphi_\kappa,\varphi_\lambda}(g_t^{i\kappa}, g_t^{i\lambda})\,dt\right]\mathbb{1}_{\varphi_\lambda<\varphi_\kappa}\right\},
\end{aligned}
$$

where, for the collection of functions $\{a(y),\, b(y)\}$ and constants $\{c, d\}$,

$$
\begin{aligned}
\mathcal{K}_i^{c,d}(a,b) = \quad & \frac{1}{c}\int_0^\infty\left[-(e^{a(y)}-1) + a(y)\,e^{a(y)+b(y)}\right]\lambda^i\,F^i(dy) \\
& + \frac{1}{d}\int_0^\infty\left[-(e^{b(y)}-1)\,e^{a(y)} + b(y)\,e^{a(y)+b(y)}\right]\lambda^i\,F^i(dy)\,.
\end{aligned}
$$

It is important to point out that the penalty (21) is not a relative entropy since each component in (21) measures log-distances between various measures yet the optimization problem is given by an expectation under the single measure $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)$.

Proposition 2 (Properties of the penalty function). *The map from $\mathcal{Q}^{\alpha,\lambda,\kappa} \to \mathbb{R}$ given by $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g) \mapsto \mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}}\left[\mathcal{H}\left(\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)\,|\,\mathbb{P}\right)\right]$ satisfies properties 1 and 2 listed above.*

Proof. See Appendix A.

Property 3 is not contained in Proposition 2 because its validity depends on performing the optimization in (11). However, it will be seen that it does indeed hold when the optimization problem is solved, and an intuitive explanation for why it should be true based on the form of (21) can be given. Consider for example if $\varphi_\alpha = 0$. Then any deviations of $\mathbb{Q}^\alpha(\eta)$ from $\mathbb{P}$ will result in the quantity $\mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}}[\log \frac{d\mathbb{Q}^\alpha(\eta)}{d\mathbb{P}}]$ being positive. Thus, unless a small change in the drift causes the expected terminal value of the MM's wealth and inventory holdings to approach $-\infty$, the infimum in (11) would require that $\mathbb{Q}^\alpha(\eta) = \mathbb{P}$, meaning the drift in the optimal measure $\mathbb{Q}^{*\alpha,\lambda,\kappa}$ will still be the constant α. Similar reasoning holds for $\varphi_\lambda = 0$ and $\varphi_\kappa = 0$.

3.3. Solving for the Value Function

With a well defined class of candidate measures and penalty function, we are now able to return to and solve the optimization problem (11). The associated HJBI equation is

$$
\begin{aligned}
&\partial_t H + \frac{1}{2}\sigma^2 \partial_{SS} H + \inf_{\eta} \left\{ \eta\, \partial_S H + \frac{1}{2\varphi_\alpha}\left(\frac{\alpha-\eta}{\sigma}\right)^2 \right\} \\
&\quad + \sum_{i=\pm} \sup_{\delta^i \geq 0} \inf_{g^{i\lambda}} \inf_{g^{i\kappa}\in\mathcal{G}^i} \Bigg\{ \lambda^i \left[\int_{\delta^i}^{\infty} e^{g^{i\lambda}+g^{i\kappa}(y)} F^i(dy) \right] \Delta^i H \\
&\quad + \mathcal{K}_i^{\varphi_\lambda,\varphi_\kappa}(g^{i\lambda}, g^{i\kappa})\, \mathbb{1}_{\varphi_\lambda \geq \varphi_\kappa} + \mathcal{K}_i^{\varphi_\kappa,\varphi_\lambda}(g^{i\kappa}, g^{i\lambda})\, \mathbb{1}_{\varphi_\lambda < \varphi_\kappa} \Bigg\} = 0\,,
\end{aligned}
\tag{22}
$$

subject to the terminal condition $H(T, x, q, S) = x + q(S - \ell(q))$, and where the set $\mathcal{G}^i$ is defined as:

$$
\mathcal{G}^i = \left\{ g : \int_0^{\infty} e^{g(y)} F^i(dy) = 1 \right\}.
$$

Proposition 3 (Solution to HJBI Equation). *The HJBI equation admits the ansatz $H(t, x, q, S) = x + q\,S + h_q(t)$, where $h_q(t)$ satisfies*

$$
\begin{aligned}
&\partial_t h_q + \inf_{\eta} \left\{ \eta q + \frac{1}{2\varphi_\alpha}\left(\frac{\alpha-\eta}{\sigma}\right)^2 \right\} \\
&\quad + \sup_{\delta^+ \geq 0} \inf_{g^{+\lambda}} \inf_{g^{+\kappa}\in\mathcal{G}^+} \Bigg\{ \lambda^+ \left[\int_{\delta^+}^{\infty} e^{g^{+\lambda}+g^{+\kappa}(y)} F^+(dy) \right] (\delta^+ + h_{q-1}(t) - h_q(t)) \\
&\qquad\qquad + \mathcal{K}_+^{\varphi_\lambda,\varphi_\kappa}(g^{+\lambda}, g^{+\kappa})\, \mathbb{1}_{\varphi_\lambda \geq \varphi_\kappa} + \mathcal{K}_+^{\varphi_\kappa,\varphi_\lambda}(g^{+\kappa}, g^{+\lambda})\, \mathbb{1}_{\varphi_\lambda < \varphi_\kappa} \Bigg\} \\
&\quad + \sup_{\delta^- \geq 0} \inf_{g^{-\lambda}} \inf_{g^{-\kappa}\in\mathcal{G}^-} \Bigg\{ \lambda^- \left[\int_{\delta^-}^{\infty} e^{g^{-\lambda}+g^{-\kappa}(y)} F^-(dy) \right] (\delta^- + h_{q+1}(t) - h_q(t)) \\
&\qquad\qquad + \mathcal{K}_-^{\varphi_\lambda,\varphi_\kappa}(g^{-\lambda}, g^{-\kappa})\, \mathbb{1}_{\varphi_\lambda \geq \varphi_\kappa} + \mathcal{K}_-^{\varphi_\kappa,\varphi_\lambda}(g^{-\kappa}, g^{-\lambda})\, \mathbb{1}_{\varphi_\lambda < \varphi_\kappa} \Bigg\} = 0\,,
\end{aligned}
\tag{23}
$$

subject to terminal conditions $h_q(T) = -q\,\ell(q)$. Moreover, the optimum in (23) *is achieved, where* *the optimizers are given by*

$$
\begin{aligned}
\delta_q^{+*}(t) &= \left(\tfrac{1}{\varphi_\kappa}\log\left(1+\tfrac{\varphi_\kappa}{\kappa^+}\right) - h_{q-1}(t) + h_q(t)\right)_+\,, \qquad q \neq \underline{q}\,,\\
\delta_q^{-*}(t) &= \left(\tfrac{1}{\varphi_\kappa}\log\left(1+\tfrac{\varphi_\kappa}{\kappa^-}\right) - h_{q+1}(t) + h_q(t)\right)_+\,, \qquad q \neq \overline{q}\,,\\
\eta_q^*(t) &= \alpha - \varphi_\alpha\,\sigma^2\,q\,, \qquad\qquad (24)\\
g_q^{\pm\lambda*}(t) &= \frac{\varphi_\lambda}{\varphi_\kappa}\log\left(1 - e^{-\kappa^\pm\delta_q^{\pm*}(t)}\left(1 - e^{-\varphi_\kappa(\delta_q^{\pm*}(t)+h_{q\mp1}(t)-h_q(t))}\right)\right),\\
g_q^{\pm\kappa*}(t,y) &= -\log\left(1 - e^{-\kappa^\pm\delta_q^{\pm*}(t)}\left(1 - e^{-\varphi_\kappa(\delta_q^{\pm*}(t)+h_{q\mp1}(t)-h_q(t))}\right)\right)\\
&\quad - \varphi_\kappa\left(\delta_q^{\pm*}(t) + h_{q\mp1}(t) - h_q(t)\right)\mathbb{1}_{y\geq\delta_q^{\pm*}(t)}\,.
\end{aligned}
$$

Furthermore, equation (23) together with its terminal conditions has as unique classical solution.

Proof. See Appendix A.

The expressions in (24) are those for which the sup inf in (23) achieve their optimal value pointwise. Given that we have a unique classical solution to (23), the function H serves as a candidate optimal solution to the control problem (11), and when the process q_t is substituted for the state variable q in (24), these serve as candidate optimal controls. Once we provide a verification theorem that the candidate H is indeed the value function defined in (11), we will have shown that the candidate optimal controls are indeed the ones which achieve the optimum for the control problem. This is the task we proceed to next.

Theorem 4 (Verification Theorem). *Let $h_q(t)$ be the solution to (23) and let $H(t,x,q,S) = x + q\,S + h_q(t)$. Also let $\delta_t^{\pm\diamond} = \delta_{q_t}^{\pm*}$, $\eta_t^\diamond = \eta_{q_t}^*(t)$, $g_t^{\pm\lambda\diamond} = g_{q_t}^{\pm\lambda*}(t)$, and $g_t^{\pm\kappa\diamond}(y) = g_{q_t}^{\pm\kappa*}(t,y)$ define processes. Then $\delta^{\pm\diamond}$, $\eta^\diamond$, $g^{\pm\lambda\diamond}$, and $g^{\pm\kappa\diamond}$ are admissible controls. Further, H is the value function to the agent's control problem (11) and the optimum is achieved by these controls.*

Proof. See Appendix A.

By scrutinizing the dependence of the optimal controls η^*, $g^{\pm\lambda*}$ and $g^{\pm\kappa*}(y)$ on φ_α, φ_λ, and φ_κ, we see that property 3 of the penalty function is indeed satisfied. If $\varphi_\alpha = 0$, then the drift in the optimal candidate measure is α, the same as the reference measure. If $\varphi_\lambda = 0$, then $g_q^{\pm\lambda*}(t) = 0$ implying that the MO intensities remain at the constants $\lambda^\pm$. Lastly, if $\varphi_\kappa = 0$, then $g_q^{\pm\kappa*}(t,y) = 0$ implying that the MO execution price distribution remains the same as in the reference measure (as well note that in the limit $\varphi_\kappa \to 0$, $g_q^{\pm\lambda*}(t)$ has a finite value).

With the above discussion in mind, we see that the value function and optimal controls are well defined for all finite φ_α, φ_λ, and φ_κ. However, we cannot explicitly allow infinite values of these parameters. See for example the expression for $\eta_q^*(t)$ in (24), in which the optimal drift would be equal to $\pm\infty$ depending on the sign of q. This can be alleviated by placing large finite bounds on η and $g^\pm$ which would allow infinite values of ambiguity parameters. The exact form of the optimal controls in (24) would be different in this case, but qualitative behaviour of the optimal spreads would be similar.

The following proposition provides a closed-form solution for the value function, and hence also the optimal depths at which the agent posts, under suitable symmetry conditions between the fill rate function and the ambiguity weights of MO and LO dynamics through the fill probabilities. The equal ambiguity weights assumed in this proposition correspond to the case where both versions of the penalty function in (21) are equivalent, and the route of the measure change decomposition from Figure 2, therefore, is irrelevant.

Proposition 5 (Closed-form solution). *If $\kappa^\pm = \kappa$ and $\varphi_\lambda = \varphi_\kappa = \varphi$, then write $\omega_q(t) = e^{\kappa h_q(t)}$. Define the vector $\boldsymbol{\omega}(t) = [\omega_{\overline{q}}(t), \omega_{\overline{q}-1}(t), \ldots, \omega_{\underline{q}}(t)]'$. Also, let $\mathbf{A} = \mathbf{diag}[\xi^+, \xi^-, \alpha\kappa\, q - \frac{1}{2}\kappa\varphi_\alpha\sigma^2\, q^2]$, where $\xi^\pm = \left(1 + \frac{\varphi}{\kappa}\right)^{-\left(1+\frac{\kappa}{\varphi}\right)}\lambda^\pm$. Then,*

$$\boldsymbol{\omega}(t) \quad = \quad e^{\mathbf{A}(T-t)}\boldsymbol{\omega}(T)$$

for all $t \in [t_0, T]$, where $\omega_q(T) = e^{-\kappa q\,\ell(q)}$ and $t_0 = \sup\{t : \exists q\ \delta_q^{\pm*}(t) = 0\} \vee 0$.

Proof. See Appendix A.

The following proposition justifies why, when market dynamics are symmetric, we focus solely on the sell side of the optimal controls in discussions surrounding the intuition of the effects of ambiguity.

Proposition 6 (Symmetry in depths). *If $\alpha = 0$, $\lambda^\pm = \lambda$, $\kappa^\pm = \kappa$, $\overline{q} = -\underline{q}$, and $q\,\ell(q)$ is an even function, then $\delta_q^{+*}(t) = \delta_{-q}^{-*}(t)$.*

Proof. See Appendix A.

4. The Effects of Ambiguity Aversion on Depth

At this point the effects of each ambiguity level parameter on the optimal depth of the MM are investigated. First, each of the three parameters are considered separately by setting the other two to zero. This allows us to analyze the effects of each type of ambiguity individually. Then we consider cases where all three parameters are non-zero. In Appendix B we outline the numerical method used to produce the figures in this section.

4.1. Robust modelling to misspecification of midprice dynamics

The MM recognizes that she does not have enough data to estimate the drift of the midprice and/or does not possess the right technology to process information fast enough to use a realistic and sophisticated model of the drift. In addition, the MM knows that she will be trading with other market participants that do have the technology and who profit from trading with less informed market participants – for example high frequency traders.

In Figure 5 ($\varphi_\alpha = 20$, $\varphi_\lambda = \varphi_\kappa = 0$) solid lines show the optimal sell depths and total depths (sell plus buy) for an MM who is ambiguous about the drift of the midprice and dashed lines show the sell and total depths for the MM who is extremely confident about the reference model; i.e. is ambiguity neutral (as in Figure 1). The picture on the left-hand side shows that as ambiguity on drift increases, the depths move in a direction which induces faster mean reversion to the optimal level of inventory. When inventory is positive (negative), the ambiguity averse MM posts smaller (larger) sell depths than the ambiguity neutral MM and by symmetry we have a similar effect on the buy depth. Thus, compared to an ambiguity neutral MM, mean reversion to the optimal level of inventory is faster. In Proposition 7 below we prove that, under symmetric dynamics, the agent will always alter her spreads in this manner when increasing ambiguity on drift. Moreover, the picture on the right-hand side shows that the total depth increases with φ_α which helps the MM to recover, on average, losses from trading with better informed market participants. In other words, ambiguity to midprice drift creates a buffer to protect the MM against adverse selection losses.

