Robust Price Formation*

Johannes Hörner† Stefano Lovo‡ Tristan Tomala§

December 10, 2015

Abstract

We analyze security price formation in a dynamic setting in which long-lived dealers repeatedly compete for trading with short-lived retail traders. We characterize equilibria in which dealers’ dynamic pricing strategies are optimal no matter the private information each dealer may possess. Thus, our model’s predictions are robust to different specifications of dealers’ information structure. These equilibria reconcile in a single and parsimonious model price dynamics that are reminiscent of well-known stylized facts: excess price volatility, price/trading-flow correlation, stochastic volatility and inventory-related trading.

Keywords: Financial Market Microstructure, Informed Dealers, Price Volatility, Belief-free Equilibria.

JEL Codes: G1, G12, C72, C73.

---

*We thank Vincent Fardeau, Thierry Foucault and Ioanid Rosu for their useful comments as well as the seminar audiences at HEC Paris, EDHEC, Warwick Business School, the SEAT 2012 meeting in Brisbane and the EFA 2012 Meeting in Copenhagen, the Roy-Adres Seminar. S. Lovo and T. Tomala gratefully acknowledge financial support from the HEC Foundation. J. Hörner gratefully acknowledges financial support from NSF Grant SES 092098.

†Yale University, 30 Hillhouse Ave., New Haven, CT06520, USA. johannes.horner@yale.edu.
‡HEC,Paris and GREGHEC, 78351 Jouy-en-Josas, France. lovo@hec.fr.
§HEC, Paris and GREGHEC, 78351 Jouy-en-Josas, France. tomala@hec.fr.
Introduction

In this paper, we consider a class of market microstructure models in which some long-lived market participants (“dealers”) repeatedly trade a risky asset with short-lived market participants (“traders”). For this class of models, we characterize equilibria that are robust to all possible specifications of information asymmetries across dealers, no matter how simple or complex these asymmetries could be.

It is standard in the theoretical literature on financial market microstructure to model the stock price-formation process as the result of a strategic interaction across market participants that are possibly asymmetrically informed. The established approach consists of specifying the information structure across agents and then solving for a Bayesian equilibrium of the incomplete information game that stems from this specification. However, information asymmetries are not directly observable in practice. Hence, it is difficult to assess the extent to which a given model’s assumptions about each agent’s private information captures real-life situations. This would not be a problem if we knew that market microstructure models’ predictions are robust to changes in information structures. Only in this case, the predictions would be reliable even if based on unverifiable assumptions. Unfortunately, tractability imposes strong informational assumptions about who in the market has superior private information and who does not. When these assumptions are relaxed, solving for the Bayesian equilibrium becomes a formidable task, particularly when dealers or market-makers that repeatedly meet in the market are asymmetrically informed.\(^1\) As a result, the current theoretical literature is silent about the robustness of the predictions of canonical microstructure models to changes in the information environment. Additionally, the canonical theoretical assumption that dealers are equally uninformed about

\(^1\)In this case, in a Bayesian equilibrium, at each point in time, each dealer anticipates how its behavior affects its current expected payoff as well as each competing dealer’s posterior beliefs and future behaviors. The problem is even more complex if a dealer is uncertain about its competitors’ prior beliefs or about who is more or less informed.
market fundamentals is not supported by the empirical analysis by Ellis, Michaely, and O’Hara (2002), who claim that dealers have access to different sources of information and that they need not be well aware of other dealers’ sources of private information. This suggests that even if information structures were observable, and they are not, they would be too complex to lead to tractable models.

In this paper, we overcome this problem by focusing on those sub-game perfect equilibria that are robust to specifications of the asymmetries of information across the market participants who repeatedly meet in the same market, such as brokers, dealers and market-makers. Thus, our model’s predictions do not rely on simplifying assumptions about the information structure of such market participants.

For this purpose, we consider a class of dynamic financial market microstructure models in which risk-neutral long-lived professionals (dealers, hereafter) interact with an infinite sequence of short-lived traders. To keep the model general, we make minimal assumptions about the trading protocol and the form of uncertainty. Namely, our trading protocol embeds quote-driven markets and limit-order markets as special cases. Market participants face uncertainty about a state of Nature that affects both the fundamental value of the asset and the composition of the population of traders.

For this class of models, we focus on equilibria for which the same dealers’ dynamic pricing strategies remains optimal no matter each dealer’s information and beliefs about the economy’s fundamentals. More specifically, the same dealer’s strategies form a sub-game perfect equilibrium ex-post, i.e., no matter the state of Nature. This guarantees that the predictions of the model are robust to the extent that they do not rely on any specific assumptions about each dealer’s private information about the fundamentals. We will refer to these equilibria as ‘belief-free equilibria’ (henceforth BFE).²

We show that surprisingly, BFE exist under very mild conditions: First, there is room for trade across traders, a condition present in all market microstructure models to avoid the no-trade theorem. Second, the discount factor between two trading rounds is large enough, which is a condition that is naturally met by market microstructure economies in which trading frequency is intra-day.

We show that BFE display the following properties. Over time, dealers make strictly positive profits no matter the fundamental asset value. Dealers’ profits come from the intermediation of traders’ demand and supply and not from large positions resulting from speculative bets on the asset being mispriced. The only information that is necessary and sufficient to achieve intermediation profits is the one about traders’ demand and supply. Tautologically, this is the information that any dealer can gather through a consistent probing and monitoring of traders’ behavior. As a consequence, in a BFE, each dealer can ignore the private information he might have. Either because this information can be statistically learned from traders’ behavior or because this information will never be backed by the hard evidence from traders’ behavior and hence is useless to intermediate their order flow. The process of consistently probing traders demand gives rise to excess volatility in the asset price.

In greater detail, we first characterize four necessary conditions that a price formation strategy must satisfy to form a BFE. This translates into the following four testable restrictions on the equilibrium outcomes that are altogether unique to our robust price formation theory.

1) Because a BFE must be an equilibrium even when dealers are uninformed, dealers’ strategies must be measurable with respect to public information, and namely, the information that can be gathered from the observation of traders’ order flow. This shows that the canonical prediction that public news and trading volume are the main drivers of changes in prices is robust to changes in dealers’ information structures. The relation between trading volume and prices has been extensively empirically documented for several markets.\(^3\)

\(^3\)See for instance Chordia, Roll and Subrahmanyam (2002) and Boehmer and Wu (2008)), bond markets
2) Each dealer can always guaranty no profit from abstaining from trading. Thus, robustness requires that no matter the state of Nature, each dealer makes strictly positive profit over time. That is, dealers might lose money over short periods, but their average long-run profit is strictly positive independent of the state of Nature. Thus, the canonical hypothesis that dealers’ expected per-trade profit is zero is not a robust feature of price formation.\(^4\) The prediction that dealers make positive profits is consistent with the empirical findings that have questioned the competitiveness of dealer pricing.\(^5\)

3) Because no large inventory comes for free, for some beliefs about the asset fundamentals, a large inventory necessarily translates into a negative profit. Thus, robustness requires that dealers maintain balanced inventories even if they are risk-neutral. The smaller the average trading volume and the larger the residual uncertainty on the asset value, the tighter the bounds on dealers’ inventories. In other words, dealers’ inventories are mean-reverting. This contrasts with the view that (absent risk aversion or institutional constraints on inventory size) inventory levels should not affect dealers’ behavior.\(^6\) Thus, our theory provides an alternative explanation of the empirical evidence of market-makers mean-reverting their inventories.\(^7\)

4) In a BFE, even if market participants are Bayesian, it is not the case that equilibrium quotes reflect Bayesian beliefs about fundamentals. Namely, price sensitivity to trading volume does not fade away as public information accumulates. Thus, long-term price volatility remains large even without exogenous shocks to fundamentals. This generates patterns that are consistent with the

\(^4\)The zero-profit condition is often justified through the hypothesis of the free entry of new dealers. However, it requires the additional assumption that incumbents and entrant dealers have exactly the same information. Without this simplifying informational assumption, it is a priori unclear what free entry would imply for dealer’s profits.


\(^6\)For instance, in Ho and Stoll (1981), balanced inventory results from dealers’ risk aversion, whereas in Gromb and Vayanos (2002) and Brunnermeier and Pedersen (2009), it results from the dealers’ institutional inability to take a position beyond a certain size. Our model displays neither factor.

\(^7\)See, for example, Madhavan and Smidt (1993), Hansch, Naik, and Viswanathan (1999), Reiss and Werner (1998), and Naik and Yadav (2003).
phenomena of excess and stochastic price volatility, two well-known and empirically established properties of stock prices.\(^8\)

Thus, BFE allow the explanation of four well-documented features of stock prices. While each feature can be explained by existing models, we are not aware of any single model that explains all of them simultaneously.

We then provide sufficient conditions about, first, the key assumptions on the market microstructure economy, and, second, the properties of dealers’ strategies, guaranteeing that these strategies form a BFE.

We assume that a state of Nature \(\omega \in \Omega\) might affect both the fundamental asset value and the composition of the population of traders and that it does not change over time. We then consider the information partition \(\hat{\Omega}\) on \(\Omega\) that dealers can derive from the statistical observation of traders’ behavior. That is, two states, \(\omega\) and \(\omega'\), belong to the same element \(\hat{\omega} \in \hat{\Omega}\) if the way traders react to dealers’ actions is the same in the two states. We assume that \(\hat{\Omega}\) is a finite partition and that for each given set \(\hat{\omega} \in \hat{\Omega}\), there is room for trade across traders. That is, there are transaction prices at which some traders would buy and others would sell the asset. This implies that for each \(\hat{\omega}\), dealers have an appropriate choice of actions, allowing them to intermediate traders’ demand and make profits without taking substantial net positions in the asset.

Given the above assumptions, we can show that if dealers are patient enough, the following dealers’ Markov strategy is a BFE. In essence, dealers’ strategy consists of two ingredients: periods where they probe traders demand to learn the true \(\hat{\omega}\) (exploring phases) and periods in which they exploit the information from the exploring phase and make profits through the intermediation of traders’ demand and supply (exploiting phases). The Markov variable affecting dealers’ actions is what we call the market measure, that is, a probability measure in \(\Delta\hat{\Omega}\). The

\(^8\)The first papers providing empirical evidence that stock price volatility does not coincide with commensurate volatility in corporations’ fundamentals are Shiller (1981) and LeRoy and Douglas (1981).
market measure is updated from the observation of how traders react to dealers’ actions. We say that a market measure points at a given \( \hat{\omega} \in \hat{\Omega} \) when it attaches large enough probability to \( \hat{\omega} \). Whenever the market measure points at any given \( \hat{\omega} \), dealers set their actions to profitably intermediate traders’ demand in state \( \hat{\omega} \). When this happens, we say that dealers are in an exploiting phase. When the market measure is not concentrated on any single \( \hat{\omega} \), dealers set actions to prompt informative order flows from traders. When this happens, we say that dealers are in an exploring phase. The way the market measure is updated leads it to point at the true \( \hat{\omega} \) relatively fast and with high probability, no matter the past history of trade. Then, no matter a dealer’s belief about the true state and no matter the past history, each dealer expects the equilibrium to move relatively fast to the exploiting phase corresponding to the true state. Therefore, each dealer expects future profits to be strictly positive no matter his beliefs or past history. To prove that dealers have no incentive to deviate, we show that there are strategies that will be played after a deviation that punish the deviating dealer while rewarding the other dealers.

Despite the restrictiveness of robust equilibria, there remains considerable leeway in specifying BFE. Rather than delineating precisely the scope of these equilibria, we take advantage of this leeway to focus on a tractable subset that accords with additional regularities documented in the empirical literature. In the second part of the paper, we illustrate the functioning of a Markov BFE in the simple framework of the Glosten and Milgrom (1985) model, and we contrast its predictions with those of the textbook canonical equilibrium.