Ambiguity on drift has important effects on how the MM manages exposure to inventory risk. We observe that the MM picks a model where the midprice drifts at $\eta^*_q(\varphi_\alpha) = \alpha - \varphi_\alpha\, \sigma^2\, q$ which clearly depends on the level of inventory and the volatility of the midprice. Note that when $q > 0$ the MM assumes that (compared to the reference model) the midprice is drifting at a lower speed, $\eta^*_q(\varphi_\alpha) < \alpha$, and the effect is to adjust the postings in the LOB so that inventories are reduced. The larger the inventory is, the closer to the midprice are sell LOs posted and mean reversion to the optimal level of inventory is quicker. The intuition is the following. Assume for simplicity that $\alpha = 0$, thus when the MM is long the asset her model assumes that the midprice will drift down and

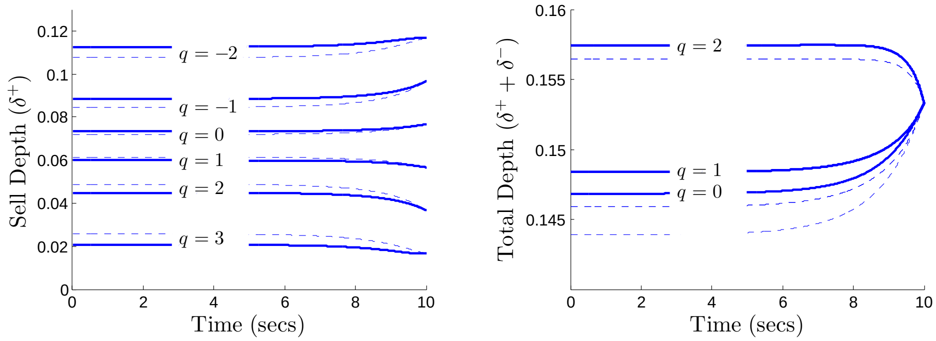


Figure 5: Optimal sell and total depths for an MM who is ambiguity averse to midprice drift (dashed lines are ambiguity neutral depths). Parameter values are $\varphi_\alpha = 20$, $\kappa^\pm = 15$, $\lambda^\pm = 2$, $\sigma = 0.01$, $\alpha = 0$, $\ell(q) = \theta\, q$, $\theta = 0.01$, $\overline{q} = -\underline{q} = 3$ and $T = 10$ seconds.

it is clear that holding long inventory in a falling market is not optimal so the strategy is to close down the positions as soon as possible – the MM will perceive a falling market until her inventory reaches $q = 0$. How quickly are positions closed depends also on the degree of ambiguity aversion and the volatility of the midprice. A similar argument applies when $q < 0$.

Another important point we remark is that only when the MM is ambiguity averse to midprice drift does the volatility of the midprice have an effect on the optimal strategy (σ^2 only appears in Proposition 3 when multiplied by φ_α). Here we see that the MM perceives more exposure to inventory risk when σ is higher because the chances of observing unfavourable price movements are higher.

Proposition 7 (Effect of Ambiguity on Drift). *Suppose $\alpha = 0$, $\lambda^\pm = \lambda$, $\kappa^\pm = \kappa$, $\overline{q} = -\underline{q}$, and $q\,\ell(q)$ is an even function. Fix φ_λ and φ_κ, and let $\varphi_\alpha < \varphi'_\alpha$. Denote by $\delta_q^{\pm*}(t;\varphi_\alpha)$ and $\delta_q^{\pm*}(t;\varphi'_\alpha)$ the spreads corresponding to the feedback forms given in (24) for the respective parameters. Then the sell spreads satisfy*

$$\begin{aligned}
\delta_q^{+*}(t;\varphi_\alpha) &\le \delta_q^{+*}(t;\varphi'_\alpha)\,, \quad \underline{q} < q \le 0\,,\\
\delta_q^{+*}(t;\varphi_\alpha) &\ge \delta_q^{+*}(t;\varphi'_\alpha)\,, \quad 0 < q \le \overline{q}\,,
\end{aligned}$$

and the buy spreads satisfy

$$\begin{aligned}
\delta_q^{-*}(t;\varphi_\alpha) &\le \delta_q^{-*}(t;\varphi'_\alpha)\,, \quad 0 \le q < \overline{q}\,,\\
\delta_q^{-*}(t;\varphi_\alpha) &\ge \delta_q^{-*}(t;\varphi'_\alpha)\,, \quad \underline{q} \le q < 0\,.
\end{aligned}$$

PROOF. See Appendix A.

4.1.1. Equivalence of ambiguity on midprice drift and inventory penalization

Consider an MM who maximizes expected terminal wealth, but rather than penalizing the payoff with a relative entropy term she directly penalizes her running inventory position, as first proposed in Cartea and Jaimungal (2015) and Guilbaud and Pham (2013). The value function is

$$H^{\phi}(t,x,q,S) = \sup_{(\delta_s^{\pm})_{t\le s\le T}\in\mathcal{A}} \mathbb{E}^{\mathbb{P}}_{t,x,q,S}\left[X_T + q_T\left(S_T - \ell(q_T)\right) - \phi\,\sigma^2\int_t^T q_s^2\,ds\right],$$

where $\phi \geq 0$ and $\phi\,\sigma^2\int_t^T q_s^2\,ds$ acts as a penalization on running inventory. The value function satisfies the HJB equation

$$\partial_t H^{\phi} + \alpha\partial_S H^{\phi} + \frac{1}{2}\sigma^2\partial_{SS}H^{\phi} - \phi\sigma^2 q^2 + \sup_{\delta^{\pm}\ge 0}\left\{\lambda^+ e^{-\kappa^+\delta^+}\Delta^+ H^{\phi} + \lambda^- e^{-\kappa^-\delta^-}\Delta^- H^{\phi}\right\} = 0\,,$$

subject to $H^{\phi}(T,x,q,S) = x + q\,(S-\ell(q))$. By making the ansatz $H^{\phi}(t,x,q,S) = x + q\,S + h_q^{\phi}(t)$ and substituting into the above equation, we obtain

$$\partial_t h_q^{\phi} + \alpha q - \phi\sigma^2 q^2 + \sup_{\delta^+,\delta^-\ge 0}\left\{\lambda^+ e^{-\kappa^+\delta^+}\left(\delta^+ + h_{q-1}^{\phi} - h_q^{\phi}\right) + \lambda^- e^{-\kappa^-\delta^-}\left(\delta^- + h_{q+1}^{\phi} - h_q^{\phi}\right)\right\} = 0\,,$$

which, if we let $\phi = \frac{1}{2}\varphi_\alpha$, is equivalent to (23) when $\varphi_\lambda = \varphi_\kappa = 0$ (also see (A.16) in the proof of Proposition 3). Moreover, the optimal posting strategies for this problem are identical to those in Proposition 3 and the value function H^{ϕ} is equal to the value function H. This means that ambiguity aversion specific to midprice drift is equivalent to imposing a penalization on running inventory.

4.2. Robust modelling to misspecification of arrival of market orders

MOs are sent by impatient traders who seek immediate execution. The informational content of these orders is important for MMs. Trading with better informed market participants will generally result in a loss to MMs who naively post orders in the book. One way in which MMs protect themselves from adverse selection costs is to post wider spreads so that on average the money they lose to better informed traders is compensated by earning a wider spread from trading with other market participants.

In the reference model the MM assumes that MOs arrive according to homogeneous Poisson processes and that these orders have no informational content that could affect the midprice dynamics. On the other hand, the MM knows that her model is misspecified. For example, it may be that on average the MM knows how many orders arrive over a period of time, but over shorter time scales

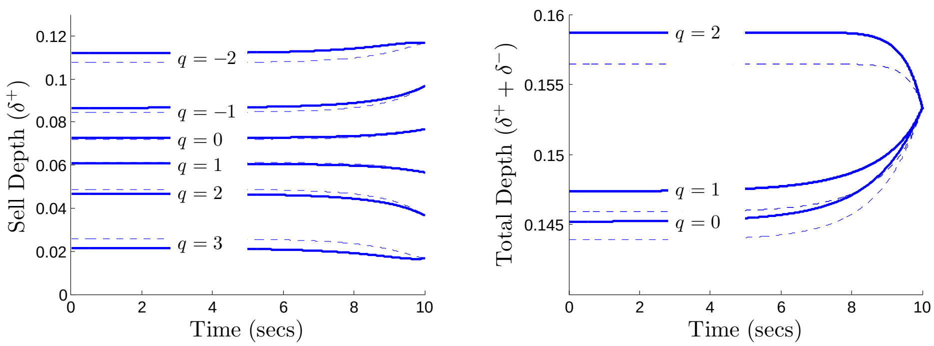


Figure 6: Optimal sell and total depths for an MM who is ambiguity averse to MO rate of arrival (dashed lines are ambiguity neutral depths). Parameter values are $\varphi_\lambda = 6$, $\kappa^\pm = 15$, $\lambda^\pm = 2$, $\sigma = 0.01$, $\alpha = 0$, $\ell(q) = \theta\, q$, $\theta = 0.01$, $\overline{q} = -\underline{q} = 3$ and $T = 10$.

the arrival of orders is more complex. When the MM misspecifies the rate of arrival of MOs she not only posts LOs which are suboptimal, but she also increases the probability of being adversely selected by better informed market participants.

Figure 6 shows the optimal sell and total spreads for $\varphi_\lambda = 6$ ($\varphi_\alpha = \varphi_\kappa = 0$). Similar to the effect of ambiguity on drift we observe that the MM sends quotes to the book that induce faster mean reversion to the optimal level of inventory, compared to the ambiguity neutral MM, and total depths are wider to recover adverse selection costs.

4.3. Robust modelling to misspecification of fill probabilities

Here we consider an MM who is ambiguity averse to the fill probabilities of LOs, but not to their rate of arrival of MOs or midprice drift. When the MM chooses a model robust to misspecification in the fill probability of orders, the distribution of MO execution price relative to the midprice corresponding to the optimal control is no longer exponential. In the reference measure $\mathbb{P}$, the fill probability for a given posting is as in (4), but under the optimal measure $\mathbb{Q}^{*\kappa}$ it inherits dependence on the MM's inventory and time and can be explicitly written:

$$
\begin{aligned}
p_t^\pm(y) &= \int_y^\infty e^{g_q^{\pm\kappa*}(t,y')}\, F^\pm(dy') \\
&= \begin{cases} \dfrac{e^{-\kappa^\pm y} - e^{-\kappa^\pm \delta_q^{\pm*}(t)}\left(1 - e^{-\varphi_\kappa(\delta_q^{\pm*}(t) + h_{q\mp1}(t) - h_q(t))}\right)}{1 - e^{-\kappa^\pm \delta_q^{\pm*}(t)}\left(1 - e^{-\varphi_\kappa(\delta_q^{\pm*}(t) + h_{q\mp1}(t) - h_q(t))}\right)}, & y < \delta_q^{\pm*}(t)\,, \\[2ex] \dfrac{e^{-\kappa^\pm y} e^{-\varphi_\kappa(\delta_q^{\pm*}(t) + h_{q\mp1}(t) - h_q(t))}}{1 - e^{-\kappa^\pm \delta_q^{\pm*}(t)}\left(1 - e^{-\varphi_\kappa(\delta_q^{\pm*}(t) + h_{q\mp1}(t) - h_q(t))}\right)}\,, & y \ge \delta_q^{\pm*}(t)\,. \end{cases}
\end{aligned}
$$

The left-hand panel of Figure 7 shows typical fill probabilities of orders for various levels of q at a

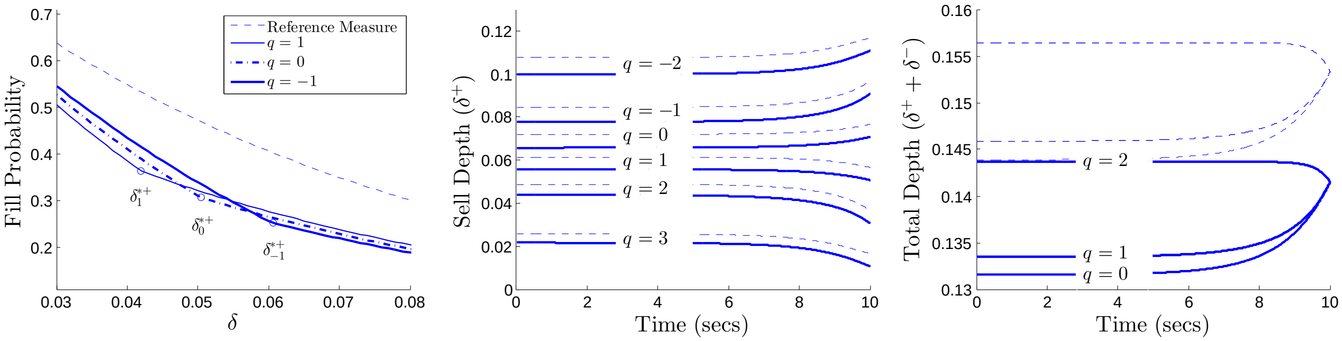


Figure 7: Typical fill probabilities of orders volume for various levels of q and optimal sell and total depths for an MM who is ambiguity averse to MO fill probabilities (dashed lines are ambiguity neutral depths). Parameter values are $\varphi_\kappa = 3$ ($\varphi_\kappa = 15$ in left-hand panel to emphasize the change), $\kappa^\pm = 15$, $\lambda^\pm = 2$, $\sigma = 0.01$, $\alpha = 0$, $\ell(q) = \theta\, q$, $\theta = 0.01$, $\overline{q} = -\underline{q} = 3$.

fixed point in time. Notice that the fill probability under the optimal measure is strictly less than the reference fill probability. The other two panels of the figure show that for $\varphi_\kappa > 0$ ($\varphi_\alpha = \varphi_\lambda = 0$) the MM posts smaller sell and total depths compared to the reference model. The MM fears that her LOs will not be filled and therefore reduces depths to increase churn, but this strategy will not help to recover adverse selection losses, as shown by narrower total depths in the right-hand side panel of the figure.

4.4. Robust modelling to misspecification of arrival rate of MOs, fill probability of MOs, and mid-price drift

In this section we allow the MM to penalize each of ambiguity to midprice drift, ambiguity to MO arrival rate, and ambiguity to LO fill probability at the same time.

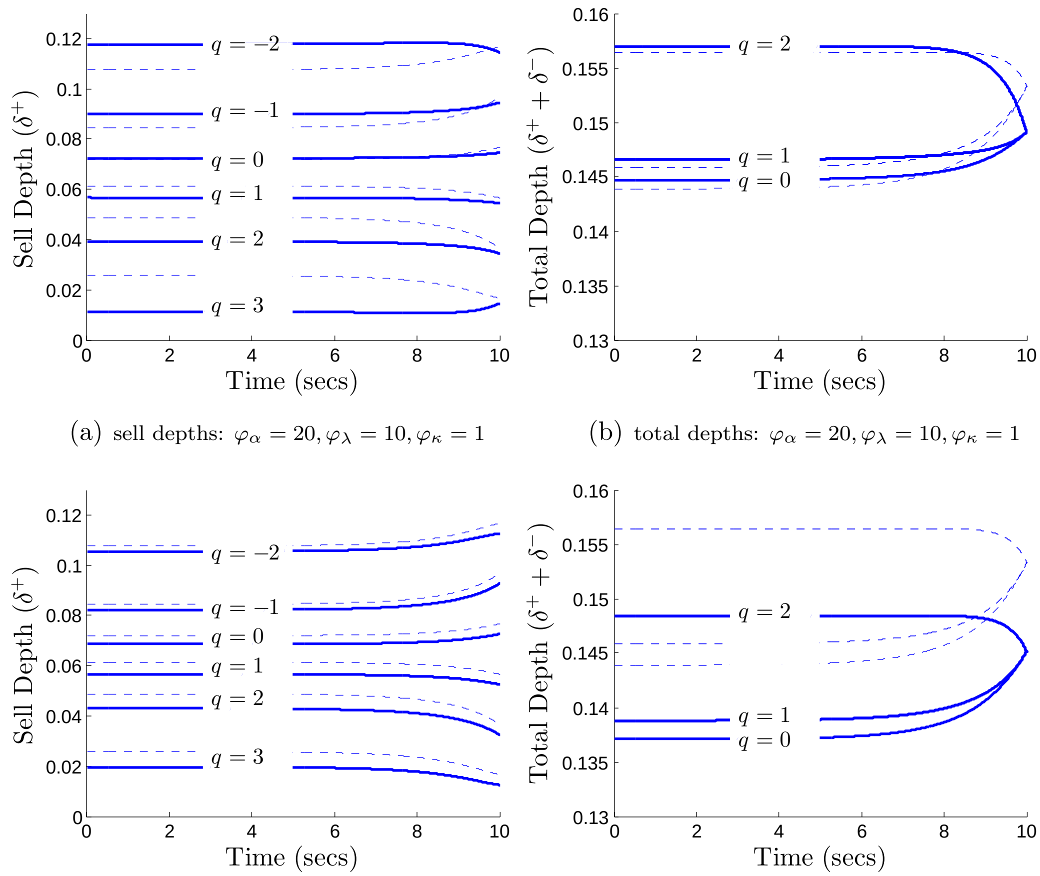
Figure 8 shows the optimal sell depths for an MM who is ambiguity averse with respect to all three factors for various degrees of ambiguity aversion. Here we see that depending on the relative values of φ_α, φ_λ, and φ_κ the effect on the optimal strategy might be different. For example, when approaching maturity when there is ambiguity on fill probabilities:

$$\delta_q^{\pm *}(t) \xrightarrow[t\to T]{} \frac{1}{\varphi_\kappa} \log\left(1 + \frac{\varphi_\kappa}{\kappa}\right) + \theta\bigg((q \mp 1)^2 - q^2\bigg)\,,$$

which is strictly less than the optimal strategy for an ambiguity neutral MM near maturity:

$$\delta_q^{\pm *}(t) \xrightarrow[t\to T]{} \frac{1}{\kappa} + \theta\bigg((q \mp 1)^2 - q^2\bigg)\,.$$

However, as time to maturity increases, ambiguity aversion with respect to midprice drift and MO arrival rate play a more significant role, and for some values of inventory, the optimal quotes can



(a) sell depths: $\varphi_\alpha = 20, \varphi_\lambda = 10, \varphi_\kappa = 1$

(b) total depths: $\varphi_\alpha = 20, \varphi_\lambda = 10, \varphi_\kappa = 1$

(c) sell depths: $\varphi_\alpha = 10, \varphi_\lambda = 1, \varphi_\kappa = 2$

(d) total depths: $\varphi_\alpha = 10, \varphi_\lambda = 1, \varphi_\kappa = 2$

Figure 8: Optimal sell and total depths for an MM who is ambiguity averse towards MO arrival rate, fill probabilities, and midprice drift (dashed lines are ambiguity neutral depths). Parameter values are $\kappa^\pm = 15$, $\lambda^\pm = 2$, $\sigma = 0.01$, $\alpha = 0$, $\ell(q) = \theta\, q$, $\theta = 0.01$, $\overline{q} = -\underline{q} = 3$ and $T = 10$.

become equal to, or even cross, those of the ambiguity neutral MM, see Figure 8 (a) for inventories $q = -2$ and $q = -1$, where $\varphi_\alpha = 20$, $\varphi_\lambda = 10$, and $\varphi_\kappa = 1$. This behaviour is not always present and depends on the relative value of the ambiguity parameters. For example Figure 8 (c) shows depths when $\varphi_\alpha = 10$, $\varphi_\lambda = 1$, and $\varphi_\kappa = 2$ and we observe that the optimal depths of the ambiguity averse MM are always less than those of the ambiguity neutral MM. This will generally be the case when the degrees of ambiguity aversion are weighted more strongly towards fill probabilities than to midprice drift and MO arrival rate.

Moreover, panels (b) and (d) of Figure 8 depict total depths. Panel (b) shows that well before expiry total depths are always wider when the MM is ambiguity averse, and as previously discussed, this helps the strategies to recover losses that stem from adverse selection. On the other hand, panel (d) shows that if ambiguity aversion to midprice drift and MO arrival rate are decreased and ambiguity aversion to fill probabilities is increased, total depths are narrower than those of the ambiguity neutral MM.

We comment that the varying modifications to the optimal quotes due to our notion of ambiguity aversion is behaviour that cannot be captured by a CARA utility function. As seen in Guéant et al. (2012), once the model is chosen and when a time and inventory level are fixed, including or increasing risk aversion can only affect the optimal posting by moving it in one direction ($\delta_q^{\pm *}(t;\gamma)$ is monotone in γ, the risk aversion coefficient). The nature of the monotonicity may change depending on the particular model chosen as well as q and t. Figure 8 and the discussion above show that ambiguity aversion can cause either increases or decreases in the optimal depth within a fixed reference model depending on an MM's levels of ambiguity aversion relative to each other.

5. Robust Market Making and the Profit and Loss of MMs

Making optimal decisions with an incorrect model is costly. But how much does robust market making help to protect the MM from model misspecification? In this section we assume that the true dynamics of the state variables are different from those in the reference measure $\mathbb{P}$, but the MM considers alternative models to reflect her ambiguity aversion. In particular, we show that the Sharpe ratio of the MM's Profit and Loss (PnL) (the ratio of expectation of PnL to standard deviation of PnL) can be considerably increased by including ambiguity aversion to the midprice drift, and the expectation can be increased with ambiguity to arrival and fill probability of MOs while keeping the Sharpe ratio essentially constant.

In this simulation, motivated by Cartea and Jaimungal (2015) and Cartea et al. (2014), where they use real market data to motivate their models, we choose the true dynamics, which are not observed by the MM, to be given by

$$\begin{aligned}
dS_t &= \alpha_t dt + \sigma dW_t\,, \\
d\alpha_t &= -\beta_\alpha \alpha_t dt + \epsilon^+ \, dM_t^+ - \epsilon^- \, dM_t^-\,, \\
d\lambda_t^\pm &= \beta_\lambda(\theta_\lambda - \lambda_t^\pm)dt + \eta_\lambda dM_t^\pm + \nu_\lambda dM_t^\mp\,, \\
d\kappa_t^\pm &= \beta_\kappa(\theta_\kappa - \kappa_t^\pm)dt + \eta_\kappa dM_t^\pm + \nu_\kappa dM_t^\mp\,.
\end{aligned}$$

Here $\epsilon^\pm = 0.001$ is the market impact of MOs which affects (temporarily) the drift of the midprice α_t (short-term alpha) and the arrival rates of MOs follow a bivariate Hawkes process to reflect trade clustering and cross-excitation. Finally, to capture the dynamics of the LOB, the true model assumes that the fill rates are as in the reference model, see (4), but the parameter $\kappa^\pm$ follows a mean-reverting process with jumps to reflect stochastic changes in the depth of the LOB as a consequence of the arrival of MOs. We note that these dynamics correspond to the measure

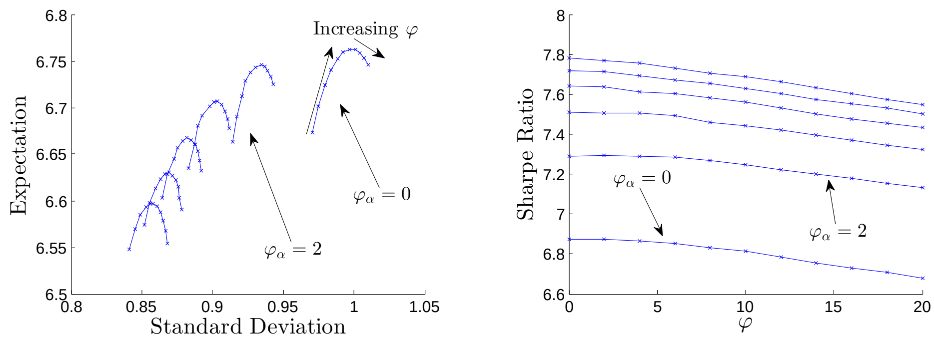


Figure 9: Expectation and standard deviation for various levels of ambiguity parameters when trading strategies are executed for 300 seconds on the simulated process outlined above. The φ_α parameter ranges from 0 to 10 and $\varphi_\lambda = \varphi_\kappa = \varphi$ ranges from 0 to 20.

$\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)$ which is equivalent to the reference measure $\mathbb{P}$, where

$$\begin{aligned}
\eta_t &= \alpha_t \,, \\
g_t^{\pm\lambda} &= \log\left(\frac{\lambda_t^\pm}{\lambda^\pm}\right), \\
g_t^{\pm\kappa}(y) &= \log\left(\frac{\kappa_t^\pm}{\kappa^\pm}\right) - (\kappa_t^\pm - \kappa^\pm)y \,.
\end{aligned}$$

We use 30,000 simulations for each set of ambiguity parameters, liquidation penalty $\ell(q) = \theta\, q$ where $\theta = 0.001$, and $\overline{q} = -\underline{q} = 8$, to calculate the mean and standard deviation of the PnL of the market making strategy. The remaining model parameters used in the simulations are given in Table 1.

σ	$\epsilon^\pm$	β_α	θ_λ	β_λ	η_λ	ν_λ	θ_κ	β_κ	η_κ	ν_κ
0.01	0.001	1	0.2	70/9	5	2	15	7/6	5	2

Table 1: Parameter values used to generate midprice, LOB, and MOs dynamics

We assume that the MM has enough data to calibrate her reference model, which assumes constant arrival rate, exponential fill probabilities and a midprice with zero drift and constant volatility, and calculates the long-run expected value of the above processes

$$\alpha = \lim_{t\to\infty} \mathbb{E}[\alpha_t|\mathcal{F}_0]\,, \quad \lambda^\pm = \lim_{t\to\infty} \mathbb{E}[\lambda_t^\pm|\mathcal{F}_0]\,, \quad \kappa^\pm = \lim_{t\to\infty} \mathbb{E}[\kappa_t^\pm|\mathcal{F}_0]\,,$$

and obtains $\alpha = 0$, $\lambda^\pm = 2$, $\kappa^\pm = 27$.

In Figure 9, each point on the graph represents the performance of a strategy corresponding to a specific set of ambiguity parameters. Each curve signifies the path traced out by a single value of φ_α for various values of φ. The figure shows that there are ranges where increasing ambiguity aversion to MO dynamics and fill probabilities increases expected profits. Moreover, ambiguity aversion to the midprice drift helps the MM in the trade-off between expected and standard deviation of profits. Overall, the result is to see that the Sharpe ratio of MM strategies can be considerably increased when the MM acknowledges that her model is an approximation and she introduces ambiguity aversion to make her decisions robust to misspecification which help her to manage inventory risks and mitigate adverse selection costs.

6. Conclusions

We show how MMs can incorporate ambiguity aversion in their choice of model so that their market making activities are robust to model misspecification. Depending on the degree of confidence that the MM places on the different building blocks of her model our framework allows to control for different degrees of ambiguity on the arrival rate of MOs, fill probability of LOs, and the drift of the fundamental value of the asset they deal. We also find that the actions of an ambiguity averse MM generalize those of an MM that is risk averse through more possible changes to behaviour depending on circumstances.

Robust market making adjusts the MMs quotes to control for exposure to inventory risk and adverse selection costs. Thus, we show that robust market making can generate significant improvements in the profitability of market making strategies. In some cases expected profits increase without increasing the standard deviation of profits and in other cases the increase in expected profits can be achieved along a reduction in standard deviation.

Appendix A. Proof of Main Results

Appendix A.1. Proof of Proposition 2

Proof. First we show property 2. If $\varphi_\alpha = \varphi_\lambda = \varphi_\kappa = \varphi$, then all three logarithm terms in the penalty function may be combined and the product of three Radon-Nikodym derivatives becomes the single measure change:

$$\mathcal{H}(\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)|\mathbb{P}) = \frac{1}{\varphi}\log\left(\frac{d\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)}{d\mathbb{P}}\right).$$

The $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)$ expectation of this expression is relative entropy with the associated aversion level φ, as desired.

Next we show property 1. Consider $\varphi_\lambda > \varphi_\kappa$ and take the $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta, g)$ expectation in two parts. First:

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\frac{1}{\varphi_\alpha}\log\left(\frac{d\mathbb{Q}^\alpha(\eta)}{d\mathbb{P}}\right)\right] \\
&= \frac{1}{\varphi_\alpha}\mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\frac{1}{2}\int_0^T\left(\frac{\alpha-\eta_t}{\sigma}\right)^2 dt\right] \\
&\geq 0,
\end{aligned}
$$

Second:

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\frac{1}{\varphi_\lambda}\log\left(\frac{d\mathbb{Q}^{\alpha,\lambda}(\eta,g)}{d\mathbb{Q}^\alpha(\eta)}\right)+\frac{1}{\varphi_\kappa}\log\left(\frac{d\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}{d\mathbb{Q}^{\alpha,\lambda}(\eta,g)}\right)\right] \\
&= \frac{1}{\varphi_\lambda}\left(\mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\log\left(\frac{d\mathbb{Q}^{\alpha,\lambda}(\eta,g)}{d\mathbb{Q}^\alpha(\eta)}\right)\right]+\frac{\varphi_\lambda}{\varphi_\kappa}\mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\log\left(\frac{d\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}{d\mathbb{Q}^{\alpha,\lambda}(\eta,g)}\right)\right]\right) \\
&\geq \frac{1}{\varphi_\lambda}\left(\mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\log\left(\frac{d\mathbb{Q}^{\alpha,\lambda}(\eta,g)}{d\mathbb{Q}^\alpha(\eta)}\right)\right]+\mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\log\left(\frac{d\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}{d\mathbb{Q}^{\alpha,\lambda}(\eta,g)}\right)\right]\right) \\
&= \frac{1}{\varphi_\lambda}\mathbb{E}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\log\left(\frac{d\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}{d\mathbb{Q}^\alpha(\eta)}\right)\right] \\
&\geq 0\,.
\end{aligned}
$$