The applications of repeated games to the market microstructure that are closest to our work are Dutta and Madhavan (1997), Benveniste et al. (1992) and Desgranges and Foucault (2005). These papers assume either no informational asymmetry or short-lived informational asymmetries. Here, instead, the state of nature is chosen once and for all so that a dealer owning some private information might possibly take advantage of it over a long horizon.\(^9\)

\(^9\)The same results hold if the frequency of trading is high compared to the frequency with which the state of
Few theoretical papers analyze the effect of asymmetric information among dealers. Even fewer do so within a dynamic framework. Some static models where dealers or, more generally, liquidity providers are asymmetrically informed are Roël (1988), Bloomfield and O'Hara (2000), de Frutos and Manzano (2005) and Boulatov and George (2010). Within a dynamic framework, Moussa Saley and De Meyer (2003) and Calcagno and Lovo (2006) study the case of one better-informed price maker. De Meyer (2010) considers the case of two-sided incomplete information. However, their findings are not robust to the specification of dealers’ information structure. The results obtained by Du and Zhu (2012) are closer in spirit to our work. Within the framework of a double auction, they show that for a specific additive functional form of bidders’ values, the static auction has an ex post equilibrium and that this property extends to the repeated auction, giving rise to an almost belief-free equilibrium.

The paper is organized as follows. Section 1 introduces the general framework. Section 2 describes some salient properties of the one-shot trading game. Section 3 defines robust equilibria in the repeated game. Sections 4 and 5 present necessary and sufficient conditions, respectively, for a strategy profile to form a BFE of the repeated game. Section 6 presents an example based on the Glosten and Milgrom model and compares BFE with the canonical equilibrium. Section 7 discusses extensions to imperfect monitoring about dealers’ actions, non-stationary states of nature and dealers’ strategies based on private information. Section 8 concludes. All proofs are in the Appendix.

1 A Model of Price Formation

A risky asset is exchanged for money over an infinite number of periods $t = 1, 2, \ldots$. There are two groups of market participants. The first consists of professional financial intermediaries, such as dealers, market-makers and brokers. They are finitely many and consistently monitor nature changes.
and participate in the market.\footnote{For example, for the NASDAQ, Ellis, Michaely, and O’Hara (2002) show that the average number of market-makers is between 6 and 10 and that only a fraction of these is actually active in the market.} We call this group of market participant dealers and model them as \( n \) infinitely lived risk-neutral agents, with \( n \geq 1 \) kept finite. The second group consists of household, fund managers and institutional investors that occasionally come to the market to rebalance their portfolios and/or exploit some private information they have about the asset fundamentals. We call this group traders and model them as an infinite sequence of short-lived, possibly risk-averse agents.

At time 0 and once and for all, Nature choses the state \( \omega \) in the set \( \Omega \). We allow the state to affect the economy in two dimensions: the fundamental value of the asset and the composition of the population of traders. We denote by \( W(\omega) \) the fundamental value of the asset in state \( \omega \). Assume that \( W(\omega) = v(\omega) + \tilde{e}(\omega) \), where \( \tilde{v} \) and \( \tilde{e} \) are independently distributed and bounded. Furthermore, we assume \( E[\tilde{e}] = 0 \) and \( \text{Var} [\tilde{e}] > 0 \). We denote \( \bar{e} := \max_\omega |\tilde{e}(\omega)| \). As in Back and Baruch (2004), a public release of information takes place at a random time \( \tau \), and conditional on it not having occurred yet, the probability that it occurs in the next period is constant. After the public announcement, all positions are liquidated at price \( W(\omega) \).

In every period, a randomly selected trader comes to the market, trades and leaves the market. A trader is characterized by its type \( (g, y, c) \), where \( g : \mathbb{R} \rightarrow \mathbb{R} \) is the trader’s utility that is an increasing and concave function of his post-trade wealth, \( y \) is the initial endowment of the risky asset, and \( c \) is the initial endowment of cash. We assume that traders’ types belong to a compact set \( \Theta \). Given any subset \( \theta \subseteq \Theta \), we denote with \( Z(\omega, \theta) \in [0, 1] \) the exogenous probability that if the state of nature is \( \omega \), period \( t \) trader’s type belongs to \( \theta \).

Traders know the realization of the \( \tilde{v} \) component of the asset value \( W(\omega) \) but not the realization of the \( \tilde{e} \) component. There is no correlation between the distribution of traders’ types and \( \tilde{e} \). This implies that information about the composition of traders’ population is useless to learn \( \tilde{e} \). Thus, \( \tilde{e} \) can be interpreted as the residual uncertainty about the fundamental asset value after
aggregating all possible information from the population of traders.

We make no assumption regarding the private information that a given dealer might have about the state \( \omega \). In a belief-free equilibrium, defined below, each dealer’s strategy must be optimal no matter the true state \( \omega \). Hence, it is optimal no matter a dealer’s beliefs about \( \omega \).

\section{Stage trading round}

\subsection{Trading protocol}

In every period \( t \), trading unfolds as follows. First, dealers simultaneously choose their actions. We use \( A_i \) to denote the finite set of actions available to dealer \( i \). A dealer’s action profile \( a \) is an element of the set \( A := \times_i A_i \). Second, a trader is drawn from a large population of traders and choses a reaction \( s \) in the finite set \( S \). Third, any given action-reaction profile \( (a, s) \in A \times S \) translates into transactions (possibly nil) across market participants. Fourth, the trader leaves the market.

The trading protocol specifies for any given \( (a, s) \in A \times S \) and any market participant \( j \), the quantities \( Q_j(a, s) \) and \( P_j(a, s) \) of shares of the risky asset and money, respectively, that are transferred to \( j \). That is, if the state of nature is \( \omega \), an action-reaction profile \( (a, s) \) leads to a change in agent \( j \)'s wealth given by

\[ W(\omega)Q_j(a, s) + P_j(a, s). \]  

(1)

Transfers of cash and assets belong to finite grids \( G \) and \( G_q \), respectively, which include 0. In the following, if \(-\frac{P_j(a,s)}{Q_j(a,s)} = p > 0\), we say that the agent \( j \) buys at price \( p \) if \( Q_j(a,s) > 0 \) and that the agent \( j \) sells at price \( p \) if \( Q_j(a,s) < 0 \).

We assume that the trading protocol is such that no agent can be forced to trade. More precisely, for any positive price \( p \in G \), and quantity \( q \in G_q \), each agent \( j \) can choose his action
so that no matter what the other agents do, if he trades, he trades no more than $q$ shares, if he buys, he does not pay more then $p$ per share, and if he sells, he does not sell for less than $p$ per share. Note that because $0 \in G_q$, each market participant can unilaterally decide not to trade.

2.1.1 Applications

Here, we illustrate how some of the trading protocols analyzed in the literature fit into this framework.

**Quote-driven markets (Glosten and Milgrom (1985))** The set of actions available to each dealer is the set of bid and ask quotes to be chosen on the price grid $G$. Formally, $A_i = G \times G$. The trader observes the dealers’ quotes and chooses whether to buy one unit, sell one unit, or not trade, that is, $G_q = \{-1, 0, 1\}$. Thus, $S = \{\text{sell, no-trade, buy}\}$ is the set of possible market orders that can be chosen by a trader. A trader’s market order is executed against the best dealers’ quotes.

**Quote-driven market (Biais, Martimort and Rochet (2000))** The set of actions available to each dealer is a schedule $T(\cdot)$, which specifies his willingness to trade $q \in G_q$ shares of the asset against the transfer of $T(q)$, on the grid $G$. Thus, $A_i = [G]^{M_q}$, where $M_q$ denotes the cardinality of $G_q$. The trader observes dealers’ schedules and chooses how many shares to trade with each dealer; thus, $S = [G_q]^n$.

**Limit order markets** There is no agreed-upon way to model limit order markets (see, for example, Foucault (1999), Goettler et al. (2005), Foucault at al. (2005), Rosu (2009)). We present one possible specification that captures the functioning of a limit order market. At the beginning of period $t$, first, each dealer can submit a limit order that enters the book at the
specified price. These limit orders form the initial book. Second, a time $t$ trader chooses whether to submit a market order that trades against the initial book and/or a limit order that enters the book. This changes the initial book into the interim book. Third, each dealer can submit a market order that trades against the interim book. Fourth, all limit orders that are not executed are cancelled.\footnote{This fourth step is without loss of generality because the short-lived trader would cancel its order, whereas dealers can immediately post just-cancelled limit orders at the beginning of the following trading round.} The set of actions for both dealers and traders is $A_i = S = G \times G_q \times G_q$. Namely, $\{(p_j, q_j), m_j\} \in A_j$ specifies agent $j$’s limit and market orders.

2.2 Traders and Dealers

**Traders** Traders come to the market for both speculative and hedging reasons. The trader knows the realization of $\bar{v}$ but has no information about $\bar{e}$. He trades to maximize the expected utility of his post-trade wealth and then leaves the market. Formally, if $v(\omega) = v$, in time $t$, a dealer’s action is $a \in A$ and the time $t$ trader’s type is $(g, y, c)$, then the trader’s reaction is

$$D(v, a) := \arg\max_{s \in S} E [g((v + \bar{e})(Q_{Tr}(a, s) + y) + P_{Tr}(a, s) + c)],$$

where we set $j = Tr$ for the trader. Let $\theta(v, a, s) \subseteq \Theta$ be the set of traders’ types such that $D(v, a) = s$; where $s \in S$. Then,

$$F(\omega, a, s) := Z(\omega, \theta(v(\omega), a, s))$$

is the probability that time $t$ trader chooses reaction $s$, given that the dealer’s action is $a$ and that the state of Nature is $\omega$. The distribution function $F : \Omega \times A \to \Delta S$ depends on the specific assumptions about the set $\Theta$ and the distribution of type $Z$.

Because the set $\Theta$ of traders’ type is compact, $E[\bar{e}] = 0$ and $Var[\bar{e}] > 0$, traders never buy the asset at a price that is too high nor sell at a price that is too low when compared to $v(\omega)$.
However, traders who are risk-averse and who initially have a sufficiently short or long position in the risky asset will be willing to buy at price strictly larger than $v(\omega)$ or to sell at price strictly smaller than $v(\omega)$, respectively. We assume that $\Theta$ and $Z$ are such that if traders were to meet simultaneously in the market, there would be room for trade in all states $\omega$. In other words, there is always a positive probability that the time $t$ trader is willing to buy or to sell the asset at price $p$ as long as $p$ is not too far from $v(\omega)$. Formally,

**Assumption: Room for trade (RFT)** The set $\Theta$ and the probability distribution $Z$ are such that there are $\overline{p} \geq \underline{p} > 0$ such that for all $\omega \in \Omega$,

1. If $(a, s)$ translates into the trader buying at price $p < v(\omega) + \underline{p}$, then $F(\omega, a, s) > 0$.
2. If $(a, s)$ translates into the trader selling at price $p > v(\omega) - \underline{p}$, then $F(\omega, a, s) > 0$.
3. If $(a, s)$ translates into the trader buying at price $p > v(\omega) + \overline{p}$, then $F(\omega, a, s) = 0$.
4. If $(a, s)$ translates into the trader selling at price $p < v(\omega) - \overline{p}$, then $F(\omega, a, s) = 0$.

Note that the RFT assumption has its parallel in most assumptions made in the market microstructure literature to avoid the no-trade theorem (Milgrom and Stokey (1982)), such as, for example, the presence of noise/liquidity traders.