The case $\varphi_\kappa \geq \varphi_\lambda$ is identical. □

Appendix A.2. Proof of Proposition 3

Proof. Substituting the ansatz $H(t, x, q, S) = x + q\,S + h_q(t)$ into the PDE (22) results in the system of ODEs (23). The minimization in η is independent of the optimization in $\delta^\pm$, $g^{\pm\lambda}$, and $g^{\pm\kappa}$ and so can be done directly. First order conditions imply that $\eta^* = \alpha - \varphi_\alpha\sigma^2 q$, as desired. This value of η^* is easily seen to be unique as it is a quadratic optimization. For the optimization over $\delta^\pm$, $g^{\pm\lambda}$, and $g^{\pm\kappa}$, first consider $\varphi_\lambda > \varphi_\kappa$. Then the term to be optimized is

$$
\begin{aligned}
\mathfrak{G}(\delta^\pm, g^{\pm\lambda}, g^{\pm\kappa}) &= \lambda^\pm\left[\int_{\delta^\pm}^\infty e^{g^{\pm\lambda}+g^{\pm\kappa}(y)}F^\pm(y)dy\right]\left(\delta^\pm + h_{q\mp1} - h_q\right) + \mathcal{K}_\pm^{\varphi_\lambda,\varphi_\kappa}(g^{\pm\lambda}, g^{\pm\kappa}) \\
&= \lambda^\pm\left[\int_{\delta^\pm}^\infty e^{g^{\pm\lambda}+g^{\pm\kappa}(y)}F^\pm(y)dy\right]\left(\delta^\pm + h_{q\mp1} - h_q\right)\ldots \\
&\quad + \frac{1}{\varphi_\lambda}\left[\lambda^\pm\int_0^\infty -(e^{g^{\pm\lambda}}-1) + g^{\pm\lambda}e^{g^{\pm\lambda}+g^{\pm\kappa}(y)}F^\pm(y)dy\right]\ldots \\
&\quad + \frac{1}{\varphi_\kappa}\left[\lambda^\pm\int_0^\infty -(e^{g^{\pm\kappa}(y)}-1)e^{g^{\pm\lambda}} + g^{\pm\kappa}(y)e^{g^{\pm\lambda}+g^{\pm\kappa}(y)}F^\pm(y)dy\right].
\end{aligned}
\tag{A.1}
$$

The remainder of the proof will proceed as follows:

1. Introduce Lagrange multipliers $\gamma^{\pm}$ corresponding to the constraints on $g^{\pm\kappa}(y)$.
2. Compute first order conditions for the unconstrained $g^{\pm\kappa}(y)$ which minimizes the Lagrange modified term.
3. Compute the values of $\gamma^{\pm}$.
4. Verify that the corresponding $g^{\pm\kappa *}(y)$ provides a minimizer of $\mathfrak{G}(\delta^{\pm}, g^{\pm\lambda}, g^{\pm\kappa})$ for all functions $g^{\pm\kappa} \in \mathcal{G}^{\pm}$.
5. Compute first order conditions for $g^{\pm\lambda}$.
6. Verify that the corresponding $g^{\pm\lambda *}$ provides a minimizer of $\mathfrak{G}(\delta^{\pm}, g^{\pm\lambda}, g^{\pm\kappa *})$.
7. Compute first order conditions for $\delta^{\pm}$ subject to the constraint $\delta^{\pm} \geq 0$.
8. Verify that the corresponding $\delta^{\pm *}$ provides a maximizer of $\mathfrak{G}(\delta^{\pm}, g^{\pm\lambda *}, g^{\pm\kappa *})$.
9. Prove existence and uniqueness for the solution h.

Parts 1 and 2: solving for $g^{\pm\kappa}$: The constraint $\int_0^\infty e^{g^{\pm\kappa}(y)} F^{\pm}(y) dy = 1$ is handled by introducing Lagrange multipliers $\gamma^{\pm}$ and then minimizing over unconstrained $g^{\pm\kappa}$. The optimization with respect to $g^{\pm\kappa}$ is handled in a pointwise fashion by minimizing the integrand with respect to $g^{\pm\kappa}(y)$ for each value of $y \in [0, \infty)$. For $y \in (\delta^{\pm}, \infty)$, the quantity to be minimized is

$$\lambda^{\pm} e^{g^{\pm\lambda} + g^{\pm\kappa}(y)} \Big(\delta^{\pm} + h_{q\mp 1} - h_q \Big) + \frac{\lambda^{\pm}}{\varphi_\lambda} \Big(-(e^{g^{\pm\lambda}} - 1) + g^{\pm\lambda} e^{g^{\pm\lambda} + g^{\pm\kappa}(y)} \Big) \\ + \frac{\lambda^{\pm}}{\varphi_\kappa} \Big(-(e^{g^{\pm\kappa}(y)} - 1) e^{g^{\pm\lambda}} + g^{\pm\kappa}(y) e^{g^{\pm\lambda} + g^{\pm\kappa}(y)} \Big) + \gamma^{\pm} \Big(e^{g^{\pm\kappa}(y)} - 1 \Big) . \tag{A.2}$$

First order conditions in $g^{\pm\kappa}(y)$ give

$$g^{\pm\kappa}(y) = -\frac{\varphi_\kappa}{\varphi_\lambda} g^{\pm\lambda} - \frac{\gamma^{\pm}\varphi_\kappa}{\lambda^{\pm}} e^{-g^{\pm\lambda}} - \varphi_\kappa(\delta^{\pm} + h_{q\mp 1} - h_q) . \tag{A.3}$$

Similarly, first order conditions in $g^{\pm\kappa}(y)$ for $y \in [0, \delta^{\pm}]$ give

$$g^{\pm\kappa}(y) = -\frac{\varphi_\kappa}{\varphi_\lambda} g^{\pm\lambda} - \frac{\gamma^{\pm}\varphi_\kappa}{\lambda^{\pm}} e^{-g^{\pm\lambda}} . \tag{A.4}$$

Combining equations (A.3) and (A.4) gives

$$g^{\pm\kappa *}(y) = -\frac{\varphi_\kappa}{\varphi_\lambda} g^{\pm\lambda} - \frac{\gamma^{\pm}\varphi_\kappa}{\lambda^{\pm}} e^{-g^{\pm\lambda}} - \varphi_\kappa(\delta^{\pm} + h_{q\mp 1} - h_q) \mathbb{1}_{y > \delta^{\pm}} . \tag{A.5}$$

Part 3: solving for $\gamma^{\pm}$: Substituting this expression into the integral constraint and performing some computations gives an expression for $\gamma^{\pm}$:

$$\gamma^{\pm} = -\frac{g^{\pm\lambda}\lambda^{\pm}}{\varphi_{\lambda}} e^{g^{\pm\lambda}} + \frac{\lambda^{\pm} e^{g^{\pm\lambda}}}{\varphi_{\kappa}} \log\left(1 - e^{-\kappa^{\pm}\delta^{\pm}} + e^{-\varphi_{\kappa}(\delta^{\pm}+h_{q\mp1}-h_q)} e^{-\kappa^{\pm}\delta^{\pm}}\right).$$

Substituting this into equation (A.5) gives

$$g^{\pm\kappa*}(y) \quad = \quad -\log\left(1 - e^{-\kappa^{\pm}\delta^{\pm}} + e^{-\varphi_{\kappa}(\delta^{\pm}+h_{q\mp1}-h_q)} e^{-\kappa^{\pm}\delta^{\pm}}\right) - \varphi_{\kappa}(\delta^{\pm} + h_{q\mp1} - h_q)\mathbb{1}_{y>\delta^{\pm}}\,. \qquad \text{(A.6)}$$

Part 4: verify that $g^{\pm\kappa*}$ is a minimizer over $\mathcal{G}^{\pm}$: To prove that this expression for $g^{\pm\kappa*}(y)$ is indeed a minimizer, it will be convenient to introduce some shorthand notation:

$$\Delta^{\pm}h_q = h_{q\mp1} - h_q\,, \qquad \text{(A.7)}$$

$$A^{\pm} = 1 - e^{-\kappa^{\pm}\delta^{\pm}} + e^{-\varphi_{\kappa}(\delta^{\pm}+\Delta^{\pm}h_q)} e^{-\kappa^{\pm}\delta^{\pm}}\,, \qquad \text{(A.8)}$$

$$\underline{g}^{\pm} = -\log A^{\pm}\,, \qquad \text{(A.9)}$$

$$\overline{g}^{\pm} = -\log A^{\pm} - \varphi_{\kappa}(\delta^{\pm} + \Delta^{\pm}h_q)\,. \qquad \text{(A.10)}$$

It is important to note that these quantities do not depend on $g^{\pm\lambda}$. Also note that $\underline{g}^{\pm}$ and $\overline{g}^{\pm}$ are the two possible values that $g^{\pm\kappa*}(y)$ can take depending on whether $y \le \delta^{\pm}$ or $y > \delta^{\pm}$. Let $f^{\pm}$ be any other function in $\mathcal{G}^{\pm}$ and define $k^{\pm}(y) = e^{f^{\pm}(y)} - e^{g^{\pm\kappa*}(y)}$. Then define $f^{\pm}_{\epsilon}(y) = \log(\epsilon k^{\pm}(y) + e^{g^{\pm\kappa*}(y)})$. One can easily check that $f^{\pm}_{\epsilon} \in \mathcal{G}^{\pm}$ for all $\epsilon \in [0,1]$ and that $f^{\pm}_0 = g^{\pm\kappa*}$ and $f^{\pm}_1 = f^{\pm}$. Let $m(\epsilon) = \mathfrak{G}(\delta^{\pm}, g^{\pm\lambda}, f^{\pm}_{\epsilon})$. We will confirm that $g^{\pm\kappa*}$ is the minimizer by showing that

$$\mathfrak{G}(\delta^{\pm}, g^{\pm\lambda}, g^{\pm\kappa*}) = m(0) \le m(1) = \mathfrak{G}(\delta^{\pm}, g^{\pm\lambda}, f^{\pm})\,. \qquad \text{(A.11)}$$

It is sufficient to show that m has non-negative second derivative for all $\epsilon \in [0,1]$. Substituting expressions for $f_\epsilon^\pm$ and (A.6) into $\mathfrak{G}(\delta^\pm, g^{\pm\lambda}, f_\epsilon^\pm)$ gives

$$\begin{aligned}
m(\epsilon) = \quad & \lambda^\pm\left[\int_{\delta^\pm}^\infty e^{g^{\pm\lambda}+f_\epsilon(y)}F^\pm(y)dy\right]\left(\delta^\pm + \Delta^\pm h_q\right)\dots \\
& + \frac{1}{\varphi_\lambda}\left[\lambda^\pm \int_0^\infty -(e^{g^{\pm\lambda}}-1) + g^{\pm\lambda}e^{g^{\pm\lambda}+f_\epsilon(y)}F^\pm(y)dy\right]\dots \\
& + \frac{1}{\varphi_\kappa}\left[\lambda^\pm \int_0^\infty -(e^{f_\epsilon(y)}-1)e^{g^{\pm\lambda}} + f_\epsilon(y)e^{g^{\pm\lambda}+f_\epsilon(y)}F^\pm(y)dy\right] \\
= \quad & \lambda^\pm\left[\int_{\delta^\pm}^\infty e^{g^{\pm\lambda}}(e^{\overline{g}^\pm}+\epsilon k(y))F^\pm(y)dy\right]\left(\delta^\pm + \Delta^\pm h_q\right)\dots \\
& + \frac{1}{\varphi_\lambda}\left[\lambda^\pm \int_0^{\delta^\pm} -(e^{g^{\pm\lambda}}-1) + g^{\pm\lambda}e^{g^{\pm\lambda}}(e^{\underline{g}^\pm}+\epsilon k(y))F^\pm(y)dy\right]\dots \\
& + \frac{1}{\varphi_\lambda}\left[\lambda^\pm \int_{\delta^\pm}^\infty -(e^{g^{\pm\lambda}}-1) + g^{\pm\lambda}e^{g^{\pm\lambda}}(e^{\overline{g}^\pm}+\epsilon k(y))F^\pm(y)dy\right]\dots \\
& + \frac{1}{\varphi_\kappa}\left[\lambda^\pm \int_0^{\delta^\pm} -(e^{\underline{g}^\pm}+\epsilon k(y)-1)e^{g^{\pm\lambda}} + \log(e^{\underline{g}^\pm}+\epsilon k(y))e^{g^{\pm\lambda}}(e^{\underline{g}^\pm}+\epsilon k(y))F^\pm(y)dy\right]\dots \\
& + \frac{1}{\varphi_\kappa}\left[\lambda^\pm \int_{\delta^\pm}^\infty -(e^{\overline{g}^\pm}+\epsilon k(y)-1)e^{g^{\pm\lambda}} + \log(e^{\overline{g}^\pm}+\epsilon k(y))e^{g^{\pm\lambda}}(e^{\overline{g}^\pm}+\epsilon k(y))F^\pm(y)dy\right],
\end{aligned}$$

and taking a derivative with respect to ϵ gives

$$\begin{aligned}
m'(\epsilon) = \quad & \lambda^\pm\left[\int_{\delta^\pm}^\infty e^{g^{\pm\lambda}}k(y)F(y)\,dy\right]\left(\delta^\pm + \Delta^\pm h_q\right) + \frac{1}{\varphi_\lambda}\left[\lambda^\pm \int_0^\infty g^{\pm\lambda}e^{g^{\pm\lambda}}k(y)F^\pm(y)dy\right]\dots \\
& + \frac{1}{\varphi_\kappa}\left[\lambda^\pm \int_0^{\delta^\pm} -k(y)e^{g^{\pm\lambda}} + k(y)e^{g^{\pm\lambda}} + k(y)\log(e^{\underline{g}^\pm}+\epsilon k(y))e^{g^{\pm\lambda}}F^\pm(y)dy\right]\dots \\
& + \frac{1}{\varphi_\kappa}\left[\lambda^\pm \int_{\delta^\pm}^\infty -k(y)e^{g^{\pm\lambda}} + k(y)e^{g^{\pm\lambda}} + k(y)\log(e^{\overline{g}^\pm}+\epsilon k(y))e^{g^{\pm\lambda}}F^\pm(y)dy\right] \\
= \quad & \lambda^\pm\left[\int_{\delta^\pm}^\infty e^{g^{\pm\lambda}}k(y)F(y)\,dy\right]\left(\delta^\pm + \Delta^\pm h_q\right)\dots \\
& + \frac{1}{\varphi_\kappa}\left[\lambda^\pm \int_0^{\delta^\pm} k(y)\log(e^{\underline{g}^\pm}+\epsilon k(y))e^{g^{\pm\lambda}}F^\pm(y)dy\right]\dots \\
& + \frac{1}{\varphi_\kappa}\left[\lambda^\pm \int_{\delta^\pm}^\infty k(y)\log(e^{\overline{g}^\pm}+\epsilon k(y))e^{g^{\pm\lambda}}F^\pm(y)dy\right].
\end{aligned}$$

Evaluating this expression at $\epsilon = 0$ gives

$$
\begin{aligned}
m'(0) = \quad & \lambda^{\pm}\left[\int_{\delta^{\pm}}^{\infty} e^{g^{\lambda\pm}} k(y) F(y)\, dy\right]\left(\delta^{\pm} + \Delta^{\pm} h_q\right)\dots \\
& + \frac{1}{\varphi_\kappa}\left[\lambda^{\pm}\int_0^{\delta^{\pm}} k(y)\underline{g}^{\pm} e^{g^{\lambda\pm}} F^{\pm}(y) dy\right] + \frac{1}{\varphi_\kappa}\left[\lambda^{\pm}\int_{\delta^{\pm}}^{\infty} k(y)\overline{g}^{\pm} e^{g^{\lambda\pm}} F^{\pm}(y) dy\right] \\
= \quad & \lambda^{\pm}\left[\int_{\delta^{\pm}}^{\infty} e^{g^{\lambda\pm}} k(y) F(y)\, dy\right]\left(\delta^{\pm} + \Delta^{\pm} h_q\right) - \frac{1}{\varphi_\kappa}\lambda^{\pm}\int_0^{\delta^{\pm}} k(y)\log(A^{\pm}) e^{g^{\lambda\pm}} F(y)\, dy \dots \\
& - \frac{1}{\varphi_\kappa}\lambda^{\pm}\int_0^{\delta^{\pm}} k(y)\Big(\log(A^{\pm}) + \varphi_\kappa(\delta^{\pm} + \Delta^{\pm} h_q)\Big) e^{g^{\lambda\pm}} F(y)\, dy \\
= \quad & 0\,,
\end{aligned}
$$

as expected. Continuing by taking a second derivative with respect to ϵ:

$$
\begin{aligned}
m''(\epsilon) &= \frac{1}{\varphi_\kappa}\lambda^{\pm}\int_0^{\delta^{\pm}} \frac{e^{g^{\pm\lambda}} k^2(y)}{e^{\underline{g}^{\pm}} + \epsilon k(y)} F(y)\, dy + \frac{1}{\varphi_\kappa}\lambda^{\pm}\int_{\delta^{\pm}}^{\infty} \frac{e^{g^{\pm\lambda}} k^2(y)}{e^{\overline{g}^{\pm}} + \epsilon k(y)} F(y)\, dy \\
&= \frac{1}{\varphi_\kappa}\lambda^{\pm}\int_0^{\infty} \frac{e^{g^{\pm\lambda}} k^2(y)}{e^{g^{\pm\kappa *}(y)} + \epsilon k(y)} F(y)\, dy \\
&= \frac{1}{\varphi_\kappa}\lambda^{\pm}\int_0^{\infty} \frac{e^{g^{\pm\lambda}} k^2(y)}{e^{f_\epsilon(y)}} F(y)\, dy\,.
\end{aligned}
$$

This expression is non-negative for all $\epsilon \in [0, 1]$, showing that indeed the expression for $g^{\pm\kappa *}(y)$ in equation (A.6) is a minimizer. This expression is strictly positive unless $k \equiv 0$, showing that the inequality in (A.11) is strict unless $f = g^{\pm\kappa *}$, therefore $g^{\pm\kappa *}$ is the unique minimizer.

Part 5: first order conditions for $g^{\pm\lambda}$: After substituting the expression (A.6) into the term to be minimized (A.1) and performing some tedious computations, we must minimize the following with respect to $g^{\pm\lambda}$:

$$
\begin{aligned}
& \lambda^{\pm} e^{g^{\pm\lambda}} e^{\overline{g}^{\pm}} (\delta^{\pm} + \Delta^{\pm} h_q) e^{-\kappa^{\pm}\delta^{\pm}} \\
& + \frac{\lambda^{\pm}}{\varphi_\lambda}\Big(-(e^{g^{\pm\lambda}} - 1) + g^{\pm\lambda} e^{g^{\pm\lambda}} e^{\overline{g}^{\pm}}\Big) e^{-\kappa^{\pm}\delta^{\pm}} + \frac{\lambda^{\pm}}{\varphi_\kappa}\Big(-(e^{\overline{g}^{\pm}} - 1) e^{g^{\pm\lambda}} + \overline{g}^{\pm} e^{g^{\pm\lambda}} e^{\overline{g}^{\pm}}\Big) e^{-\kappa^{\pm}\delta^{\pm}} \\
& + \frac{\lambda^{\pm}}{\varphi_\lambda}\Big(-(e^{g^{\pm\lambda}} - 1) + g^{\pm\lambda} e^{g^{\pm\lambda}} e^{\underline{g}^{\pm}}\Big)(1 - e^{-\kappa^{\pm}\delta^{\pm}}) + \frac{\lambda^{\pm}}{\varphi_\kappa}\Big(-(e^{\underline{g}^{\pm}} - 1) e^{g^{\pm\lambda}} + \underline{g}^{\pm} e^{g^{\pm\lambda}} e^{\underline{g}^{\pm}}\Big)(1 - e^{-\kappa^{\pm}\delta^{\pm}})\,.
\end{aligned}
\tag{A.12}
$$

Applying first order conditions in $g^{\pm\lambda}$ and carrying out some tedious computations gives the candidate minimizer:

$$g^{\pm\lambda *} = \frac{\varphi_\lambda}{\varphi_\kappa} \log A^{\pm} = \frac{\varphi_\lambda}{\varphi_\kappa} \log\left(1 - e^{-\kappa^{\pm}\delta^{\pm}} + e^{-\varphi_\kappa(\delta^{\pm} + h_{q\mp 1} - h_q)} e^{-\kappa^{\pm}\delta^{\pm}}\right) . \tag{A.13}$$

This is the unique root corresponding to the first order conditions.

Part 6: verify that $g^{\pm\lambda *}$ is a minimizer: Taking two derivatives of (A.12) with respect to $g^{\pm\lambda}$ and cancelling terms gives

$$\frac{g^{\pm\lambda} e^{g^{\pm\lambda}}}{\varphi_\lambda} + \frac{e^{g^{\pm\lambda}}}{\varphi_\lambda} - \frac{\log(A^{\pm}) e^{g^{\pm\lambda}}}{\varphi_\kappa} .$$

When the expression (A.13) is substituted above, this becomes $\frac{A^{\pm \frac{\varphi_\lambda}{\varphi_\kappa}}}{\varphi_\lambda}$, which is always positive because $A^{\pm} > 0$. Thus, this value of $g^{\pm\lambda *}$ provides a minimizer. Uniqueness of the root corresponding to first order conditions (and the fact that it is the only critical value) implies that $g^{\pm\lambda *}$ is the unique minimizer.

Part 7: solving for $\delta^{\pm}$: Substituting expressions (A.7) to (A.10) and (A.13) into (A.12), after some tedious computations we must maximize the following expression over $\delta^{\pm}$:

$$\frac{\lambda^{\pm}}{\varphi_\lambda}\left(1 - \exp\left\{\frac{\varphi_\lambda}{\varphi_\kappa} \log(1 - e^{-\kappa^{\pm}\delta^{\pm}} + e^{-\varphi_\kappa(\delta^{\pm} + h_{q\mp 1} - h_q)} e^{-\kappa^{\pm}\delta^{\pm}})\right\}\right) .$$

Maximizing this term is equivalent to minimizing

$$\exp\left\{\frac{\varphi_\lambda}{\varphi_\kappa} \log(1 - e^{-\kappa^{\pm}\delta^{\pm}} + e^{-\varphi_\kappa(\delta^{\pm} + h_{q\mp 1} - h_q)} e^{-\kappa^{\pm}\delta^{\pm}})\right\},$$

which is equivalent to minimizing:

$$1 - e^{-\kappa^{\pm}\delta^{\pm}} + e^{-\varphi_\kappa(\delta^{\pm} + h_{q\mp 1} - h_q)} e^{-\kappa^{\pm}\delta^{\pm}} . \tag{A.14}$$

Computing first order conditions for $\delta^{\pm}$ gives

$$\delta^{\pm *} = \frac{1}{\varphi_\kappa} \log\left(1 + \frac{\varphi_\kappa}{\kappa^{\pm}}\right) - h_{q\mp 1} + h_q . \tag{A.15}$$

If this value is positive, we check that it is a minimizer of (A.14) by taking a second derivative. If it is non-negative, we show that the first derivative of (A.14) is positive for all $\delta^{\pm} > 0$, meaning that the desired value of $\delta^{\pm *}$ is 0.

Part 8: verify that $\delta^{\pm *}$ is a minimizer of (A.14): Suppose the value given by (A.15) is positive. Taking two derivatives of (A.14) with respect to $\delta^{\pm}$ gives

$$-\kappa^{\pm 2} e^{-\kappa^{\pm}\delta^{\pm}} + (\varphi_\kappa + \kappa^{\pm})^2 e^{-\varphi_\kappa(\delta^{\pm} + h_{q\mp 1} - h_q)} e^{-\kappa^{\pm}\delta^{\pm}}.$$

Substituting (A.15) into this expression gives

$$\kappa^{\pm}\varphi_\kappa e^{-\kappa^{\pm}\delta^{\pm *}} > 0$$

and so the value in (A.15) minimizes (A.14). Now suppose the value in (A.15) is non-positive. This means the following inequality holds:

$$e^{-\varphi_\kappa(h_{q\mp 1} - h_q)} \le \frac{\kappa^{\pm}}{\kappa^{\pm} + \varphi_\kappa}.$$

The first derivative of (A.14) with respect to $\delta^{\pm}$ is

$$\left(\kappa^{\pm} - (\kappa^{\pm} + \varphi_\kappa) e^{-\varphi_\kappa(\delta^{\pm} + h_{q\mp 1} - h_q)}\right) e^{-\kappa^{\pm}\delta^{\pm}},$$

and the preceding inequality implies that this is non-negative for all $\delta^{\pm} \ge 0$, implying that $\delta^{\pm *} = 0$ is the minimizer of (A.14). Thus, the value of $\delta^{\pm}$ which maximizes the original term of interest is

$$\delta^{\pm *} = \left(\frac{1}{\varphi_\kappa}\log(1 + \frac{\varphi_\kappa}{\kappa^{\pm}}) - h_{q\mp 1} + h_q\right)_+,$$

as desired. The case of $\varphi_\kappa > \varphi_\lambda$ is essentially identical.

Part 9: existence and uniqueness of h: Begin by substituting the optimal feedback controls, η^*, $g^{\pm\lambda*}$, and $g^{\pm\kappa*}(y)$ into equation (23). This results in:

$$\begin{aligned}
&\partial_t h_q + \alpha q - \frac{1}{2}\varphi_\alpha \sigma^2 q^2 \\
&\quad + \sup_{\delta^+\geq 0}\left\{\frac{\lambda^+}{\varphi_\lambda}\left(1-\exp\left\{\frac{\varphi_\lambda}{\varphi_\kappa}\log(1-e^{-\kappa^+\delta^+}+e^{-\kappa^+\delta^+-\varphi_\kappa(\delta^++h_{q-1}-h_q)})\right\}\right)\right\}\mathbb{1}_{q\neq \underline{q}} \\
&\quad + \sup_{\delta^-\geq 0}\left\{\frac{\lambda^-}{\varphi_\lambda}\left(1-\exp\left\{\frac{\varphi_\lambda}{\varphi_\kappa}\log(1-e^{-\kappa^-\delta^-}+e^{-\kappa^-\delta^--\varphi_\kappa(\delta^-+h_{q+1}-h_q)})\right\}\right)\right\}\mathbb{1}_{q\neq \overline{q}} = 0\,, \\
&\qquad\qquad h_q(T) = -q\,\ell(q)\,.
\end{aligned} \tag{A.16}$$

This is a system of ODEs of the form $\partial_t \mathbf{h} = \mathbf{F}(\mathbf{h})$. To show existence and uniqueness of the solution to this equation, the function $\mathbf{F}$ will be shown to be bounded and globally Lipschitz. It suffices to show that the function f is bounded and globally Lipschitz, where f is given by

$$f(x,y) = \sup_{\delta\geq 0}\left\{\frac{\lambda}{\varphi_\lambda}\left(1-\exp\left\{\frac{\varphi_\lambda}{\varphi_\kappa}\log(1-e^{-\kappa\delta}+e^{-\kappa\delta-\varphi_\kappa(\delta+x-y)})\right\}\right)\right\}\,.$$

Boundedness and the global Lipschitz property of f implies the same for $\mathbf{F}$, and so existence and uniqueness follows from the Picard-Lindelöf theorem. The global Lipschitz property will be a result of showing that all directional derivatives of f exist and are bounded for all $(x,y)\in\mathbb{R}^2$.

The supremum is attained at $\delta^* = (\frac{1}{\varphi_\kappa}\log(1+\frac{\varphi_\kappa}{\kappa}) - x + y)_+$. Thus, two separate domains for f must be considered: $\frac{1}{\varphi_\kappa}\log(1+\frac{\varphi_\kappa}{\kappa}) > x-y$ and $\frac{1}{\varphi_\kappa}\log(1+\frac{\varphi_\kappa}{\kappa}) \leq x-y$. First consider $\frac{1}{\varphi_\kappa}\log(1+\frac{\varphi_\kappa}{\kappa}) > x-y$ so that $\delta^* = \frac{1}{\varphi_\kappa}\log(1+\frac{\varphi_\kappa}{\kappa}) - x + y$. Substituting this into the expression for f yields:

$$\begin{aligned}
f(x,y) &= \frac{\lambda}{\varphi_\lambda}\left(1-\exp\left\{\frac{\varphi_\lambda}{\varphi_\kappa}\log\left(1-e^{-\frac{\kappa}{\varphi_\kappa}\log(1+\frac{\varphi_\kappa}{\kappa})+\kappa(x-y)}(1-e^{-\log(1+\frac{\varphi_\kappa}{\kappa})})\right)\right\}\right) \\
&= \frac{\lambda}{\varphi_\lambda}\left(1-\exp\left\{\frac{\varphi_\lambda}{\varphi_\kappa}\log\left(1-Be^{\kappa(x-y)}\right)\right\}\right),
\end{aligned}$$

where $B = (\frac{\kappa}{\varphi_\kappa+\kappa})^{\frac{\kappa}{\varphi_\kappa}}\frac{\varphi_\kappa}{\varphi_\kappa+\kappa} > 0$. Letting $z = Be^{\kappa(x-y)}$, the inequality $\frac{1}{\varphi_\kappa}\log(1+\frac{\varphi_\kappa}{\kappa}) > x-y$ implies

$$\begin{aligned}
0 < z &< (\frac{\kappa}{\varphi_\kappa+\kappa})^{\frac{\kappa}{\varphi_\kappa}}\frac{\varphi_\kappa}{\varphi_\kappa+\kappa}e^{\frac{\kappa}{\varphi_\kappa}\log(1+\frac{\varphi_\kappa}{\kappa})} \\
&= (\frac{\kappa}{\varphi_\kappa+\kappa})^{\frac{\kappa}{\varphi_\kappa}}\frac{\varphi_\kappa}{\varphi_\kappa+\kappa}(\frac{\varphi_\kappa+\kappa}{\kappa})^{\frac{\kappa}{\varphi_\kappa}} = \frac{\varphi_\kappa}{\varphi_\kappa+\kappa} < 1\,.
\end{aligned}$$

Since z is positive we have:

$$f(x,y) = \frac{\lambda}{\varphi_\lambda}\left(1 - \exp\left\{\frac{\varphi_\lambda}{\varphi_\kappa}\log\left(1-z\right)\right\}\right) < \frac{\lambda}{\varphi_\lambda}\ .$$

Taking partial derivatives of f in this domain gives

$$\begin{aligned}\partial_x f(x,y) = -\partial_y f(x,y) &= \frac{\lambda}{\varphi_\kappa} e^{\frac{\varphi_\lambda}{\varphi_\kappa}\log\left(1-Be^{\kappa(x-y)}\right)}\frac{B\kappa e^{\kappa(x-y)}}{1-Be^{\kappa(x-y)}} \\ &= \frac{\lambda}{\varphi_\kappa} e^{\frac{\varphi_\lambda}{\varphi_\kappa}\log\left(1-z\right)}\frac{z}{1-z}\ . \end{aligned} \tag{A.17}$$

This expression is non-negative and continuous for $0 \le z \le \frac{\varphi_\kappa}{\varphi_\kappa+\kappa}$, and therefore achieves a finite maximum somewhere on that interval. Thus, $\partial_x f$ and $\partial_y f$ are bounded in this domain, and so directional derivatives exist and are also bounded everywhere in the interior of the domain. On the boundary, directional derivatives exist and are bounded if the direction is towards the interior of the domain.