**Dealers** There are $n$ dealers who are long-lived and risk-neutral. We will call *reward* a dealer’s profit in a single trading round and *payoff* the discounted sum of the current and future rewards. We make no assumption about the private information of any given dealer regarding the state $\omega$. However, for each realization of $\omega$, we can compute a dealer’s reward resulting from any given dealer’s action profile $a \in A$. Let $u_i(\omega, a)$ denote dealer $i$’s expected profit or reward in a single period $t$ given the state $\omega$ and dealer’s action profile $a$. Here, expectations are taken with respect to the possible trader’s reactions $s$. Namely,
\[ u_i(\omega, a) = \sum_{s \in S} F(\omega, a, s) (W(\omega)Q_i(a, s) + P_i(a, s)) \]  \hspace{1cm} (2)

We denote with \( \tilde{a} \in \Delta A \) a distribution over dealers’ action profiles, with \( \tilde{a}(a) \) representing the probability that the action profile is \( a \in A \) when dealers play \( \tilde{a} \). Thus, for \( \tilde{a} \in \Delta A \), dealer \( i \)'s reward in state \( \omega \) is

\[ u_i(\omega, \tilde{a}) := \sum_{a \in A} \tilde{a}(a) u_i(\omega, a), \]

We conclude this section with an example that illustrates how our general model can be specified to obtain a Glosten and Milgrom type model.

**Example 1** Consider the Glosten and Milgrom trading mechanism described above, i.e., \( A = G \times G \) and \( S = \{sell, no\text{-}trade, buy\} \). Let \( \alpha_i, \beta_i \) be the bid and ask quotes set by dealer \( i \), and let \( \alpha \) and \( \beta \) be the best ask and best bid across dealers, respectively. Let \( v(\omega) \in \{v_1, v_2\} \) and \( e(\omega) \in \{-e, e\} \), with \( v_1 < v_2 \), \( e > 0 \) and \( \Pr[e(\omega) = e] = 1/2 \).

**Traders** Traders are mean-variance investors with utility \( E[\tilde{x}] - \frac{\gamma}{2} \text{Var}[\tilde{x}] \), where \( \tilde{x} \) is a trader’s post-trade wealth. Traders only differ in their initial inventory \( y \). Let the distribution of traders’ inventory be state-independent and uniform on the finite interval \( \Theta = [-\overline{y}, \overline{y}] \) with \( \overline{y} > 1/2 \). Then, it results

\[ F(\omega, a, sell) = \max \left\{ 0, \min \left\{ 1, 1 - \left( 1 - \frac{1}{2} + \frac{v(\omega) - \beta}{\gamma \text{Var}[\tilde{e}] + \overline{y}} \right) \right\} \right\}, \]  \hspace{1cm} (3)

\[ F(\omega, a, buy) = \max \left\{ 0, \min \left\{ 1, \left( -\frac{1}{2} + \frac{v(\omega) - \alpha}{\gamma \text{Var}[\tilde{e}] + \overline{y}} \right) \right\} \right\}, \]  \hspace{1cm} (4)

\[ F(\omega, a, no\text{-}trade) = \max \{0, 1 - F(\omega, \beta, sell) - F(\omega, \alpha, buy)\}. \]  \hspace{1cm} (5)

It is straightforward to verify that \( F \) satisfy Assumption RFT for \( \overline{\rho} = \rho := (\overline{y} - 1/2) \gamma \text{Var}[\tilde{e}]. \)
Dealers Given a dealer’s quote profile $a$ and a state $\omega$, dealer $i$’s reward is

$$ u_i(\omega, a) = (W(\omega) - \beta_i)F(\omega, a, \text{sell})1_{\{\beta_i = \beta\}}\eta_\beta(a) + (\alpha_i - W(\omega))F(\omega, \text{buy})1_{\{\alpha_i = \alpha\}}\eta_\alpha(a) $$

where $\eta_\beta > 0$ and $\eta_\alpha > 0$ are tie-breaking rules applied in case more than one dealer sets the best bid or ask, respectively. The symbol $1_{\{\}}$ denotes the indicator function.

2.3 The repeated game

We now move to the repeated game. The rewards of the dealers are discounted at the common factor $\delta < 1$, and the overall payoff is the average discounted sum of rewards. The discount factor $\delta$ accounts for both the dealer’s impatience and the possibility that public information is released in the current period. In each period, dealers’ actions and traders’ reactions are observed by all dealers. Let $H^t$ denote the set of public histories $h^t = \{a^\tau, s^\tau\}_{\tau=0}^{t-1}$. Given some sequence of action profiles $\{a^t\}_{t=1}^\infty$ by the dealers, dealer $i$’s payoff in state $\omega$ is

$$ \sum_{t=1}^\infty (1 - \delta)\delta^{t-1}u_i(\omega, a^t). \tag{6} $$

A public strategy profile (strategy, henceforth) is a mapping $\sigma : \cup_t H^t \to \times_i \Delta A_i$ that maps each public history $h^t$ in the action profile that dealers play at time $t$. For any given state $\omega$ and any history $h^t$, a strategy $\sigma$ induces a probability distribution over future histories in the standard fashion and hence an occupation measure over action profiles that we denote $\tilde{a}_{(\omega, \sigma, h^t)} \in \Delta A$.\footnote{Allowing for a stochastic discount factor complicates exposition but does not affect results as long as the discount factor remains close enough to one.} \footnote{Here, expectation is taken with respect to the possible realization of traders’ orders $\{s^\tau\}_{t=1}^\infty$, taking the state $\omega$ as given.} \footnote{Formally, an occupation measure of $a \in A$ is the discounted expected frequency with which action profile $a$ will be played:}

$$ \tilde{a}_{(\omega, \sigma, h^t)}(a) := E_{\sigma}\left[ \sum_{\tau \geq t} (1 - \delta)\delta^{\tau-t}1_{\{a^\tau = a\}} \mid \omega, h^t \right]. $$
Let $V_i(\omega, \sigma| h^t)$ denote dealer $i$’s expected continuation payoff after history $h^t$ given $\omega$ and $\sigma$. We have

$$V_i(\omega, \sigma| h^t) = \sum_{a \in A} \tilde{a}_{(\omega, \sigma, h^t)}(a) u_i(\omega, a) = W(\omega) Q_i(\omega, \sigma| h^t) + P_i(\omega, \sigma| h^t),$$

where

$$Q_i(\omega, \sigma| h^t) := \sum_{a \in A} \tilde{a}_{(\omega, \sigma, h^t)}(a) \sum_{s \in S} Q_i(a, s) F(\omega, a, s)$$

is the expected change in dealer $i$’s asset inventory, and

$$P_i(\omega, \sigma| h^t) := \sum_{a \in A} \tilde{a}_{(\omega, \sigma, h^t)}(a) \sum_{s \in S} P_i(a, s) F(\omega, a, s)$$

is the expected change in dealer $i$’s cash holding.

We are interested in the belief-free equilibria of the repeated game (hereafter, BFE). These are sub-game perfect equilibria for any possible underlying state and robust to the specifications of each dealer’s information about the true $\omega$. In other words, we look for dealers’ strategy profiles $\sigma$ such that at any time $t$ after any history $h^t$ for any dealer $i$, it is optimal to choose an action according to $\sigma_i(h^t)$ no matter what the dealer believes about the true state of Nature $\omega$. Formally,

**Definition 1** A belief-free equilibrium is a strategy profile $\sigma^*$ such that for every state $\omega$, $\sigma^*$ is a sub-game-perfect Nash equilibrium of the repeated game with rewards $u(\omega, \cdot)$, that is, of the repeated game with complete information in which the state $\omega$ is common knowledge among dealers:

$$\sigma^*_i \in \arg\max_{\sigma_i} V_i(\omega, \sigma_i, \sigma^*_{-i}| h^t),$$

for all players $i$, all $\omega \in \Omega$, all $t$ and all $h^t \in H^t$.

Some remarks are in order. First, a BFE is a perfect Bayesian equilibrium given any initial prior distribution of dealers’ belief about $\omega$ and any additional private information a dealer
might possess. Thus, a BFE is a sub-game-perfect equilibrium no matter the specific dealers’ information structure. Second, a BFE is an equilibrium no matter whether dealers are Bayesian. In particular, a BFE remains an equilibrium even if dealers are ambiguity-averse as long as ambiguity pertains to the distribution of the possible states of nature $\omega \in \Omega$.

3 Learnable and non-learnable states and achievable payoffs

To be able to construct a belief-free equilibrium of the repeated game, it is important to clarify what rewards dealers can achieve in a single trading round. Of course, that depends on what they know about the state. Note, however, that a BFE must remain an equilibrium even under the canonical assumption that dealers have no private information about $\omega$, a situation we cannot rule out from our analysis. In this case, the only information about $\omega$ that can emerge and that a dealer can exploit is that coming from traders’ behavior. In this section, we first formalize what dealers can learn from traders about the state of Nature. Second, we describe some rewards dealers can achieve once they have learned as much as possible from traders.

The way the function $F$ is affected by the state determines the information about the true $\omega$ that dealers can gather from traders’ behavior. For example, if traders’ behavior is identical in state $\omega$ and in state $\omega'$, then it is impossible to tell these two states apart even after observing an infinite history of how they react to dealers’ actions. Namely, $F$ defines a partition that we assume to be finite, $\hat{\Omega} := \{\hat{\omega}_1, \hat{\omega}_2, \ldots, \hat{\omega}_k\}$, of the space $\Omega$. Formally, the partition $\hat{\Omega}$ satisfies the following two conditions:

15 To see this, note that in a perfect Bayesian equilibrium, dealers’ strategies satisfy

$$\sigma^*_i \in \arg\max_{\sigma_i} E \left[ V_i(\omega, \sigma_i, \sigma^*_j|h^t)|I_i \right],$$

where expectations are taken with respect to both the possible states $\omega$ and the possible realizations of traders’ orders $\{s^t\}_{t=1}^\infty$, and $I_i$ is dealer $i$’s private information. Hence, a BFE is also a perfect Bayesian equilibrium, but a perfect Bayesian equilibrium need not be belief-free.

16 Unlike in Easley and O’Hara (2010), where some of the traders are ambiguity-averse, here, ambiguity aversion applies to dealers.
1. Two states $\omega, \omega' \in \Omega$ belong to the same element $\hat{\omega} \in \hat{\Omega}$ if and only if, for any given $(a, s) \in A \times S$, it holds that

$$F(\omega, a, s) = F(\omega', a, s).$$

In words, traders’ reactions to dealers’ actions are identical in states $\omega$ and $\omega'$. In this case, we say that $\omega$ and $\omega'$ are not statistically distinguishable as a result of traders’ behavior.

2. Conversely, suppose that two states $\omega, \omega' \in \Omega$ satisfy $\omega \in \hat{\omega}, \omega' \in \hat{\omega}'$, where $\hat{\omega}, \hat{\omega}' \in \hat{\Omega}$ and $\hat{\omega} \neq \hat{\omega}'$. Then, there exists a non-empty set $A(\hat{\omega}, \hat{\omega}') \subseteq A$ such that if $a \in A(\hat{\omega}, \hat{\omega}')$, then

$$F(\omega, a, s) \neq F(\omega', a, s)$$

for some $s \in S$.

In words, there are suitable choices for dealers’ actions (i.e., for $a \in A(\hat{\omega}, \hat{\omega}')$), for which the distribution of traders’ reactions $s$ differs for $\omega \in \hat{\omega}$ and $\omega \in \hat{\omega}'$. In this case, we say that $\omega$ and $\omega'$ are statistically distinguishable as a result of traders’ behavior.

In particular, because traders have no information about the $\tilde{e}$ component of the asset value, and the distribution of their types does not depend on $\tilde{e}$, their behavior does not depend on the actual value of $\tilde{e}$. In contrast, because they observe $\tilde{v}$, their behavior will be statistically measurable with respect to $\tilde{v}$. In summary, $F$ satisfies the following property:

**Learnable States (LS):** The set $\Theta$ and the probability distribution $Z$ are such that: the partition $\hat{\Omega}$ of learnable states is finite; if $v(\omega) \neq v(\omega')$, then $\omega$ and $\omega'$ are statistically distinguishable, but traders’ behavior does not allow dealers to learn the $e(\omega)$ component of $W(\omega)$.

It is important to clarify what rewards dealers can achieve once they have learned the element of $\hat{\Omega}$ containing the true state $\omega$. The next proposition shows that in a single trading round, dealers can achieve different reward profiles, namely, profiles such that all dealers make positive
profits, profiles such that all dealers make negative profits, profiles such that a dealer cannot make a positive profit once the other dealers know his beliefs on $\omega$, and profiles such that dealers’ profits differ.

**Proposition 1** If the set $\Theta$ and the probability distribution $Z$ are such that properties $\textbf{RFT}$ and $\textbf{LS}$ are satisfied, then for any given $\hat{\omega} \in \hat{\Omega}$,

1. **(Positive maximum payoffs)** There exists a non-empty set $A^*(\hat{\omega}) \subseteq \Delta A$ such that $u_i(\omega, a) > 0$ for all $\omega \in \hat{\omega}$ and dealer $i$ if and only if $a \in A^*(\hat{\omega})$.

2. **(Negative minimum payoffs)** There exists an action profile $a(\hat{\omega}) \in \Delta A$ such that $u_i(\omega, a(\hat{\omega})) < 0$ for all $\omega \in \hat{\omega}$ and dealer $i$.

3. **(Non-positive expected payoffs)** For any given dealer $i$ and any distribution $\mu_{\hat{\omega}} \in \Delta \hat{\omega}$, there exists $a^i_{-i}(\mu_{\hat{\omega}}) \in \times_{j \neq i} \Delta A_j$ such that

$$\max_{a_i} \sum_{\omega \in \hat{\omega}} \mu_{\hat{\omega}}(\omega) u_i(\omega, a_i, a^i_{-i}(\mu_{\hat{\omega}})) \leq 0.$$ 

4. **(Non-equivalent payoffs)** There exist $n$ action profiles $\{a^1(\hat{\omega}), \ldots, a^n(\hat{\omega})\} \in [\Delta A]^n$ such that $u_i(\omega, a^i(\hat{\omega})) < u_i(\omega, a^j(\hat{\omega}))$ for all $i \neq j$ and $\omega \in \hat{\omega}$.

In words, the set $A^*(\hat{\omega})$ is the set of dealers’ action profiles guaranteeing that each dealer $i$ makes strictly positive profits in all states $\omega \in \hat{\omega}$. Let $u^* > 0$ denote a lower bound on payoffs from actions in $A^*(\hat{\omega})$. Roughly speaking, **Positive maximum payoffs** and **Negative minimum payoffs** guarantee that for each statistically distinguishable state $\hat{\omega}$, there are action profiles providing each dealer with at least $u^* > 0$ and action profiles leading to strictly negative payoffs, respectively. **Non-positive expected payoffs** guarantee that when fixing a dealer $i$ and his beliefs about the true $\omega$, the other dealers can make sure that this dealer’s expected payoff is non-positive. For example, if the other dealers provide the maximum liquidity at a price equal to
the expected value of the asset for dealer $i$. Non-equivalent payoffs indicate that for each $\hat{\omega}$, one can find as many action profiles as there are dealers such that dealer $i$ prefers all the other $n - 1$ action profiles to the $i$-th action profile.

4 Necessary conditions for robust price formation

We now identify the features of dealers’ strategies that are necessary for these strategies to form a BFE. This is useful to distinguish the predictions of market-microstructure theories that are robust from those that rely on specific assumptions about dealers’ information structure. Put differently, a strategy that forms a perfect Bayesian equilibrium for a specification of a dealer’s beliefs but that does not satisfy at least one of the properties described below is not robust to changes in dealers’ information structure.

First, as we mentioned above, a BFE must remain a sub-game perfect equilibrium even when dealers have no private information about $\omega$. In this case, there is no market participant who can tell apart two states $\omega$ and $\omega'$ belonging to the same element $\hat{\omega}$ of $\hat{\Omega}$. Hence, in a BFE, the play must be the same for such two states. Formally,

**Lemma 1 (Measurability with respect to traders’ behavior)** Let $\sigma^*$ be a BFE, $\hat{\omega}$ an element of $\hat{\Omega}$ and $h^t$ a finite history, then

$$\tilde{a}(\omega, \sigma^*, h^t) = \tilde{a}(\hat{\omega}, \sigma^*, h^t),$$

for all states $\omega \in \hat{\omega}$.

Second, note that each dealer can guarantee zero profit by abstaining from trading. This implies that in a BFE, for each state $\omega \in \Omega$, each dealer’s payoff cannot be strictly negative. Otherwise, for some beliefs, a dealer would prefer to deviate to the no-trade action.
Lemma 2 (Strictly positive dealer payoffs) Let $\sigma^*$ be a BFE, $\hat{\omega}$ an element of $\hat{\Omega}$ and $h^t$ a finite history, then for all states $\omega \in \hat{\omega}$,

$$V_i(\omega, \sigma^*|h^t) \geq 0.$$ 

Moreover, if $Q_i(\hat{\omega}, \sigma^*|h^t) \neq 0$, then the weak inequality holds for at most one $\omega \in \hat{\omega}$.

In other words, in a BFE, dealers make non-negative profits state-by-state. In addition, when the equilibrium leads to changes in a dealer’s inventory, he makes strictly positive profits in most states.

Third, Lemma 1 and 2 imply that there is a bound on the size of the inventories that dealers accumulate in a BFE. Namely, the same $\sigma^*$ strategy can lead to positive payoffs in different states only if the equilibrium trading volume is relatively balanced and maintains dealers’ inventories bounded. In particular, the larger the residual uncertainty resulting from the $\tilde{e}$ component and the smaller the trading volume between traders and dealers, the tighter the bound imposed by robustness to the dealers’ average inventory will be.

The intuition is simple. Dealers’ payoffs can be decomposed into two components. The first component is what they gain from intermediating traders’ demand and supply without taking a net position in the asset. Because of RFT 1.- 2., on average, for each $\hat{\omega}$, dealers can sell the shares for more than what they pay. The larger the trading volume generated in this way, the larger dealers’ profits no matter the true value $W(\omega)$. The second component is what dealers gain or lose from loading on net positions in the asset. This is proportional to the size of their net position and to the difference between the average transaction price for each share and the asset value $W(\omega)$. Because of RFT 3.- 4. for some state $\omega$, this can result in a loss as large as $\tau$ times the inventory size. Thus, for the overall payoff to be positive in all states, the dealers’ inventory cannot explode.

Formally, let $Q(\omega, \sigma^*|h^t) := \sum_i Q_i(\omega, \sigma^*|h^t)$ be the expected change in dealers’ aggregate
inventory in state \( \omega \) after history \( h^t \). This is the sum of two components: a negative component denoted \( Q^-(\omega, \sigma^*|h^t) \leq 0 \), which are transactions in which traders buy from dealers, and a positive component denoted \( Q^+(\omega, \sigma^*|h^t) \geq 0 \), which are transactions in which traders sell to dealers. The volume of trade between traders and dealers is \( Q^+(\omega, \sigma^*|h^t) - Q^-(\omega, \sigma^*|h^t) \geq 0 \). Then, we have

**Lemma 3 (Bounded dealers’ inventories)** Let \( \sigma^* \) be a BFE and \( h^t \) an arbitrary history. Then, for all states \( \omega \in \Omega \),

\[
|Q(\omega, \sigma^*|h^t)| \leq \frac{\rho (Q^+(\omega, \sigma^*|h^t) - Q^-(\omega, \sigma^*|h^t))}{\varepsilon}. \tag{8}
\]

Fourth, the fact that dealers’ payoffs must be positive (Lemma 2) implies that if \( \omega \in \hat{\omega} \), then in equilibrium, dealers actions are in \( A^*(\hat{\omega}) \) sufficiently often. When \( a^t \in A^*(\hat{\omega}) \), we say that the dealers are in a \( \hat{\omega} \)-exploiting period. If for two elements \( \hat{\omega}, \hat{\omega}' \in \hat{\Omega} \) one had \( A^*(\hat{\omega}) \cap A^*(\hat{\omega}') = \emptyset \), then the action profile leading all dealers to make positive payoffs when \( \omega \in \hat{\omega} \) would lead at least one dealer to make negative payoffs if \( \omega \in \hat{\omega}' \). Because of Lemma 1, traders’ behavior must allow the play to distinguish the two states and hence allow the use of the appropriate exploiting actions so that the play can satisfy Lemma 2. This is possible only if dealers’ actions belong to \( A(\hat{\omega}, \hat{\omega}') \) sufficiently often. When \( a^t \in A(\hat{\omega}, \hat{\omega}') \), we say that dealers are in an exploring period. Thus, both exploiting and exploring periods are necessary ingredients of a BFE and need to be played sufficiently often.

**Lemma 4 (Frequent exploring)** Suppose that \( A^*(\hat{\omega}) \cap A^*(\hat{\omega}') = \emptyset \), and let \( \sigma^* \) be a BFE and \( h^t \) a finite history. Then, for any \( \omega \in \Omega \) action profiles in \( A(\hat{\omega}, \hat{\omega}') \) are played with strictly positive probability:

\[
\tilde{a}_{(\omega, \sigma^*, h^t)}(A(\hat{\omega}, \hat{\omega}')) > 0.
\]

This lemma states that exploring periods must be relatively frequent, no matter the history.
As we illustrate below, because exploring periods are associated with higher sensitivity of prices to trading volume, their recurrence makes prices more volatile than Bayesian beliefs on $\hat{\Omega}$.

5 Sufficient Conditions for BFE Pricing

In this section, we establish the existence of a BFE by constructing one. We first introduce the building blocks of our candidate strategy. We then show how to combine these ingredients to obtain a BFE.

**Ingredients:** We start by defining a *market measure* $\pi$. Let $\Pi \subseteq \Delta\hat{\Omega}$ be a closed set of probability distributions over $\hat{\Omega}$ and $\pi$ an element in $\Pi$. Let $\pi(\hat{\omega})$ denote the probability that $\pi$ assigns to $\hat{\omega}$. Fix $\varepsilon > 0$. We say that the market measure $\pi$ *points to* a state $\hat{\omega}$ if it assigns a probability $\pi(\hat{\omega}) \geq 1 - \varepsilon$ to state $\hat{\omega}$.

Let $\phi : \Pi \times A \times S \to \Pi$ be a probability-updating rule. That is, time $t + 1$ market measure $\pi^{t+1} = \phi(\pi^t, a^t, s^t)$ is a function $\phi$ of time $t$ market measure, dealers’ actions and trader reaction. Thus, $\pi^t$ can be recursively computed from the map $\phi$, given the sequence $\{a^\tau, s^\tau\}_{\tau=0}^{t-1}$ of actions and signals, and the initial value $\pi^0$. We will specify a particular $\phi$ below. We are interested in strategies such that on the equilibrium path and in each period $t$, dealers’ actions depend on $\pi^t$ only. That is, given $\phi$, we define a *partial strategy* to be a map $\sigma : \Pi \to \Delta A$. A public strategy profile (strategy henceforth) is a mapping $\hat{\sigma} : \bigcup_t H^t \to \times_i \Delta A_i$.