Now consider $\frac{1}{\varphi_\kappa}\log(1+\frac{\varphi_\kappa}{\kappa}) \le x-y$, which implies $\delta^* = 0$. The expression for $f(x,y)$ in this domain is

$$f(x,y) = \frac{\lambda}{\varphi_\lambda}\left(1 - e^{-\varphi_\lambda(x-y)}\right)$$

which is bounded by $\frac{\lambda}{\varphi_\lambda}\left(1 - e^{-\frac{\varphi_\lambda}{\varphi_\kappa}\log(1+\frac{\varphi_\kappa}{\kappa})}\right)$. Partial derivatives of f are given by

$$\partial_x f(x,y) = -\partial_y f(x,y) = \lambda e^{-\varphi_\lambda(x-y)}\ . \tag{A.18}$$

In this domain, the derivatives $\partial_x f$ and $\partial_y f$ are bounded by $\lambda e^{-\frac{\varphi_\lambda}{\varphi_\kappa}\log(1+\frac{\varphi_\kappa}{\kappa})}$. So similarly to the first domain, directional derivatives exist and are bounded in the interior. On the boundary, they exist and are bounded in the direction towards the interior of the domain. Thus, we have existence and boundedness on the boundary towards either of the two domains. The directional derivative on the boundary is zero when the direction is parallel to the boundary. Existence and boundedness of directional derivatives for all $(x,y) \in \mathbb{R}^2$ allows us to show the Lipschitz condition easily:

$$|f(x_2,y_2) - f(x_1,y_1)| = \left|\int_C \nabla f(x,y)\cdot d\vec{r}\right| \le \int_C |\nabla f(x,y)| ds \le \int_C A ds = A|(x_2,y_2)-(x_1,y_1)|$$

where C is the curve which connects (x_1,y_1) to (x_2,y_2) in a straight line and A is a uniform bound on the gradient of f. This proves that there exists a unique solution h to equation (23).

□

Appendix A.3. Proof of Theorem (4)

Proof. Let h be the solution to equation (23) with terminal conditions $h_q(T) = -q\,\ell(q)$, and define a candidate value function by $\hat{H}(t,x,q,S) = x + q\,S + h_q(t)$. From Ito's lemma we have

$$\begin{aligned}\hat{H}(T, X^{\delta^\pm}_{T^-}, S_{T^-}, q^{\delta^\pm}_{T^-}) = \hat{H}(t,x,S,q) &+ \int_t^T \partial_t h_{q_s}(s)ds + \int_t^T \alpha q_s ds + \sigma \int_t^T q_s dW_s \\ &+ \int_t^T \int_{\delta_s^+}^\infty \Big(\delta_s^+ + h_{q_{s^-}-1}(s) - h_{q_{s^-}}(s)\Big)\mu^+(dy,ds) \\ &+ \int_t^T \int_{\delta_s^-}^\infty \Big(\delta_s^- + h_{q_{s^-}+1}(s) - h_{q_{s^-}}(s)\Big)\mu^-(dy,ds)\,.\end{aligned}$$

Note that for any admissible measure $\mathbb{Q}(\eta,g)$ and admissible control $\delta^\pm$ we have

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}(\eta,g)}\left[\int_0^T \int_{\delta_t^\pm}^\infty (\delta_t^\pm)^2\, \nu^\pm_{\mathbb{Q}(\eta,g)}(dy,dt)\right] &= \mathbb{E}^{\mathbb{Q}(\eta,g)}\left[\int_0^T \int_{\delta_t^\pm}^\infty (\delta_t^\pm)^2 e^{g_t^\pm(y)}\, \nu^\pm(dy,dt)\right] \\ &\le \mathbb{E}^{\mathbb{Q}(\eta,g)}\left[\int_0^T \int_{\delta_t^\pm}^\infty y^2 e^{g_t^\pm(y)}\, \nu^\pm(dy,dt)\right] \\ &\le \mathbb{E}^{\mathbb{Q}(\eta,g)}\left[\int_0^T \int_0^\infty y^2 e^{g_t^\pm(y)}\, \nu^\pm(dy,dt)\right] < \infty.\end{aligned}$$

The remainder of the proof proceeds as follows:

1. Show that the feedback forms of $\delta^\diamond$, $\eta^\diamond$, $g^{\pm\lambda\diamond}$, and $g^{\pm\kappa\diamond}$ are admissible.
2. For an arbitrary admissible $\delta = (\delta_t^\pm)_{0\le t\le T}$, we define an admissible response measure indexed by M which will be denoted $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta), g_M(\delta))$.
3. We show that M can be taken sufficiently large, independent of t and δ, such that the response measure is pointwise (in t) ϵ-optimal.
4. We show that the candidate function $\hat{H}$ satisfies $\hat{H}(t,x,S,q) \ge H(t,x,S,q)$.
5. We show that the candidate function $\hat{H}$ satisfies $\hat{H}(t,x,S,q) \le H(t,x,S,q)$.

Step 1: $\delta^\diamond$, $\eta^\diamond$, $g^{\pm\lambda\diamond}$, and $g^{\pm\kappa\diamond}$ are admissible: Since q is bounded between $\underline{q}$ and $\overline{q}$, $\eta^\diamond$ is bounded and therefore admissible. The existence and uniqueness of a classical solution for h means that it achieves a finite maximum and minimum for some $q \in \{\underline{q}, \dots, \overline{q}\}$ and $t \in [0, T]$. Thus, from the feedback expressions for $g^{\pm\lambda\diamond}$ and $g^{\pm\kappa\diamond}$, we see that they are also bounded and therefore admissible. Admissibility of $\delta^\diamond$ is clear.

Step 2: Defining admissible response measure: Let $\delta = (\delta_t^\pm)_{0 \le t \le T}$ be an arbitrary admissible control and define pointwise minimizing response controls by

$$\begin{aligned}
\eta_t(\delta) &= \alpha - \varphi_\alpha\, \sigma^2\, q_t\,, \\
g_t^{\pm\lambda}(\delta) &= \frac{\varphi_\lambda}{\varphi_\kappa} \log\left(1 - e^{-\kappa^\pm \delta_t^\pm}\left(1 - e^{-\varphi_\kappa(\delta_t^\pm + h_{q_t \mp 1}(t) - h_{q_t}(t))}\right)\right), \\
g_t^{\pm\kappa}(y;\delta) &= -\log\left(1 - e^{-\kappa^\pm \delta_t^\pm}\left(1 - e^{-\varphi_\kappa(\delta_t^\pm + h_{q_t \mp 1}(t) - h_{q_t}(t))}\right)\right) \\
&\qquad - \varphi_\kappa\left(\delta_t^\pm + h_{q_t \mp 1}(t) - h_{q_t}(t)\right)\mathbb{1}_{y \ge \delta_t^\pm}\,.
\end{aligned}$$

These processes each have the same form as the pointwise minimizers found in Proposition 3, and so for a given $\delta = (\delta_t^\pm)_{0 \le t \le T}$ these controls achieve the pointwise infimum in equation (23). Since h is a classical solution to equation (23), it is bounded for $t \in [0, T]$ and $\underline{q} \le q \le \overline{q}$. Using the boundedness of h, we see that $g_t^{\pm\lambda}(0)$ is finite and bounded with respect to t, and $\lim_{\delta \to \infty} g_t^{\pm\lambda}(\delta) = 0$, therefore $g_t^{\pm\lambda}(\delta)$ is bounded. It is also clear that $\eta_t(\delta)$ is bounded. However, $g_t^{\pm\kappa}(y;\delta)$ is only bounded from above, so it is possible that the pair $(\eta_t(\delta), g_t(\delta))$ does not define an admissible measure as per the definition in (16). In order to proceed, we use a modification of $g_t^{\pm\kappa}$:

$$\begin{aligned}
g_{t,M}^{\pm\kappa}(y;\delta) &= -\log\left(1 - e^{-\kappa^\pm \delta_t^\pm}\left(1 - e^{-\varphi_\kappa(\delta_t^\pm + h_{q_t \mp 1}(t) - h_{q_t}(t))}\right)\right) \\
&\qquad - \varphi_\kappa \min\left(\delta_t^\pm + h_{q_t \mp 1}(t) - h_{q_t}(t), M\right)\mathbb{1}_{y \ge \delta_t^\pm}\,.
\end{aligned}$$

Since $g_{t,M}^{\pm\kappa}$ is bounded, letting $g_{t,M}^{\pm}(y;\delta) = g_t^{\pm\lambda}(\delta) + g_{t,M}^{\pm\kappa}(y;\delta)$, the pair $(\eta_t(\delta), g_M(\delta))$ does define an admissible measure $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta), g_M(\delta))$. Note that for a fixed t and δ_t, $g_{t,M}^{\pm\kappa}(y;\delta) \to g_t^{\pm\kappa}(y;\delta)$ as $M \to \infty$ pointwise in y and in $L^1(F^\pm(dy))$.

Step 3: Showing pointwise ϵ-optimality: As in the proof of Proposition 3 in Section Appendix A.2, consider the functional

$$\begin{aligned}
\mathfrak{G}(t, \delta^\pm, g^{\pm\lambda}, g^{\pm\kappa}) = \quad & \lambda^\pm\left[\int_{\delta^\pm}^\infty e^{g^{\pm\lambda} + g^{\pm\kappa}(y)} F^\pm(y)dy\right]\left(\delta^\pm + h_{q\mp 1}(t) - h_q(t)\right) \\
& + \mathcal{K}_\pm^{\varphi_\lambda,\varphi_\kappa}(g^{\pm\lambda}, g^{\pm\kappa})\mathbb{1}_{\varphi_\lambda > \varphi_\kappa} + \mathcal{K}_\pm^{\varphi_\kappa,\varphi_\lambda}(g^{\pm\kappa}, g^{\pm\lambda})\mathbb{1}_{\varphi_\kappa > \varphi_\lambda}\,.
\end{aligned}$$

We will now show

$$\lim_{M\to\infty} \mathfrak{G}(t,\delta_t^{\pm}, g_t^{\pm\lambda}(\delta), g_{t,M}^{\pm\kappa}(\cdot;\delta)) = \mathfrak{G}(t,\delta_t^{\pm}, g_t^{\pm\lambda}(\delta), g_t^{\pm\kappa}(\cdot;\delta))$$

uniformly in t and δ. Consider the first term only, and compute the difference when evaluated at both $g_{t,M}^{\pm\kappa}(\cdot;\delta_t)$ and $g_t^{\pm\kappa}(\cdot;\delta_t)$. which we will denote by $\mathfrak{J}(t,\delta,M)$:

$$\begin{aligned}
&\mathfrak{J}(t,\delta,M) \\
&\quad = \lambda^{\pm} e^{g_t^{\pm\lambda}(\delta)} \left|\delta_t^{\pm} + h_{q\mp1}(t) - h_q(t)\right| \left| \int_{\delta_t^{\pm}}^{\infty} e^{g_{t,M}^{\pm\kappa}(y;\delta)} F^{\pm}(y)dy - \int_{\delta_t^{\pm}}^{\infty} e^{g_t^{\pm\kappa}(y;\delta)} F^{\pm}(y)dy \right| \\
&\quad = \lambda^{\pm} e^{(1-\frac{\varphi_\kappa}{\varphi_\lambda}) g_t^{\pm\lambda}(\delta)} \left|\delta_t^{\pm} + h_{q\mp1}(t) - h_q(t)\right| e^{-\kappa^{\pm}\delta_t^{\pm}} \left| e^{-\varphi_\kappa \min(\delta_t^{\pm}+h_{q\mp1}(t)-h_q(t),M)} - e^{-\varphi_\kappa(\delta_t^{\pm}+h_{q\mp1}(t)-h_q(t))} \right| \\
&\quad = \lambda^{\pm} e^{(1-\frac{\varphi_\kappa}{\varphi_\lambda}) g_t^{\pm\lambda}(\delta)} \left|\delta_t^{\pm} + h_{q\mp1}(t) - h_q(t)\right| e^{-\kappa^{\pm}\delta_t^{\pm}} e^{-\varphi_\kappa M} \left| 1 - e^{-\varphi_\kappa(\delta_t^{\pm}+h_{q\mp1}(t)-h_q(t)-M)} \right| \mathbb{1}_{\delta_t^{\pm}+h_{q\mp1}(t)-h_q(t)\geq M} \\
&\quad \leq \lambda^{\pm} e^{(1-\frac{\varphi_\kappa}{\varphi_\lambda}) g_t^{\pm\lambda}(\delta)} \left|\delta_t^{\pm} + h_{q\mp1}(t) - h_q(t)\right| e^{-\kappa^{\pm}\delta_t^{\pm}} e^{-\varphi_\kappa M}.
\end{aligned}$$

As previously noted, both h and $g^{\pm\lambda}$ are uniformly bounded, say by C and D respectively, so clearly $\mathfrak{J}(t,\delta,M)$ is bounded. For an arbitrary $\epsilon' > 0$, we may choose M sufficiently large such that

$$\mathfrak{J}(t,\delta,M) \leq \lambda^{\pm} e^{|1-\frac{\varphi_\kappa}{\varphi_\lambda}|D} \left(\delta_t^{\pm} + 2C\right) e^{-\kappa^{\pm}\delta_t^{\pm}} e^{-\varphi_\kappa M} < \epsilon' \quad \text{for all} \quad \delta_t^{\pm} \geq 0\,.$$

Showing uniform convergence of $\mathcal{K}_{\pm}^{\varphi_\lambda,\varphi_\kappa}(g_t^{\pm\lambda}(\delta), g_{t,M}^{\pm\kappa}(\cdot;\delta))\mathbb{1}_{\varphi_\lambda>\varphi_\kappa}$ and $\mathcal{K}_{\pm}^{\varphi_\kappa,\varphi_\lambda}(g_{t,M}^{\pm\kappa}(\cdot,\delta), g^{\pm\lambda}(\delta))\mathbb{1}_{\varphi_\kappa>\varphi_\lambda}$ is essentially the same and so the details are omitted.