Fix an arbitrary starting market measure $\pi^0 \in \Pi$, an updating rule $\phi$ and a partial strategy $\sigma$, and consider a situation in which dealers use the partial strategy $\sigma$ and the market measure evolves according to $\phi$. We say that the pair $(\phi, \sigma)$ is $\varepsilon$-*learning* if, over many periods, the market measure points at the $\hat{\omega}$ that contains the true state $\omega$ with a frequency that is at least $1 - \varepsilon$. In other words, the market measure is rarely far away from the truth in terms of long-run frequency. Formally:
Definition 2 The pair $(\phi, \sigma)$ is $\varepsilon$-learning, for $\varepsilon > 0$, if for any $\hat{\omega} \in \hat{\Omega}$, any $\omega \in \hat{\omega}$ and any $\pi^0 \in \Pi$,\[\Pr_{\omega, \sigma} \left[ \liminf_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} 1_{\{\pi^t(\hat{\omega}) > 1 - \varepsilon\}} < 1 - \varepsilon \right] < \varepsilon, \quad (9)\]

We say that a pair $(\phi, \sigma)$ is $\varepsilon$-exploiting if whenever the market measure points at some $\hat{\omega}$, play is such that each dealer’s payoff is strictly positive in all states $\omega$ included in $\hat{\omega}$. Formally:

Definition 3 The pair $(\phi, \sigma)$ is $\varepsilon$-exploiting, for $\varepsilon > 0$, if for all $\hat{\omega} \in \hat{\Omega}$ and all $h^t$ such that $\pi^t(\hat{\omega}) \geq 1 - \varepsilon$, we have $\Pr_{\sigma} [a^t \in A^*(\hat{\omega}) | h^t] > 1 - \varepsilon$.

The following theorem shows that a pair $(\phi, \sigma)$ that is both $\varepsilon$-exploring and $\varepsilon$-exploiting forms a BFE if dealers are patient enough.

Theorem 1 There exists $\bar{\varepsilon} > 0$ such that for any $\varepsilon < \bar{\varepsilon}$, if $(\phi, \sigma)$ is $\varepsilon$-learning and $\varepsilon$-exploiting, then there exists $\delta < 1$ such that the outcome induced by $\sigma$ is a belief-free equilibrium outcome for all $\delta \in (\delta, 1)$.

That is, there exists a belief-free equilibrium $\sigma^*$ that specifies the same action profile as the partial strategy $\sigma$, after any history after which no player has deviated. Observe that a pair $(\phi, \sigma)$ that is both $\varepsilon$-exploring and $\varepsilon$-exploiting forms a strategy profile satisfying the necessary conditions for a BFE described in Section 4. Namely, first, the way dealers set their actions is clearly measurable with respect to traders’ behavior (Lemma 1) because dealers’ actions at $t$ only depend on $\pi^t$, which is itself a function of the public history. Second, this strategy leads to positive payoffs (Lemma 2) because the market measure frequently points to the right $\hat{\omega}$ ($\varepsilon$-exploring), and when this happens, the dealers’ payoff is positive ($\varepsilon$-exploiting). Third, the fact that dealers’ inventories are bounded (Lemma 3) is a consequence of the fact that dealers’ payoffs remain positive for all values of $W(\omega)$. Fourth, the strategy generates exploring (Lemma
4). In fact, for \((\phi, \sigma)\) to be \(\varepsilon\)-exploring, it is necessary that no matter the past history, the actions that allow the distinction of the true \(\hat{\omega}\) from the others \(\hat{\omega} \in \hat{\Omega}\) are played with strictly positive frequency. Thus, when dealers behave according to a pair \((\phi, \sigma)\) that is both \(\varepsilon\)-exploring and \(\varepsilon\)-exploiting, they all achieve long-term positive profits independently of the state \(\omega\). In the proof of Theorem 1, we show that dealers have no incentive to deviate. In fact, because of the properties detailed in Proposition 1, there are strategies that will be played after a deviation and that punish the deviating dealer while rewarding the other dealers. For this threat of punishment to be an effective deterrent, dealers should care enough about their future payoffs, \(i.e.,\) be patient enough.

6 BFE vs. canonical equilibrium: An example and some empirical implications

In this section, we analyze a particular class of BFE for the specific quote-driven market of Example 1. We compare this equilibrium with the canonical zero-profit equilibrium that can be obtained if one makes the additional assumption that all dealers are equally uninformed.

Let us first consider the canonical equilibrium. This equilibrium relies on the assumption that all dealers start with a common prior belief that \(\Pr[v(\omega) = v_2] = p^0\). Then, there is a perfect Bayesian equilibrium in which in any period \(t\), each dealer’s expected profit is nil. Furthermore, recalling that \(\rho := (\pi - 1/2)\gamma \text{Var}[\varepsilon]\) and focusing on the case \(\rho > (v_2 - v_1)\), one can write period \(t\), best bid and ask quotes as follows

\[
\alpha^t = \alpha(p^t) := E[\tilde{v} | h^t, s^t = \text{buy}] = E[\tilde{v} | h^t] + \frac{\rho}{2} - \sqrt{\frac{\rho^2}{4} - \text{Var}[\tilde{v} | h^t]},
\]

(10)

\[
\beta^t = \beta(p^t) := E[\tilde{v} | h^t, s^t = \text{sell}] = E[\tilde{v} | h^t] - \frac{\rho}{2} + \sqrt{\frac{\rho^2}{4} - \text{Var}[\tilde{v} | h^t]},
\]

(11)

\[
p^{t+1} = \phi_B(p^t, a^t, s^t),
\]

(12)
where $\phi_B(p^t, (\alpha^t, \beta^t), s^t)$ denotes the Bayesian posterior probability that $v(\omega) = v_2$ resulting from the prior probability $p^t$ and from the trader’s reaction $s^t$ to dealers’ quotes $a^t$.\footnote{In order to simplify the exposition and notation, we neglect the rounding required from the fact that quotes belong to a grid.} The r.h.s. of equations (10) and (11) are obtained by considering that the probability of a trader buying, selling or not trading are given by equations (4), (3), and (5), respectively. Expressions $E[v|h^t]$ and $\text{Var}[v|h^t]$ are the expectation and the variance of $v(\omega)$, respectively, computed using the belief $p^t$ that evolves according to (12). This equilibrium has the advantage of being “Markov”: first, in every period $t$, best bid and ask quotes only depend on the dealers’ common belief $p^t$; second, the dealers’ common posterior belief $p^{t+1}$ only depends on the common time-$t$ prior $p^t$ and on $(a^t, s^t)$, dealers’ and trader’s actions at time $t$. Thus, this quoting strategy satisfies the measurability condition of Lemma 1. However, because at any $t$, a dealer’s expected profit is nil for the belief $p^t$, there is at least one value of $W(\omega)$ for which dealers’ profits are strictly negative. Thus, this strategy does not satisfy Lemma 2. Hence, the canonical equilibrium strategy cannot form a BFE.

We now illustrate how a BFE can be achieved with a modification of the canonical equilibrium strategy. For this, it is necessary and sufficient that: first, each dealer makes strictly positive profits no matter the value of $W(\omega)$ (Lemma 2). This is achieved with a bid-ask spread that is larger than that in canonical equilibrium. Second, dealers’ inventory remains bounded (Lemma 3). This is achieved by having quotes that are decreasing in dealers’ aggregate inventory. Third, the strategy is $\varepsilon$-learning and $\varepsilon$-exploiting (Lemma 4 and Theorem 1). This is achieved by having appropriate market measures on the partition $\hat{\Omega}$, probability-updating rule $\phi$ and strategy $\sigma$ mapping the market measure and dealers’ aggregate inventory into dealers’ quotes.

Recall that for Example 1, we have $\hat{\Omega} = \{\hat{\omega}_1, \hat{\omega}_2\}$, where $\hat{\omega}_k$ is the set of states $\omega \in \Omega$ such that $v(\omega) = v_k$, $k = 1, 2$. Let us first define a market measure $\pi$ on $\hat{\Omega}$ and a probability-updating rule $\phi$. Fix some small $\varepsilon > 0$ and let $\Pi := [\varepsilon/4, 1 - \varepsilon/4]$. Let $\pi^t \in \Pi$ denote the probability
that the market measure assigns to $\hat{\omega}_2$ at time $t$. Fix an arbitrary $\pi^0 \in [\varepsilon, 1 - \varepsilon]$ as the initial market measure. Afterwards, the market measure evolves according to the following updating rule $\phi : \Pi \times A \times S \rightarrow \Pi$:

$$\pi^{t+1} = \phi(\pi^t, a^t, s^t) := \arg \min_{\pi \in \Pi} \| \pi - \phi_B(\pi^t, a^t, s^t) \|,$$

(13)

where $\phi_B(\pi^t, a^t, s^t)$ is the Bayesian posterior computed from prior $\pi^t$. In words, $\phi(\pi^t, a^t, s^t)$ maps a probability $\pi^t \in \Pi$ and an action-reaction profile $(a^t, s^t)$ into the probability $\pi^{t+1}$ that is the element in $\Pi$ that is closest to the Bayesian posterior computed using a prior equal to $\pi^t$ and the information provided by a trader’s reaction $s^t$ to dealers’ action $a^t$. Note that from the definition of a BFE, the market measure need not reflect any of the dealer’s beliefs. In particular, $\pi^0$ can be chosen arbitrarily in the interval $[\varepsilon, 1 - \varepsilon]$, and afterwards, $\pi^t$ does not actually follow the Bayesian rule. Nevertheless, the level of the market measure affects the equilibrium quotes set by rational Bayesian dealers.

Namely, we say that for $\pi^t > 1 - \varepsilon$ (resp. for $\pi^t < \varepsilon$), the game is in a $\hat{\omega}_1$-exploiting phase (resp. $\hat{\omega}_2$-exploiting phase). For $\pi^t \in [\varepsilon, 1 - \varepsilon]$, the game is in the exploring phase. Let $Y^t$ denote the dealers’ aggregate inventory at time $t$. We can now describe the mapping $\sigma$ that maps $\pi^t$ and $Y^t$ into the dealers’ quotes. Fix $b$ and $d$ such that $0 < b, d < \rho$. During a $\hat{\omega}$-exploiting phase, the best ask and bid equilibrium quotes satisfy

$$\alpha^t = v(\hat{\omega}) + d - bY^t,$$

(14)

$$\beta^t = v(\hat{\omega}) - d - bY^t,$$

(15)
respectively. During an exploring phase, the best ask and bid quotes satisfy

\[ \alpha^t = \alpha(\pi^t) + d - bY^t, \]

\[ \beta^t = \beta(\pi^t) - d - bY^t, \]

where \( \alpha(\cdot) \) and \( \beta(\cdot) \) are the functions defined in equations (10) and (11), respectively. Each dealer sets the best bid and ask quotes as a strictly positive fraction of the time according to an arbitrary sharing rule, leading each dealer to obtain a strictly positive fraction of dealers’ aggregate profit and inventory.

The (on-path) equilibrium play can then be seen as the alternation of two types of phases: exploring phases and exploiting phases. Whenever \( \pi^t \in [\varepsilon, 1 - \varepsilon] \), the game is in an exploring phase: dealers’ quotes induce an informative flow of trades. Thus, as time passes, the market measure attaches more and more weight to the true \( \hat{\omega} \). An exploiting phase starts as soon as the market measure attaches enough weight to a particular state. Namely, whenever \( \pi^t < \varepsilon \) (resp., \( \pi^t > 1 - \varepsilon \)), the game is in the \( \hat{\omega}_1 \)-exploiting phase (resp. \( \hat{\omega}_2 \)-exploiting phase). In this phase, \( a^t \in A^*(\hat{\omega}_1) \) (resp. \( a^t \in A^*(\hat{\omega}_2) \)). This guarantees that dealers gain the spread while keeping their inventory small in absolute value.

In response to the order flow during an exploiting phase, however, play can revert to the exploring phase, and so on. The reason why an exploiting phase cannot last forever is that a dealer whose beliefs differ from the market measure must be given incentives to play along and wait for play to shift towards the exploiting phase corresponding to the asset value that he might believe in. At the same time, no matter the current level of the market measure and a dealer’s belief about \( \omega \), the dealer must expect that the play will shift toward the correct exploiting phase within a bounded length of time. Otherwise, even a patient dealer would prefer to deviate and generate extra profits in the current trading round (even if held to zero profits afterwards), rather than making losses during the long transition period required for the market measure to adjust to
what he thinks the right exploiting phase is. This is possible only if the market measure evolves in a way that satisfies two conditions. First, during an exploiting phase, the probability attached by the market measure to states that are unlikely in view of the flow of information provided by traders’ orders decreases over time. This is satisfied by our choice of $\phi$: in the $\hat{\omega}_k$-exploiting phase quotes straddle $v(\hat{\omega}_k)$ and induce a balanced flow of trade only if the true state $\omega \in \hat{\omega}_k$. Thus, an unbalanced order flow changes the market measure and eventually leads to a new exploring phase. Second, during an exploiting phase, the market measure cannot be persistent but instead must be sensitive to the new public information provided by traders’ orders. Bayesian updating, for instance, does not satisfy these two properties: while it allows traders to pin down the true $v(\hat{\omega})$ almost surely eventually, it is too persistent for our purpose: once a prior probability is sufficiently concentrated on a state, it takes arbitrarily long for the Bayesian posterior to budge.

To have a sensitive market measure no matter the history, we specified that $\pi^t$ remained in the interval $\Pi = [\varepsilon/4, 1 - \varepsilon/4]$: $\pi^t$ evolves as a Bayesian belief as long as the posterior remains in the interior of $\Pi$. Because the resulting market measure is never too concentrated on a state, the time it requires to point to the correct $\hat{\omega}$ is bounded above, no matter the history.

Note that if dealers follow these strategies, at any time $t$, dealers expect their future profits to be strictly positive independently of the true $\omega$ and independently of their past history. Thus, even if a dealer’s belief differs from the market measure, he will not deviate because (using standard repeated-game logic), the result of “Non-positive expected payoffs” from Proposition 1 guarantees that other dealers can ensure that the deviating dealer makes zero profits for a long enough but finite length of time. Thus, a deviation cannot be profitable if the discount rate is sufficiently low.
6.1 Simulation and explanatory power of BFE and the canonical equilibrium

To illustrate some of the salient differences between BFE and the canonical equilibrium (hereafter, CE), we simulate price behavior resulting from these two equilibria.\footnote{For this simulation, the parameters of the fundamentals are: $v_1 = $15, $v_2 = $18, $\tau = 8$, $Var[\tilde{e}] = 4$, $\gamma = 1$, $\overline{\gamma} = 13/6$ (thus, $\rho = 15$), whereas the BFE parameters are $d = $0.05, $b = 0.01$, $\varepsilon = 0.03$, $\rho^0 = \pi^0 = 0.5$. The figure reports the time series of 3000 trades.}

**Excess price volatility:** In a BFE, quotes are intrinsically more volatile than they are in CE. This is because in CE, dealers’ quotes reflect the common Bayesian belief, which eventually attaches probability arbitrarily close to 1 to the true value of $v$. This leads to a vanishing volatility and bid-ask spread, with quotes that remain arbitrarily close to $v$. This cannot happen for the BFE market measure, which is never too concentrated on a given state and hence remains unstable. Thus, independently of the previous history of trade and of dealers’ actual beliefs about $v(\omega)$, the market measure and quotes remain sensitive to the trading volume. This is illustrated in Figure 1, which reports a simulation of the two equilibria for $v(\omega) = v_1$. The sequence of traders is the same for the two equilibria. The right panel of Figure 1 reports the evolution of quotes in BFE. The left panel reports the evolution quotes in CE.

**Volatility clustering:** The recurrence of exploring and exploiting phases gives rise to price volatility clusters. In exploring phases, dealers attract informative unbalanced traders’ order flow, whereas in exploiting phases, dealers make profits from the intermediation of relatively balanced order flow. In exploring phases, quotes react more sharply to the trading volume. Thus, quotes’ volatility is higher in exploring than in exploiting phases. The volatility-clustering effect is apparent in the right panel of Figure 1. The alternation of these phases endogenously generates price volatility regime shifts, a pattern that has been extensively documented in the asset-pricing literature. Price sensitivity to the order flow in exploiting and exploring phases is illustrated in Figure 2, which shows how the market measure reacts to the trading volume in an exploiting...
Figure 1: Evolution of bid (red dots) and ask (blue dots) quotes in the canonical equilibrium (left panel) and in BFE (right panel). Ask quotes and bid quotes are in blue, and magenta, respectively.
Figure 2: Market measure and dealers’ inventory in an exploiting phase (left panel) and an exploring phase (right panel). Interestingly, volatility regime shifts are anticipated by precise patterns in the order flow and evolution of dealers’ inventory. A shift from low to high price volatility tends to be preceded by consistent traders’ order imbalance and significant changes in dealers’ inventory. The transition from high to low volatility phases follows the fading of traders’ order imbalance and a stabilization of dealers’ inventory. Note that in BFE, the volatility regime shift is completely endogenous and occurs in the absence of news. This is not the case in CE, which predicts that in the absence of news, price volatility is bound to fade.

Quote volatility vs. trading flow, bid-ask spread and profits: This ‘Markov” BFE has some interesting implications regarding the correlation between price volatility, liquidity (as
measured by bid-ask spread), dealers’ aggregate inventory and profits. In an exploring phase, orders are more informative; in comparison to exploiting phases, exploring phases are associated with larger bid-ask spread, price volatility and aggregate inventory and with lower profits. This is consistent with the empirical regularities observed by Comerton-Forde et al. (2010): liquidity is negatively correlated with dealers’ profits and inventories as well as with price volatility.

**News and volatility:** Whereas this benchmark model does not allow for exogenous shocks in information, it is straightforward to extend the model to allow for the exogenous arrival of public news about the fundamentals. Our BFE can accommodate this extension by having the market measure depend on all public information, i.e., public news as well as order flow. Unexpected news arriving when the market is an exploiting phase moves the market measure and may trigger an exploring phase. As a result, following news, price volatility increases. This may generate price overshooting and/or undershooting with respect to the level of quotes that will be reached once a new exploiting phase starts.

**Dealers’ profits:** One of the necessary conditions for an equilibrium to be belief-free is that dealers’ long-term profits are strictly positive state by state. In BFE, this is achieved by maintaining a spread that is larger than the one predicted in CE. Note, for example, that in BFE, the spread remains bounded away from 0 even when the market measure is relatively concentrated. As a result, while in CE, the average dealers’ aggregate per-period profit quickly converges to 0; in BFE, it is of the same magnitude as $d$ (see Figure 3). Note that a dealer’s *ex post* profit also depends on the value of $e(\omega) \in \{-e, e\}$. Figure 4 represents the *ex post* cumulative profit for $e(\omega) = -e$ and $e(\omega) = e$. In CE (left panel of Figure 4), the dealers’ cumulative profit remains negative for at least one realization of $e(\omega)$. In BFE, the dealers’ cumulative profit eventually becomes positive no matter the realized $e(\omega)$ (right panel of Figure 4).

**Dealers’ inventories:** A strictly positive profit in all states can only be achieved when aggregate inventory does not explode. For this reason, in BFE, dealers’ inventory must remain
Figure 3: Evolution of the average per-period profit taking $e(\omega) = 0$ in GM (red dashed line) and in BFE (blue solid line).

Figure 4: Cumulative dealers' profits for $e(\omega) = 0$ (dashed line), $e(\omega) = \underline{e}$ (blue line) and $e(\omega) = \overline{e}$ (red line).
bounded. This is not the case in CE. For example, in the simulation, \( v(\omega) = v_1 \). Hence, traders tend to sell more than buy the asset. The right panel of Figure 5 reports the evolution of aggregate inventory in CE (red line) and in BFE (blue line). In CE, dealers’ aggregate inventory tends to explode. To the contrary, in BFE, dealers’ aggregate inventory remains more balanced thanks to the bias in quotes \(-bY^t\). Equation (8) provides the upper and lower bound for the ratio between the average change in dealers’ inventory and the average trading volume.

The larger \( e \) and the smaller \( \rho \), the smaller this ratio is on average. The variable \( e \) can be interpreted as the asset-intrinsic uncertainty, i.e., the remaining uncertainty after incorporating traders’ information.\(^\text{19}\) The parameter \( \rho \) is a measure of traders’ willingness to trade for hedging rather than speculative reasons. In standard market microstructure parlance, \( \rho \) increases with liquidity traders’ activity. Thus, BFE predicts that dealers’ inventories are more balanced in the presence of intrinsic uncertainty and for companies that are seldom traded by institutional investors.

![Dealers' inventory: CE vs BFE](image)

**Figure 5:** Evolution of dealers’ aggregate inventory in CE and in BFE. Red dashed lines represent CE, and blue solid lines BFE.

\(^{19}\)Some possible proxies for the presence of intrinsic uncertainty are growth companies vs. utility companies, the youth of the firm, the firm’s sector, product market innovations, R&D investments, business sensitivity to exogenous risks, such as weather or other natural risks, and foreign country risk.
7 Extensions

Our environment is restrictive in several dimensions. In particular, dealers’ actions are observed by all other dealers. Furthermore, the state of the world that determines the fundamentals is fixed once and for all. Additionally, long-term market participants do not take advantage of their private information. Here, we sketch how the model can be extended and the analysis adapted to address such features.

A restriction of our model is that dealers’ actions are observable. This might not be realistic for some opaque markets, such as, for instance, when dealers’ quotes are anonymous. The imperfect monitoring of actions makes it more difficult to detect a dealer’s deviation from the mutually profitable strategy profile. This reduces the threat of punishment, but it does not eliminate the dealers’ ability to sustain a BFE. For this, it is sufficient that equilibrium strategies are built in a way that makes deviations detectable. For example, Christie and Schultz (1994) and Christie, Harris and Schultz (1994) document how Nasdaq dealers used to quote only even-eight quotes, thus guaranteeing a larger bid-ask spread and higher dealer profits. Deviations from such a scheme can easily be detected even when quotes are anonymous. More generally, the imperfect monitoring of players’ actions is not an issue for the existence of a BFE (see Fudenberg and Yamamoto (2011)). However, the imperfect monitoring of dealers’ actions might impose further restrictions on the type of equilibrium strategies that can be sustained in a BFE.

Allowing for fluctuations in the value of the asset raises no difficulty as long as these fluctuations take place at a slower rate than the learning process. That is, in the definition that \((\phi, \sigma)\) be \(\varepsilon\)-learning, we must now account for the fact that \(\hat{\omega}(\omega_t)\) depends on time \(t\). Hence, the learning requirement is considerably stronger. We must think that learning the fundamental value occurs on the same time scale as the fluctuations of the value itself; perhaps learning occurs within a day of trading, an interval of time over which the fluctuations in the fundamental value are sufficiently small to be considered negligible. If trading periods are at high frequency (say,
milliseconds), fundamentals hardly change from one such period to the next. Of course, we have in mind that the flow of trade itself does not affect fundamentals. The verification that $\sigma$ is a belief-free equilibrium follows exactly the same steps as in the main proof.