Let $\epsilon > 0$ be arbitrary and let M be sufficiently large (chosen independently of t and δ) so that

$$0 \quad < \quad \mathfrak{G}(t,\delta_t^{\pm}, g_t^{\pm\lambda}(\delta), g_{t,M}^{\pm\kappa}(\cdot;\delta)) - \mathfrak{G}(t,\delta_t^{\pm}, g_t^{\pm\lambda}(\delta), g_t^{\pm\kappa}(\cdot;\delta)) \quad < \quad \epsilon\,.$$

Then since δ is arbitrary and h satisfies equation (23), the following inequality holds almost surely for every t:

$$
\begin{aligned}
\partial_t h_{q_t} + \eta_t(\delta) q_t + \frac{1}{2\varphi_\alpha}\left(\frac{\alpha - \eta_t(\delta)}{\sigma}\right)^2 \\
&+ \lambda^+ \left[\int_{\delta_t^+}^{\infty} e^{g_t^{+\lambda}(\delta) + g_{t,M}^{+\kappa}(y;\delta)} F^+(dy)\right] (\delta_t^+ + h_{q_t-1}(t) - h_{q_t}(t)) \\
&+ \mathcal{K}_+^{\varphi_\lambda,\varphi_\kappa}(g_t^{+\lambda}(\delta), g_{t,M}^{+\kappa}(\cdot;\delta))\, \mathbb{1}_{\varphi_\lambda \geq \varphi_\kappa} + \mathcal{K}_+^{\varphi_\kappa,\varphi_\lambda}(g_{t,M}^{+\kappa}(\cdot;\delta), g_t^{+\lambda}(\delta))\, \mathbb{1}_{\varphi_\lambda < \varphi_\kappa} \\
&+ \lambda^- \left[\int_{\delta_t^-}^{\infty} e^{g_t^{-\lambda}(\delta) + g_{t,M}^{-\kappa}(y;\delta)} F^-(dy)\right] (\delta_t^- + h_{q_t+1}(t) - h_{q_t}(t)) \\
&+ \mathcal{K}_-^{\varphi_\lambda,\varphi_\kappa}(g_t^{-\lambda}(\delta), g_{t,M}^{-\kappa}(\cdot;\delta))\, \mathbb{1}_{\varphi_\lambda \geq \varphi_\kappa} + \mathcal{K}_-^{\varphi_\kappa,\varphi_\lambda}(g_{t,M}^{-\kappa}(\cdot;\delta), g_t^{-\lambda}(\delta))\, \mathbb{1}_{\varphi_\lambda < \varphi_\kappa} < \epsilon\,.
\end{aligned}
\tag{A.19}
$$

Thus, the measure $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta), g_M(\delta))$ is pointwise (in t) ϵ-optimal, uniformly in δ.

Step 4: Showing $\hat{H}(t,x,S,q) \geq H(t,x,S,q)$: Taking an expectation of $\hat{H}(T, X_{T^-}^{\delta^\pm}, S_{T^-}, q_{T^-}^{\delta^\pm})$ in the measure $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta), g_M(\delta))$, and using (A.19), gives

$$
\begin{aligned}
&\mathbb{E}_{t,x,q,S}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta),g_M(\delta))}\left[\hat{H}(T, X_{T^-}^{\delta^\pm}, S_{T^-}, q_{T^-}^{\delta^\pm})\right] \\
&\quad = \hat{H}(t,x,S,q) + \mathbb{E}_{t,x,q,S}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta),g_M(\delta))}\Bigg[\int_t^T \partial_t h_{q_s}(s) ds + \int_t^T \alpha q_s ds + \sigma \int_t^T q_s dW_s \\
&\qquad\qquad + \int_t^T \int_{\delta_s^+}^{\infty} \left(\delta_s^+ + h_{q_s-1}(s) - h_{q_s}(s)\right) \nu^+_{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta),g_M(\delta))}(dy, ds) \\
&\qquad\qquad + \int_t^T \int_{\delta_s^-}^{\infty} \left(\delta_s^- + h_{q_s+1}(s) - h_{q_s}(s)\right) \nu^-_{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta),g_M(\delta))}(dy, ds)\Bigg]. \\
&\quad \leq \hat{H}(t,x,S,q) + \epsilon(T-t) + \mathbb{E}_{t,x,q,S}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta),g_M(\delta))}\Bigg[-\frac{1}{2\varphi_\alpha}\int_t^T \left(\frac{\alpha - \eta_s(\delta)}{\sigma}\right)^2 ds \\
&\qquad\qquad - \int_t^T \left(\mathcal{K}_+^{\varphi_\lambda,\varphi_\kappa}(g_s^{+\lambda}(\delta), g_{s,M}^{+\kappa}(\cdot;\delta))\, \mathbb{1}_{\varphi_\lambda \geq \varphi_\kappa} + \mathcal{K}_+^{\varphi_\kappa,\varphi_\lambda}(g_{s,M}^{+\kappa}(\cdot;\delta), g_s^{+\lambda}(\delta))\, \mathbb{1}_{\varphi_\lambda < \varphi_\kappa}\right) ds \\
&\qquad\qquad - \int_t^T \left(\mathcal{K}_-^{\varphi_\lambda,\varphi_\kappa}(g_s^{-\lambda}(\delta), g_{s,M}^{-\kappa}(\cdot;\delta))\, \mathbb{1}_{\varphi_\lambda \geq \varphi_\kappa} + \mathcal{K}_-^{\varphi_\kappa,\varphi_\lambda}(g_{s,M}^{-\kappa}(\cdot;\delta), g_s^{-\lambda}(\delta))\, \mathbb{1}_{\varphi_\lambda < \varphi_\kappa}\right) ds\Bigg].
\end{aligned}
$$

Therefore, the candidate function satisfies

$$
\begin{aligned}
&\hat{H}(t,x,S,q)+\epsilon(T-t)\\
&\geq \mathbb{E}_{t,x,q,S}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta),g_M(\delta))}\Bigg[\hat{H}(T,X_{T^-}^{\delta^\pm},S_{T^-},q_{T^-}^{\delta^\pm})+\frac{1}{2\varphi_\alpha}\int_t^T\left(\frac{\alpha-\eta_s(\delta)}{\sigma}\right)^2 ds\\
&\qquad+\sum_{i=\pm}\int_t^T\left(\mathcal{K}_i^{\varphi_\lambda,\varphi_\kappa}(g_s^{i\lambda}(\delta),g_{s,M}^{i\kappa}(\cdot;\delta))\,\mathbb{1}_{\varphi_\lambda\geq\varphi_\kappa}+\mathcal{K}_i^{\varphi_\kappa,\varphi_\lambda}(g_{s,M}^{i\kappa}(\cdot;\delta),g_s^{i\lambda}(\delta))\,\mathbb{1}_{\varphi_\lambda<\varphi_\kappa}\right)ds\Bigg]\\
&=\mathbb{E}_{t,x,q,S}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta),g_M(\delta))}\left[\hat{H}(T,X_T^{\delta^\pm},S_T,q_T^{\delta^\pm})+\mathcal{H}_{t,T}\left(\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta),g_M(\delta))|\mathbb{P}\right)\right]\\
&=\mathbb{E}_{t,x,q,S}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta),g_M(\delta))}\left[X_T^{\delta^\pm}+q_T^{\delta^\pm}(S_T-\ell(q_T^{\delta^\pm}))+\mathcal{H}_{t,T}\left(\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta),g_M(\delta))|\mathbb{P}\right)\right].
\end{aligned}
$$

Since this holds for one particular choice of admissible measure $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta(\delta),g_M(\delta))$, we have

$$
\hat{H}(t,x,S,q)+\epsilon(T-t)\geq\inf_{\mathbb{Q}\in\mathcal{Q}}\mathbb{E}_{t,x,q,S}^{\mathbb{Q}}\left[X_T^{\delta^\pm}+q_T^{\delta^\pm}(S_T-\ell(q_T^{\delta^\pm}))+\mathcal{H}_{t,T}(\mathbb{Q}|\mathbb{P})\right].
$$

This inequality holds for the arbitrarily chosen control $\delta_t^\pm$, therefore

$$
\begin{aligned}
\hat{H}(t,x,S,q)+\epsilon(T-t)&\geq\sup_{(\delta_s^\pm)_{t\leq s\leq T}\in\mathcal{A}}\inf_{\mathbb{Q}\in\mathcal{Q}}\mathbb{E}_{t,x,q,S}^{\mathbb{Q}}\left[X_T^{\delta^\pm}+q_T^{\delta^\pm}(S_T-\ell(q_T^{\delta^\pm}))+\mathcal{H}_{t,T}(\mathbb{Q}|\mathbb{P})\right]\\
&=H(t,x,S,q)\,,
\end{aligned}
$$

and letting $\epsilon\to 0$ we finally obtain

$$
\hat{H}(t,x,S,q)\geq H(t,x,S,q)\,. \tag{A.20}
$$

Step 5: Showing $\hat{H}(t,x,S,q)\leq H(t,x,S,q)$: Now, let $\delta^\diamond=(\delta_t^{\pm\diamond})_{0\leq t\leq T}$ be the control process defined in the statement of the theorem, and let η_t, $g_t^{\pm\lambda}$ and $g_t^{\pm\kappa}(y)$ be arbitrary such that they induce an admissible measure $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g) \in \mathcal{Q}^{\alpha,\lambda,\kappa}$. Then from Ito's lemma and the fact that h satisfies equation (23):

$$
\begin{aligned}
&\mathbb{E}_{t,x,q,S}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\hat{H}(T, X_{T^-}^{\delta^{\pm\diamond}}, S_{T^-}, q_{T^-}^{\delta^{\pm\diamond}})\right] \\
&= \hat{H}(t,x,S,q) + \mathbb{E}_{t,x,q,S}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\Bigg[\int_t^T \partial_t h_{q_s}(s)ds + \int_t^T \alpha q_s^{\delta^{\pm\diamond}} ds + \sigma \int_t^T q_s^{\delta^{\pm\diamond}} dW_s \\
&\qquad + \int_t^T \int_{\delta_s^{+\diamond}}^{\infty} \left(\delta_s^{+\diamond} + h_{q_s-1}(s) - h_{q_s}(s)\right) \nu^+_{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}(dy,ds)\Bigg] \\
&\qquad + \int_t^T \int_{\delta_s^{-\diamond}}^{\infty} \left(\delta_s^{-\diamond} + h_{q_s+1}(s) - h_{q_s}(s)\right) \nu^-_{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}(dy,ds)\Bigg] \\
&\geq \hat{H}(t,x,S,q) + \mathbb{E}_{t,x,q,S}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\Bigg[-\frac{1}{2\varphi_\alpha}\int_t^T \left(\frac{\alpha-\eta_s}{\sigma}\right)^2 ds \\
&\qquad - \sum_{i=\pm}\int_t^T \left(\mathcal{K}_i^{\varphi_\lambda,\varphi_\kappa}(g_s^{i\lambda}, g_s^{i\kappa})\, \mathbb{1}_{\varphi_\lambda \geq \varphi_\kappa} + \mathcal{K}_i^{\varphi_\kappa,\varphi_\lambda}(g_s^{i\kappa}, g_s^{i\lambda})\, \mathbb{1}_{\varphi_\lambda < \varphi_\kappa}\right) ds\Bigg].
\end{aligned}
$$

And so the candidate function satisfies

$$
\begin{aligned}
\hat{H}(t,x,S,q) &\leq \mathbb{E}_{t,x,q,S}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\Bigg[\hat{H}(T, X_{T^-}^{\delta^{\pm\diamond}}, S_{T^-}, q_{T^-}^{\delta^{\pm\diamond}}) + \frac{1}{2\varphi_\alpha}\int_t^T \left(\frac{\alpha-\eta_s}{\sigma}\right)^2 ds \\
&\qquad + \sum_{i=\pm}\int_t^T \left(\mathcal{K}_i^{\varphi_\lambda,\varphi_\kappa}(g_s^{i\lambda}, g_s^{i\kappa})\, \mathbb{1}_{\varphi_\lambda \geq \varphi_\kappa} + \mathcal{K}_i^{\varphi_\kappa,\varphi_\lambda}(g_s^{i\kappa}, g_s^{i\lambda})\, \mathbb{1}_{\varphi_\lambda < \varphi_\kappa}\right) ds\Bigg] \\
&= \mathbb{E}_{t,x,q,S}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[\hat{H}(T, X_T^{\delta^{\pm\diamond}}, S_T, q_T^{\delta^{\pm\diamond}}) + \mathcal{H}_{t,T}(\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)|\mathbb{P})\right] \\
&= \mathbb{E}_{t,x,q,S}^{\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)}\left[X_T^{\delta^{\pm\diamond}} + q_T^{\delta^{\pm\diamond}}(S_T - \ell(q_T^{\delta^{\pm\diamond}})) + \mathcal{H}_{t,T}(\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)|\mathbb{P})\right] .
\end{aligned}
$$

Since this holds for any arbitrary admissible measure $\mathbb{Q}^{\alpha,\lambda,\kappa}(\eta,g)$, we have

$$
\hat{H}(t,x,S,q) \leq \inf_{\mathbb{Q}\in\mathcal{Q}^{\alpha,\lambda,\kappa}} \mathbb{E}_{t,x,q,S}^{\mathbb{Q}}\left[X_T^{\delta^{\pm\diamond}} + q_T^{\delta^{\pm\diamond}}(S_T - \ell(q_T^{\delta^{\pm\diamond}})) + \mathcal{H}_{t,T}(\mathbb{Q}|\mathbb{P})\right] .
$$

Therefore,

$$
\begin{aligned}
\hat{H}(t,x,S,q) &\leq \sup_{(\delta_s^{\pm})_{t\leq s\leq T}\in\mathcal{A}} \inf_{\mathbb{Q}\in\mathcal{Q}^{\alpha,\lambda,\kappa}} \mathbb{E}_{t,x,q,S}^{\mathbb{Q}}\left[X_T^{\delta^{\pm}} + q_T^{\delta^{\pm}}(S_T - \ell(q_T^{\delta^{\pm}})) + \mathcal{H}_{t,T}(\mathbb{Q}|\mathbb{P})\right] \\
&= H(t,x,S,q)\,. \qquad \text{(A.21)}
\end{aligned}
$$

Combining (A.20) and (A.21) gives

$$\hat{H}(t,x,q,S) = H(t,x,q,S),$$

as desired. □

Appendix A.4. Proof of Proposition 5

Proof. Under the stated hypotheses, and assuming $\delta_q^{\pm*}(t) > 0$, substituting the optimizers from Proposition 3 into equation (23) gives the following:

$$\partial_t h_q + \alpha q - \frac{1}{2}\varphi_\alpha \sigma^2 q^2 + \frac{\xi^+}{\kappa} e^{-\kappa(-h_{q-1}+h_q)} \mathbb{1}_{q\neq \underline{q}} + \frac{\xi^-}{\kappa} e^{-\kappa(-h_{q+1}+h_q)} \mathbb{1}_{q\neq \overline{q}} = 0\,. \tag{A.22}$$

The substitution $\omega_q(t) = e^{\kappa h_q(t)}$ yields an ODE for ω_q:

$$\partial_t \omega_q + (\alpha\kappa q - \frac{1}{2}\varphi_\alpha \kappa \sigma^2 q^2)\omega_q + \xi^+ \omega_{q-1} \mathbb{1}_{q\neq \underline{q}} + \xi^- \omega_{q+1} \mathbb{1}_{q\neq \overline{q}} = 0\,.$$

In matrix form this reads

$$\partial_t \boldsymbol{\omega}(t) + \mathbf{A}\boldsymbol{\omega}(t) = \mathbf{0}\,.$$

The boundary condition $h_q(T) = -q\,\ell(q)$ implies $\omega_q(T) = e^{-\kappa q\,\ell(q)}$ and this first order coupled system has solution

$$\boldsymbol{\omega}(t) = e^{\mathbf{A}(T-t)}\boldsymbol{\omega}(T).$$

Equation (A.22) only holds for $t \in [t_0, T]$, so this solution only holds in the same interval. □