A third restriction is that long-term market participants do not take advantage of their private information, if any. This is an implication of our definition of belief-free equilibrium, which requires robustness to any possible information structure. What really matters for dealers is identifying the set of quotes that balance the supply and demand coming from the mass of investors. As these quotes can be ultimately learned from the observation of the trading flow, dealers’ private information is not crucial. The fact that in our equilibrium, dealers do not take advantage of their private information might be counter-intuitive, but there is no difficulty in re-defining our model to accommodate such behavior without abandoning the belief-free assumption altogether. Rather than taking the asset value as a primitive that determines a distribution over the players’ private signals, one can think of the players’ private signals as a primitive that determines the asset’s value. In that case, we can re-define a strategy profile to be belief-free if it is the case that for every player, given his private signal, his strategy (that can depend on his private signal) is optimal independently of the other players’ possible strategies. That is, given a player’s signal, there is a set of signal profiles of his opponents that are consistent with his; for each such signal profile, his opponents play some strategy profile. Belief-freeness requires the player’s strategy to be optimal against all of these profiles. In fact, it is clear that we do not need to require that the players’ combined signals pin down the value of the asset. Rather, it pins down a set of possible values, with respect to all of which, the best-reply property must hold.

This provides a natural extension of the definition of belief-free equilibrium that allows dealers to take advantage of their private information. While this extension of belief-free equilibrium can be characterized, it is clearly less robust to changes in the information structure; hence, it can help explain price formation only in situations in which information asymmetries are limited and known a priori. We believe that such an extension raises interesting questions and technical
8 Conclusion

This paper considers market microstructure models in which long-lived dealers interact with short-lived traders. We characterize equilibrium price formation strategies that are robust to the specification of dealers’ beliefs about fundamentals. Belief-free equilibria feature two key ingredients. First, dealers collectively learn from trading flow the value of those fundamentals that affect traders’ demand. Second, for any given value of these fundamentals, dealers generate positive profits from the intermediation of traders’ demand. This has three robust implications that contrast with those delivered by canonical microstructure models relying on the assumption of equally uninformed competitive dealers. First, dealers’ long-term profits are strictly positive independently of the asset’s fundamental value. These profits are obtained through the intermediation of traders’ demand. Second, the trading price need not reflect any of the dealers’ beliefs, and it is generally more volatile than prices that reflect Bayesian beliefs. Third, dealers’ inventories tend to be balanced even in the absence of risk aversion or institutional constraints. Given that belief-free equilibrium is more stringent than traditional solution concepts, it might be surprising that so much flexibility remains—in particular, the equilibrium is not unique. Hence, we have focused on a belief-free equilibrium with a simple Markovian structure. When applied to a version of the Glosten and Milgrom model, it explains well-documented stylized empirical facts. For specific microstructure games, it might therefore be reasonable to focus on belief-free equilibria that satisfy further criteria. For example, depending on the specific trading model considered, one could analyze equilibria that maximize the dealers’ aggregate payoff, those that minimize the expected time required for the market measure to point to the true state or even equilibria that minimize the aggregate cost of learning or, more generally, strategies that form a belief-free equilibrium for the lowest possible level of dealers’ patience.
Appendix

Proof of Proposition 1

1. **Positive maximum payoffs:** Any $\hat{\omega}$, it is sufficient to show that the set $A^*(\hat{\omega})$ is not empty. Fix dealer $i$ and consider the following two action profiles $a(i)$ and $a'(i)$ in which all dealers different from $i$ set the no-trade action. In $a(i)$, dealer $i$ sets his action so that if a trader trades, he can only buy at price $v(\hat{\omega}) + \rho$, with $0 < \rho < P$ and he cannot sell. Conversely, in $a'(i)$, dealer $i$ sets his action so that a trade can only consist of the trader selling at price $v(\hat{\omega}) - \rho$. Because of RTF 1. and 2., we have that for all $\omega \in \hat{\omega}$, the expected asset transfer to dealer $i$ is equal to $Q_i(\hat{\omega}, a(i)) > 0$ and $Q_i(\hat{\omega}, a'(i)) < 0$ for action $a(i)$ and $a'(i)$, respectively. Now, let $q_i := Q_i(\hat{\omega}, a'(i))/(Q_i(\hat{\omega}, a'(i)) - Q_i(\hat{\omega}, a(i))) \in [0,1]$ and consider $\tilde{a}(i)$ obtained by playing $a(i)$ with probability $q$ and $a'(i)$ with probability $1 - q$. This translates in dealer $i$’s expected payoff of $q_i Q_i(\hat{\omega}, a(i)) 2\rho > 0$ no matter the value of $e(\omega)$. In fact, in expectation, he buys $q_i Q_i(\hat{\omega}, a(i))$ shares for $v(\hat{\omega}) - \rho$, and he sells the same quantity for $v(\hat{\omega}) + \rho$ per share. Now, consider the random strategy $\tilde{a}$ obtained by first selecting a dealer $i$ with probability $1/n$ and then playing $\tilde{a}(i)$. This guarantees that $u_i(\omega, \tilde{a}) = q_i Q_i(\hat{\omega}, a(i)) 2\rho/n > 0$ for every $i$ and every $\omega \in \hat{\omega}$, no matter the value of the $e(\omega)$ component. Thus, $\tilde{a} \in A^*(\hat{\omega})$.

2. **Negative minimum payoffs:** Fix dealer $i$ and consider the action $\underline{a}(i)$ in which all dealers different from $i$ set the no-trade action. In $\underline{a}(i)$, dealer $i$ sets his action so that if a trader trades, he can only buy at a price strictly smaller than $v(\hat{\omega}) - \bar{e}$, and he cannot sell. Because of RTF 1., there will be a trader willing to sell at this price, implying that dealer $i$’s payoff is negative no matter the true value of $\omega$ and hence for all $\omega \in \hat{\omega}$. Consider the random strategy $\underline{a}(\hat{\omega})$ obtained by first selecting a dealer $i$ with probability $1/n$ and then playing $\underline{a}(i)$. Clearly, $u_i(\omega, \underline{a}(\hat{\omega})) < 0$ for all $\omega \in \hat{\omega}$ and dealer $i$.

3. **Non-positive expected payoffs:** Fix dealer $i$ and a distribution $\mu_\omega \in \Delta \hat{\omega}$. Let $W_{\mu_\omega} := v(\hat{\omega}) + \sum_{\omega \in \Delta \hat{\omega}} \mu_\omega(\omega) e(\omega)$ be the expected fundamental value of the asset computed using probability measure $\mu_\omega$. Let $p_1$ and $p_2$ be the two points on the price grid $G$ that are closest to $W_{\mu_\omega}$, with $p_1 \leq W_{\mu_\omega} \leq p_2$. Define $a'_{i-1}(\mu_\omega)$ as follows. Each dealer $j \neq i$ sets an action such that any other market participant can buy and sell up to the maximum tradable quantity at price $p = W_{\mu_\omega}$. Consider dealer $i$’s expected payoff when his belief
that the state is $\omega$ is equal to $\hat{\mu}(\omega)$. His expected payoff from playing $a_i$ when the other dealers play $a_{-i}(\hat{\mu})$ is

$$
\sum_{\omega \in \hat{\omega}} \mu(\omega)u_i(\omega, a_i, a_{-i}(\hat{\mu})) = \sum_{\omega \in \hat{\omega}} \mu(\omega)(v(\omega) + e(\omega))Q_i(\omega, a_i, a_{-i}) + P_i(\omega, a_i, a_{-i})
$$

$$
= W_{\mu}\hat{Q}_i(\hat{\omega}, a_i, a_{-i}) + P_i(\hat{\omega}, a_i, a_{-i})
$$

(18)

where the second equality follows from the fact that for any $\omega \in \hat{\omega}$ and $a \in A$, property LS implies that $v(\omega) = v(\hat{\omega})$, $Q_i(\omega, a) = Q_i(\hat{\omega}, a)$ and $P_i(\omega, a) = P_i(\hat{\omega}, a)$. The last expression can be interpreted as the payoff of a dealer who values the asset exactly $W_{\mu}$ and buys a quantity $Q_i(\hat{\omega}, a_i, a_{-i})$ of the asset in exchange for $P_i(\hat{\omega}, a_i, a_{-i})$. To see that this expression cannot be strictly positive, note first that if $a_i$ is such that dealer $i$ trades with some other dealer, the other dealers’ actions are such that he can only trade at price $p = W_{\mu}$, implying that dealer $i$ profit is nil. Suppose that $a_i$ is such that dealer $i$ trades with the trader. Because the trader can trade any quantity at price $p = W_{\mu}$ from the other dealers, he will trade with dealer $i$ only if he offers better trading conditions, that is, only if he can buy from dealer $i$ for less than $W_{\mu}$ or sell for more than $W_{\mu}$. In both cases, dealer $i$’s payoff of expression (18) cannot be strictly positive. Note that it could be that because of the price grid, trade cannot occur at a price exactly equal to $W_{\mu}$. In this case, it is possible for dealer $i$ to make some strictly positive payoff. However, this payoff can be made arbitrarily small for a sufficiently dense price grid.

4. Non-equivalent payoffs: Consider the strategy $\tilde{a}(i)$ defined in point 1 above. When dealers play $\tilde{a}(i)$, dealer $i$’s payoff is positive, whereas all other dealers’ payoff is nil. Let $a^i(\hat{\omega})$ be obtained by first selecting a dealer $j \neq i$ with probability $1/(n-1)$ and then playing $\tilde{a}(j)$. Because in this strategy, dealer $i$’s payoff is nil, whereas all other dealers’ payoff is strictly positive, we have $u_i(\omega, a^i(\hat{\omega})) < u_i(\omega, a^j(\hat{\omega}))$ for all $i \neq j$ and $\omega \in \hat{\omega}$.

\[\blacksquare\]

**Proof of Lemma 1**

A BFE must be an equilibrium even when dealers have no private information about $\omega$. In this case, the distribution over histories can be different in two states $\omega$ and $\omega'$ only if traders’ behavior differs in those two states. \[\blacksquare\]

**Proof of Lemma 2**
To see that the equilibrium payoff cannot be negative, suppose that for some state \( \omega \), dealer \( i \) and history \( h_t \), we have \( V_i(\omega, \sigma^* | h_t) < 0 \). Then, if dealer \( i \) believes that the true state is \( \omega \), he would be better off by deviating and playing the no-trade action guaranteeing him a nil payoff. Thus, \( \sigma^* \) cannot be a BFE. To see that the equilibrium payoff is strictly positive in most states, take any \( \hat{\omega} \in \hat{\Omega} \). Note that for all \( \omega \in \hat{\omega} \), \( v(\omega) \) is the same because of \( \text{LS} \), and the equilibrium play is the same because of Lemma 1. Thus, we have

\[
V_i(\omega, \sigma^* | h_t) = (v(\hat{\omega}) + e(\omega))Q_i(\hat{\omega}, \sigma^* | h_t) + P_i(\hat{\omega}, \sigma^* | h_t)
\]

Suppose \( Q_i(\hat{\omega}, \sigma^* | h_t) \neq 0 \) and take two different states \( \omega, \omega' \in \hat{\omega} \). Because \( e(\omega) \neq e(\omega') \), then it must be \( V_i(\omega, \sigma^* | h_t) \neq V_i(\omega', \sigma^* | h_t) \). Therefore, \( V_i(\omega, \sigma^* | h_t) = 0 \) for at most one state \( \omega \in \hat{\omega} \).