Appendix A.5. Proof of Proposition 6

PROOF. Let $\hat{h}_q(t) = h_{-q}(t)$. Then:

$$\begin{aligned}
&\partial_t \hat{h}_q - \varphi_\alpha \sigma^2 q^2 \\
&\quad + \sup_{\delta^+ \geq 0} \inf_{g^{+\lambda}} \inf_{g^{+\kappa} \in \mathcal{G}^+} \Bigg\{ \lambda \left[\int_{\delta^+}^\infty e^{g^{+\lambda} + g^{+\kappa}(y)} F(dy) \right] (\delta^+ + \hat{h}_{q+1}(t) - \hat{h}_q(t)) \\
&\qquad\qquad + \mathcal{K}_+^{\varphi_\lambda, \varphi_\kappa}(g^{+\lambda}, g^{+\kappa})\, \mathbb{1}_{\varphi_\lambda \geq \varphi_\kappa} + \mathcal{K}_+^{\varphi_\kappa, \varphi_\lambda}(g^{+\kappa}, g^{+\lambda})\, \mathbb{1}_{\varphi_\lambda < \varphi_\kappa} \Bigg\} \\
&\quad + \sup_{\delta^- \geq 0} \inf_{g^{-\lambda}} \inf_{g^{-\kappa} \in \mathcal{G}^-} \Bigg\{ \lambda \left[\int_{\delta^-}^\infty e^{g^{-\lambda} + g^{-\kappa}(y)} F(dy) \right] (\delta^- + \hat{h}_{q-1}(t) - \hat{h}_q(t)) \\
&\qquad\qquad + \mathcal{K}_-^{\varphi_\lambda, \varphi_\kappa}(g^{-\lambda}, g^{-\kappa})\, \mathbb{1}_{\varphi_\lambda \geq \varphi_\kappa} + \mathcal{K}_-^{\varphi_\kappa, \varphi_\lambda}(g^{-\kappa}, g^{-\lambda})\, \mathbb{1}_{\varphi_\lambda < \varphi_\kappa} \Bigg\} = 0\,.
\end{aligned}$$

Since $\delta^\pm$, $g^{\pm\lambda}$, and $g^{\pm\kappa}$ appearing inside the optimization are dummy variables, their labels can be changed by making the substitution $\pm \to \mp$, and the hypotheses of the proposition also imply $\mathcal{K}_- = \mathcal{K}_+$. Then the functional form of the ODE for $\hat{h}_q$ is equivalent to that of h_q, and they also share the same terminal conditions, $\hat{h}_q(T) = h_q(T) = -q\,\ell(q)$ so we have $h_{-q}(t) = \hat{h}_q(t) = h_q(t)$. Then from the feedback form of $\delta_q^{\pm *}(t)$ in Proposition 3, it is clear that $\delta_q^{+*}(t) = \delta_{-q}^{-*}(t)$, as desired. □

Appendix A.6. Proof of Proposition 7

PROOF. We prove the result for the sell spread and for $\underline{q} < q \leq 0$ only, as the proof for the buy spread and other values of q follows analogously.

Define $\chi_q(t;\varphi) = h_q(t;\varphi) - h_{q-1}(t;\varphi)$ for $\underline{q} < q \leq \overline{q}$, where $h_q(t;\varphi)$ is the solution to equation (23) corresponding to drift ambiguity parameter φ. From (A.16) in the proof of Proposition 3 we have

$$\begin{aligned}
\partial_t \chi_q = &- \varphi \sigma^2 (1 - 2q) - \sup_{\delta^+ \geq 0} \left\{ \frac{\lambda}{\varphi_\lambda} \left(1 - \exp\left\{ \frac{\varphi_\lambda}{\varphi_\kappa} \log(1 - e^{-\kappa\delta^+} + e^{-\kappa\delta^+ - \varphi_\kappa(\delta^+ - \chi_q)}) \right\} \right) \right\} \\
&- \sup_{\delta^- \geq 0} \left\{ \frac{\lambda}{\varphi_\lambda} \left(1 - \exp\left\{ \frac{\varphi_\lambda}{\varphi_\kappa} \log(1 - e^{-\kappa\delta^-} + e^{-\kappa\delta^- - \varphi_\kappa(\delta^- + \chi_{q+1})}) \right\} \right) \right\} \mathbb{1}_{q \neq \overline{q}} \\
&+ \sup_{\delta^+ \geq 0} \left\{ \frac{\lambda}{\varphi_\lambda} \left(1 - \exp\left\{ \frac{\varphi_\lambda}{\varphi_\kappa} \log(1 - e^{-\kappa\delta^+} + e^{-\kappa\delta^+ - \varphi_\kappa(\delta^+ - \chi_{q-1})}) \right\} \right) \right\} \mathbb{1}_{q-1 \neq \underline{q}} \\
&+ \sup_{\delta^- \geq 0} \left\{ \frac{\lambda}{\varphi_\lambda} \left(1 - \exp\left\{ \frac{\varphi_\lambda}{\varphi_\kappa} \log(1 - e^{-\kappa\delta^-} + e^{-\kappa\delta^- - \varphi_\kappa(\delta^- + \chi_q)}) \right\} \right) \right\},
\end{aligned} \tag{A.23}$$

with terminal condition $\chi_q(T;\varphi) = 0$. In the proof of Proposition 6 we also show that under the stated hypotheses we have the symmetry result $h_q(t;\varphi) = h_{-q}(t;\varphi)$, which implies $\chi_q(t;\varphi) = -\chi_{-q+1}(t;\varphi)$. Using this symmetry property, we further have $\chi_1 = -\chi_0$. Hence, we can rewrite (A.23) by treating the $q = 0$ case separately as follows

$$\begin{aligned}
\partial_t \chi_q = & -\varphi\sigma^2(1-2q) - \sup_{\delta^+\geq 0}\left\{\frac{\lambda}{\varphi_\lambda}\left(1-\exp\left\{\frac{\varphi_\lambda}{\varphi_\kappa}\log(1-e^{-\kappa\delta^+}+e^{-\kappa\delta^+-\varphi_\kappa(\delta^+-\chi_q)})\right\}\right)\right\} \\
& - \sup_{\delta^-\geq 0}\left\{\frac{\lambda}{\varphi_\lambda}\left(1-\exp\left\{\frac{\varphi_\lambda}{\varphi_\kappa}\log(1-e^{-\kappa\delta^-}+e^{-\kappa\delta^--\varphi_\kappa(\delta^-+\chi_{q+1})})\right\}\right)\right\}\mathbb{1}_{q\neq\overline{q}}\mathbb{1}_{q\neq 0} \\
& - \sup_{\delta^-\geq 0}\left\{\frac{\lambda}{\varphi_\lambda}\left(1-\exp\left\{\frac{\varphi_\lambda}{\varphi_\kappa}\log(1-e^{-\kappa\delta^-}+e^{-\kappa\delta^--\varphi_\kappa(\delta^--\chi_0)})\right\}\right)\right\}\mathbb{1}_{q=0} \\
& + \sup_{\delta^+\geq 0}\left\{\frac{\lambda}{\varphi_\lambda}\left(1-\exp\left\{\frac{\varphi_\lambda}{\varphi_\kappa}\log(1-e^{-\kappa\delta^+}+e^{-\kappa\delta^+-\varphi_\kappa(\delta^+-\chi_{q-1})})\right\}\right)\right\}\mathbb{1}_{q-1\neq\underline{q}} \\
& + \sup_{\delta^-\geq 0}\left\{\frac{\lambda}{\varphi_\lambda}\left(1-\exp\left\{\frac{\varphi_\lambda}{\varphi_\kappa}\log(1-e^{-\kappa\delta^-}+e^{-\kappa\delta^--\varphi_\kappa(\delta^-+\chi_q)})\right\}\right)\right\},
\end{aligned} \tag{A.24}$$

In this form, the ODEs for $\underline{q} < q \leq 0$ are decoupled from those of $0 < q \leq \overline{q}$. Focusing on $\underline{q} < q \leq 0$, we can write the collection of ODEs in matrix form as follows

$$\partial_t \boldsymbol{\chi} = \boldsymbol{G}(\boldsymbol{\chi};\varphi)\,,$$

for a vector valued function $\boldsymbol{G}(\cdot;\cdot)$, parameterized by its second argument, with the property that (i) $\boldsymbol{G}(\boldsymbol{\chi};\varphi_\alpha) > \boldsymbol{G}(\boldsymbol{\chi};\varphi'_\alpha)$ componentwise, and (ii)

$$\frac{\partial G_q}{\partial \chi_{q'}} \leq 0, \quad q \neq q'\,,$$

(this follows from (A.17) and (A.18) in the proof of Proposition 3). Thus, due to a classical comparison principle (see Ważewski (1950)) we have $\chi_q(t;\varphi'_\alpha) \geq \chi_q(t;\varphi_\alpha)$ which implies $\delta_q^{+*}(t;\varphi'_\alpha) \geq \delta_q^{+*}(t;\varphi_\alpha)$ for $\underline{q} < q \leq 0$. □

Appendix B. Outline of Numerical Scheme

Here we outline a numerical scheme for solving equation (23), and which we employ to produce the figures in Section 4. The only exception to this is in Figure 5 where we apply Proposition 5 to solve for the optimal depth exactly since this example satisfies the required symmetry conditions (we also use this closed-form solution to compute the ambiguity-neutral spreads, which use the same set of parameters as those in Figure 1).

As mentioned in the proof of Proposition 3 (see equation (A.16)), the coupled differential equations (23) can be written in the form:

$$\partial_t \boldsymbol{h} = \boldsymbol{F}(\boldsymbol{h}) ,$$

with terminal condition $h_q(T) = -q\,\ell(q)$. To approximate the solution to this system, we employ a fully explicit finite-difference method. In particular, let $\Delta t = \frac{T}{N}$ and $t_n = n\,\Delta t$, denote the approximation of the classical solution at time t_n by $\hat{\boldsymbol{h}}(n)$, we then recursively compute $\hat{\boldsymbol{h}}(n)$ by

$$\begin{aligned} \hat{\boldsymbol{h}}(n) &= \hat{\boldsymbol{h}}(n+1) - \boldsymbol{F}(\hat{\boldsymbol{h}}(n+1))\Delta t , \qquad 0 \le n < N , \\ \hat{h}_q(N) &= -q\,\ell(q) . \end{aligned}$$

References

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

Avellaneda, M. and S. Stoikov (2008, November). High-frequency trading in a limit order book. Quantitative Finance 8, 217–224.

Bayraktar, E. and M. Ludkovski (2011). Optimal trade execution in illiquid markets. Mathematical Finance 21(4), 681–701.

Bayraktar, E. and M. Ludkovski (2014). Liquidation in limit order books with controlled intensity. Mathematical Finance 24(4), 627–650.

Cartea, Á. and S. Jaimungal (2013). Modeling asset prices for algorithmic and high frequency trading. Forthcoming in Applied Mathematical Finance.

Cartea, Á. and S. Jaimungal (2015). Risk metrics and fine tuning of high-frequency trading strategies. Mathematical Finance 25(3), 576–611.

Cartea, A., S. Jaimungal, and D. Kinzebulatov (2016). Algorithmic trading with learning. International Journal of Theoretical and Applied Finance 19(04), 1650028.

Cartea, Á., S. Jaimungal, and J. Penalva (2015). Algorithmic and High-Frequency Trading. Cambridge University Press.

Cartea, Á., S. Jaimungal, and J. Ricci (2014). Buy low, sell high: A high frequency trading perspective. SIAM Journal on Mathematical Finance 5(1), 415–444.

Fleming, W. H. and H. M. Soner (2006). Controlled Markov Processes and Viscosity Solutions (2nd ed.). Springer.

Fodra, P. and M. Labadie (2012). High-frequency market-making with inventory constraints and directional bets. Working papers, HAL.

Gatheral, J., A. Schied, and A. Slynko (2012). Transient linear price impact and Fredholm integral equations. Mathematical Finance 22(3), 445–474.

Guéant, O. and C.-A. Lehalle (2015). General intensity shapes in optimal liquidation. Mathematical Finance 25(3), 457–495.

Guéant, O., C.-A. Lehalle, and J. Fernandez-Tapia (2012). Dealing with the inventory risk: a solution to the market making problem. Mathematics and Financial Economics, 1–31.

Guilbaud, F. and H. Pham (2013). Optimal high frequency trading with limit and market orders. Quantitative Finance 13(1), 79–94.

Guilbaud, F. and H. Pham (2015). Optimal high-frequency trading in a pro rata microstructure with predictive information. Mathematical Finance 25(3), 545–575.

Hansen, L. P. and T. J. Sargent (2007). Robustness. Princeton.

Hansen, L. P. and T. J. Sargent (2011). Robustness and ambiguity in continuous time. Journal of Economic Theory 146(3), 1195 – 1223. Incompleteness and Uncertainty in Economics.

Hansen, L. P., T. J. Sargent, G. Turmuhambetova, and N. Williams (2006). Robust control and model misspecification. Journal of Economic Theory 128(1), 45 – 90.

Hugonnier, J., F. Pelgrin, and P. St-Amour (2013). Health and (Other) Asset Holdings. Review of Economic Studies 80(2), 663–710.

Jacod, J. and A. N. Shiryaev (1987). Limit theorems for stochastic processes. Grundlehren der mathematischen Wissenschaften. Springer-Verlag.

Jaimungal, S. and D. Kinzebulatov (2014). Optimal Execution with a Price Limiter. Risk July.

Jaimungal, S. and G. Sigloch (2012). Incorporating risk and ambiguity aversion into a hybrid model of default. Mathematical Finance 22(1), 57–81.

Kharroubi, I. and H. Pham (2010). Optimal portfolio liquidation with execution cost and risk. SIAM Journal on Financial Mathematics 1, 897–931.

Large, J. (2007). Measuring the resiliency of an electronic limit order book. Journal of Financial Markets 10(1), 1–25.

Lim, A. E. B. and J. G. Shanthikumar (2007). Relative Entropy, Exponential Utility, and Robust Dynamic Pricing. Operations Research 55(2), 198–214.

Maenhout, P. J. (2004). Robust portfolio rules and asset pricing. Review of Financial Studies 17(4), 951–983.

Pham, H. (2009). Continuous-time Stochastic Control and Optimization with Financial Applications (1st ed.), Volume 61 of Stochastic Modelling and Applied Probability. Berlin Heidelberg: Springer-Verlag.

Schied, A. (2013). Robust strategies for optimal order execution in the Almgren-Chriss framework. Applied Mathematical Finance 20(3), 264–286.

Skiadas, C. (2013). Smooth Ambiguity Aversion Toward Small Risks and Continuous-Time Recursive Utility. Journal of Political Economy 121(4), 775–792.

Uppal, R. and T. Wang (2003, December). Model misspecification and underdiversification. The Journal of Finance 58(6), 2465–2486.

Ważewski, T. (1950). Systèmes de équations et des inégalités différentielles ordinaires aux deuxièmes membres monotones et leurs application. In Annales de la Société polonaise de mathématique, pp. 112–166.