**Proof of Lemma 3**

Fix \( \hat{\omega} \in \hat{\Omega} \). Property \( \text{LS} \) implies that \( v(\omega) = v(\hat{\omega}) \) for all \( \omega \in \hat{\omega} \) and that knowing \( v(\omega) \) does not allow any inference about \( e(\omega) \). Because of Lemma 1, for any state \( \omega \in \hat{\omega} \), we have that the change in dealers’ aggregate inventory is the same \( Q(\hat{\omega}, \sigma^* | h_t) \). This change in inventory is mirrored by the change in traders’ aggregate inventory, i.e., \( Q(\hat{\omega}, \sigma^* | h_t) = -Q_{Tr}(\hat{\omega}, \sigma^* | h_t) \). Consider the dealers’ aggregate payoff. Because of property \( \text{RTF} \), traders will never buy for more than \( v(\hat{\omega}) + \bar{p} \) nor sell for less \( v(\hat{\omega}) - \bar{p} \). Hence, the dealers’ aggregate payoff cannot be greater than:

\[
(v(\hat{\omega}) + e(\omega) - (v(\hat{\omega}) - \bar{p}))Q^+(\omega, \sigma^* | h_t) + (v(\hat{\omega}) + e(\omega) - (v(\hat{\omega}) + \bar{p}))Q^-(\omega, \sigma^* | h_t),
\]

which is non-negative only if

\[
\bar{p} > \frac{Q(\hat{\omega}, \sigma^* | h_t)e(\omega)}{Q^+(\hat{\omega}, \sigma^* | h_t) - Q^-(\hat{\omega}, \sigma^* | h_t)}
\]

Lemma 2 requires each dealer’s payoff to be non-negative and hence a fortiori their aggregate payoff. Therefore, the above expression must be positive for all realizations of \( e(\omega) \). Then, the result follows from the fact that \( Q, Q^+ \) and \( Q^- \) do not depend on \( e(\omega) \) and that \( e(\omega) \in [-\bar{e}, \bar{e}] \).

**Proof of Lemma 4**

Suppose that after some history \( h_t \), dealers’ actions never belong to \( A(\hat{\omega}, \hat{\omega}') \), that is, in all
periods $t' > t$, the dealers' action profile $a' \notin A(\hat{\omega}, \hat{\omega}')$, implying $F(\hat{\omega}, a', s) = F(\hat{\omega}', a', s)$ for all $s \in S$. In other words, after $h^t$, the observation of traders' behavior does not distinguish $\hat{\omega}$ from $\hat{\omega}'$. Now, the measurability of the equilibrium play with respect to traders' behavior (Lemma 1) implies that the equilibrium occupation measure after $h^t$ is the same no matter whether the true state $\omega$ is in $\hat{\omega}$ or $\hat{\omega}'$. Let $\tilde{a}(h^t)$ be this occupation measure. Because $A^*(\hat{\omega}) \cap A^*(\hat{\omega}') = \emptyset$, one must have either $\tilde{a} \notin A^*(\hat{\omega})$ or $\tilde{a} \notin A^*(\hat{\omega}')$ or both. However, this implies that after history $h^t$, dealers' continuation payoffs are strictly negative for some $\omega$. Thus, the continuation strategy cannot be a BFE because it would contradict Lemma 2. □

Proof of Theorem 1

Let $\pi > 0$, finite, be such that $|u_i(\omega, a)| < \pi$, for all dealers $i$, $\omega \in \Omega$, and $a \in \Delta A$. A dealer's reward cannot be smaller than $-\pi$. Let $u^* > 0$ denote the minimum possible reward for a dealer when $a \in A^*(\hat{\omega})$ and the true state $\omega \in \hat{\omega}$. Formally, for all $i$ and $\hat{\omega} \in \hat{\Omega}$, one has $u_i(\omega, a) > u^*$, as long as $\omega \in \hat{\omega}$ and $a \in A^*(\hat{\omega})$.

Fix a profile $(\phi, \sigma)$ satisfying the assumptions of the theorem, and let $\omega$ be the true state. We first show that on the equilibrium path, each dealer's payoff is strictly positive. Consider the play on the equilibrium path. Let $q^t$ be the probability that at time $t$, the market measure satisfies $\pi^t(\hat{\omega}(\omega)) > 1 - \varepsilon$. Because $(\phi, \sigma)$ is $\varepsilon$-exploiting, Definition 3 implies that with a probability $q_t$, dealer $i$'s stage $t$ reward is at least $(1 - \varepsilon)u^* - \varepsilon\pi$. Then, at time $\tau \geq 0$, dealer $i$'s payoff satisfies

$$V^\delta_i(\omega, \sigma|h^\tau) > (1 - \delta) \sum_{t=\tau}^{\infty} \delta^{t-\tau} (q^t((1 - \varepsilon)u^* - \varepsilon\pi) - (1 - q^t)\pi)$$

$$= (1 - \varepsilon)(u^* + \pi)(1 - \delta) \sum_{t=\tau}^{\infty} \delta^t q^{t-\tau} - \pi. \quad (19)$$

Now, because $(\phi, \sigma)$ is $\varepsilon$-exploring, from Definition 2, one has that

$$\Pr_{\omega, \sigma} \left[ \lim_{\delta \to 1} (1 - \delta) \sum_{t=\tau}^{\infty} \delta^t q^{t-\tau} > 1 - \varepsilon \right] > 1 - \varepsilon. \quad (20)$$

Hence, we have that

$$\lim_{\delta \to 1} V^\delta_i(\omega, \sigma|h^\tau) > (1 - \varepsilon)^3(u^* + \pi) - (1 + \varepsilon)\pi. \quad (21)$$

As the r.h.s. is strictly positive for $\varepsilon = 0$, it is also positive for all $\varepsilon$ smaller than some $\bar{\varepsilon} > 0$. The continuity of $V^\delta_i$ in $\delta$ implies there exists $\tilde{\delta} < 1$ such that for $\varepsilon < \bar{\varepsilon}$, dealer $i$'s continuation
payoff $V^*_i(\omega, \sigma|h_\epsilon)$ is strictly positive.

The next step is to show that dealers have no profitable deviations. For this purpose, we first establish a simple lemma.

**Lemma 5** For any $\hat{\omega} \in \hat{\Omega}$, all $\omega \in \hat{\omega}$, and all $a \in A^*(\hat{\omega})$, there exist $n$ action profiles $\{\tilde{a}^1(\hat{\omega}), \ldots, \tilde{a}^n(\hat{\omega})\} \in [\Delta A]^n$ such that

$$0 < u_i(\omega, \tilde{a}^i(\omega)) < u_i(\omega, \tilde{a}^j(\omega)) < u_i(\omega, a),$$  \hspace{1cm} (22)

for all $i \neq j$.

**Proof.** Consider the convex combination

$$\tilde{a}^i(\hat{\omega}) := \beta_1(\hat{\omega})\beta_2(\hat{\omega})a(\hat{\omega}) + \beta_1(\hat{\omega})(1 - \beta_2(\hat{\omega}))a^i(\hat{\omega}) + (1 - \beta_1(\hat{\omega}))a, \hspace{1cm} (23)$$

for some $\beta_1(\hat{\omega}), \beta_2(\hat{\omega}) \in [0, 1]$, where $a(\hat{\omega})$ is defined in Proposition 1-2, and $a^i(\hat{\omega})$ is as in Proposition 1-4. Note that the $n$ action profiles $\{\tilde{a}^i(\hat{\omega})\}_{i=1,\ldots,n}$ also satisfy the condition in Proposition 1-4 as long as $\beta_1(\hat{\omega}) > 0$, $\beta_2(\hat{\omega}) < 1$. Furthermore, because $u(\omega, a(\hat{\omega})) < 0$, we can pick $\beta_2(\hat{\omega})$ close enough to one, and $\beta_1(\hat{\omega})$ is close enough to zero to guarantee that all payoffs are between 0 and $u(\omega, a)$.

We may now define $n$ partial strategy profiles $\sigma^i_{\epsilon}$ as follows. Let $A_L$ denote a set of learning action profiles satisfying $A(\hat{\omega}, \hat{\omega}') \cap A_L \neq \emptyset$ for each couple $\hat{\omega} \neq \hat{\omega}'$. Let $L$ denote the cardinality of $A_L$, and let $D_{\omega}$ denote the Dirac measure attaching probability 1 to $\hat{\omega}$. If $h^t$ is such that $\|\pi^t - D_{\hat{\omega}}\| < \epsilon$, then let $\sigma^i_{\epsilon}(h^t) = (1 - \epsilon)\tilde{a}(\hat{\omega}) + (\epsilon/L)\Sigma_{a \in A_L}a$. For all other $h^t$, let $\sigma^i_{\epsilon}(h^t) = (1/L)\Sigma_{a \in A_L}a$. Note that the strategy $\sigma^i_{\epsilon}$ is exploratory, it is such that all dealers’ continuation payoffs are strictly positive, but dealer $i$ would prefer the play to follow strategy $\sigma^{j,\epsilon}$ rather than $\sigma^i_{\epsilon}$, with $j \neq i$.

In addition, define $n$ partial “punishment” strategies $\sigma^i_{\epsilon}$ as follows.

Fix any $\hat{\omega} \in \hat{\Omega}$. Proposition 1-3 guarantees that we can extend Blackwell (1956)’s approachability argument to the discounted case: for any $\eta > 0$, there is $\delta^\eta < 1$, $m^\eta < \infty$ and $m^\eta$-period strategy $a_{-i}(\dot{\omega})$ for player $-i$ such that if $\delta > \delta^\eta$, for any sequence $\{a_1^1, \ldots, a_i^{m^\eta}\}$, player $i$’s discounted payoff during these $m^\eta$ periods is smaller than $\eta$ in each $\omega \in \dot{\omega}$. This strategy is then an ingredient for the punishment partial strategy $\sigma^i_{\epsilon}$. If $h^t$ is such that for some $\dot{\omega}_i$, $\pi^t$ assigns probability no more than $\epsilon$ to states outside of $\dot{\omega}_i$ but probability at least $\epsilon$ to all $\omega \in \dot{\omega}_i$, then $\sigma^i_{\epsilon}(h^t) = (1 - \epsilon)a^i(\dot{\omega}_i)(h^t) + (\epsilon/L)\Sigma_{a \in A_L}a$, where $a_{-i}(\dot{\omega}_i)(h^t)$ as defined above and $a^i(\dot{\omega}_i)$
is some fixed action. Note that for $\varepsilon > 0$, each of these strategies is exploratory. Furthermore, given any $\sigma_i$, any $\omega$, and any history $h^t$, the continuation payoff $V^\delta_i (\omega, \sigma_i, \sigma^i_{-i} | h^t)$ is such that

$$\lim_{\delta \to 1, \varepsilon \to 0} V^\delta_i (\omega, \sigma_i, \sigma^i_{-i} | h^t) \leq 0.$$  \hspace{1cm} (24)

From here on, the proof is standard; see Fudenberg and Maskin (1986). Given the partial strategy $\sigma$, define a strategy $\hat{\sigma}$ as follows. As long as no player unilaterally deviates, actions are specified by $\sigma$. As soon as a player (say $i$) unilaterally deviates, play proceeds according to $\sigma^{i,\varepsilon}$ for $T$ periods (for some $\varepsilon > 0$, $T \in \mathbb{N}$ to be specified). If during this $i$-punishment phase, some player (say $j$) unilaterally deviates from $\sigma^{i,\varepsilon}$, play switches to the $j$ punishment phase, in which $\sigma^{j,\varepsilon}$ is played for $T$ periods. If $T$ periods elapse without unilateral deviation during the $i$-punishment phase, play is then given by $\sigma^{i,\varepsilon}$. If there is a unilateral deviation from $\sigma^{i,\varepsilon}$ by $j$, play switches to the $j$-punishment phase, etc. It is now standard to show that for $T$ that is sufficiently large, and $\varepsilon$ that is sufficiently small, there exists $\delta \in (\hat{\delta}, 1)$ such that for all $\delta \in (\hat{\delta}, 1)$, players do not gain from deviating.

Note that this construction yields a belief-free equilibrium: The strategy is optimal irrespective of dealers’ beliefs about $\omega$ on and off the equilibrium path.

References


