Price Discovery for Derivatives

Christian Keller Michael C. Tseng*

Latest Version Available Here

Abstract

We obtain a basic theory of price discovery across derivative markets with respect to higher-order information, using a model where an agent with general private information regarding state probabilities is allowed to trade arbitrary portfolios of state-contingent claims. In an equivalent options formulation, the informed agent has private information regarding arbitrary aspects of the payoff distribution of an underlying asset and is allowed to trade arbitrary option portfolios. We characterize, in closed form, the informed demand, price impact, and information efficiency of prices. Our results offer a theory of insider trading on higher moments of the underlying payoff as a special case. The informed demand formula prescribes option strategies for trading on any given moment and extends those used in practice for, e.g. volatility trading.

JEL classification: C61; G13; G14.

Keywords: Derivatives; Price Discovery; Higher Moments Trading; Insider Smile.

*Christian Keller is from the Department of Mathematics, University of Central Florida. Michael Tseng is from the Department of Economics, University of Central Florida. We thank Roberto Burguet, Michael Caputo, Alain Chaboud, Yannick Dillschneider, Michael Gallmeyer (discussant), Sergei Glebkin (discussant), Thomas Gresik, Vladimir Gatchev, Kei Kawakami, Melody Lo (discussant), Christo Pirinsky, Marzena Rostek, and conference participants of the 2022 Northern Finance Association Meeting, 2022 Midwest Economic Theory Meeting, 2022 Econometric Society European Winter Meeting, and 10th HEC-McGill Winter Finance Workshop for questions, comments, and suggestions. We are responsible for all remaining errors. A previous version of the paper was circulated under the title "Arrow-Debreu Meets Kyle." The research of Christian Keller was supported in part by NSF-grant DMS-2106077. Corresponding author: Michael Tseng, E-mail: michael.tseng@ucf.edu. Address: Department of Economics, College of Business, University of Central Florida, 12744 Pegasus Dr. Orlando, Florida 32816-1400.

1 Introduction

A basic economic function of prices is to convey information. As informed agents trade to exploit their private information regarding the payoff of an asset, agents' private information is impounded into the asset price. However, agents may have private information beyond just the (expected) asset payoff, such as higher-order moments of the payoff distribution. Exploiting such higher-order information requires trading derivatives. For example, an informed agent trading on volatility information may need to trade two or more suitable option contracts. As financial markets matured, such derivatives trades have long become common practice. Indeed, the significant size and importance of derivative markets derive from their principal utility of providing investors with exposure to aspects of asset payoff distributions beyond the expected payoffs. This means higher-order information is impounded into derivatives prices. Yet, there has been no theory of derivatives price discovery that articulates the microfoundation of this basic economic logic. This paper considers a general tractable model that captures the fundamental aspects of derivatives price discovery—the nonlinear nature of both higher-order information and derivatives payoffs, and the resulting strategic interdependence of trades and prices within and between derivative markets. We characterize, in closed form, the informed demand, price impact within and between markets, and information efficiency of prices.

In our model, the informed trader has private information regarding the probabilities of future states of the world and is allowed to trade arbitrary portfolios of state-contingent claims. In other words, the informed trader has private information regarding the payoffs of Arrow-Debreu securities and can trade in the Arrow-Debreu markets. No assumption is made on the informed trader's private information. The model is the simplest possible model of general price discovery across contingent claim markets and encompasses various derivative market settings.

When only the underlying asset is traded and there are no derivative markets, our model specializes to the single asset model of Kyle (1985).[1] In this single asset setting, the uninformed trader has a linear supply schedule. The slope of the supply schedule—known as *Kyle's lambda*—characterizes equilibrium and is proportional to the (exogenous) variation of the asset payoff across signals. In the general contingent claims setting, the endogenous relationships between quantities and prices, under information asymmetry, emerge both *within* and *between* markets. The price of a given security depends not only on the quantity demanded for that security itself but on the quantities demanded across the *entire set* of securities. For two states x and y, the quantity demanded for security y in general causes

[1] Previous extensions of Kyle (1985) include, for example, Caballe and Krishnan (1994), Foster and Viswanathan (1996), Back et al. (2000), Rochet and Vila (1994), and Collin-Dufresne and Fos (2016). Caballe and Krishnan (1994), Foster and Viswanathan (1996), and Back et al. (2000) consider settings where multiple insiders have differential private information about the fundamental values of a finite set of assets. Rochet and Vila (1994) drops the normality assumption of Kyle (1985) on asset value and assumes the insider observes noise trades. Collin-Dufresne and Fos (2016) consider a setting where noise trading intensity is a stochastic process. None considers derivatives or private information beyond the (first moment of) asset fundamental value.

price adjustment for security x, to the extent that the demand for y reflects the informed trader's private information regarding x. The set of supply schedules for the Arrow-Debreu securities is equivalent to a *pricing kernel* that prices all state-contingent claims, e.g. options.

Generalizing the intuition of Kyle's lambda, we show that the marginal effect of the quantity demanded for one security on the price of another security is proportional to the covariance of the payoffs of the two securities. Kyle's lambda generalizes to an infinite-dimensional matrix of (endogenous) payoff covariances.[2] The informed trader's trading strategy also admits a very intuitive characterization. The informed trader's optimal portfolio of state-contingent claims is a linear combination of possible payoff distributions, with positive weight for the true payoff distribution, which is privately known to him, and negative weights for the other distributions. In other words, the informed trader buys the true payoff distribution, as a contingent claim portfolio, and sells the other distributions.

While certain parallels can be drawn with the single asset special case, our general setting also makes clear the former's incompleteness. In general, the nonlinear nature of higher-order information means the optimal portfolio is also nonlinear, i.e. requires derivatives. The absence of derivative markets restricts the informed trader to sub-optimal linear demands and therefore reduces his welfare.[3] This is the case even when private information only concerns the expected payoffs (as in Kyle (1985)), because derivatives allows the informed trader to better hide his private information by reducing his demand for low probability states.[4]

As a consequence of the fact that price discovery takes place jointly across all markets, the information efficiency of all contingent claim prices can be summarized in a single measure. We show that the information efficiency of prices does not depend on the specification of possible payoff distributions.[5] That is, the extent to which higher-order information is impounded into contingent claim prices is independent of the specification of private information. This general result points to a basic aspect of complete markets that has so far gone unnoticed. Just as risk is always shared efficiently in a complete market, information is always (partially) impounded into prices in a complete market, and the extent to which this occurs is *invariant* with respect to the nature of information.[6]

By the standard Breeden-Litzenberger (1978) formula, our model has an equivalent options formulation.[7] In this options formulation, the informed trader has private information regarding arbitrary aspects of the payoff distribution of the underlying asset, under arbitrary specification of the possible payoff distributions, and is allowed to trade arbitrary option portfolios. Private information regarding

[2]In contrast to the single asset setting, the information spillover across markets endogenizes the variance-covariance matrix of asset payoffs.

[3]As a corollary, trading derivatives increases the loss of noise traders whose trades are unrelated to the asset payoff.

[4]See Example 6.1.

[5]Nor does the information efficiency of contingent claim prices depend on the amount of noise trading that is taking place.

[6]See Remark 6.12.

[7]See Breeden and Litzenberger (1978) for the Breeden-Litzenberger formula.

high-order moments is a special case. This is the first paper to obtain a unified theory of insider trading on high-order moments.

The closed-form formula for inform demand prescribes option strategies for trading on any given moment. For trading on volatility and skewness, the informed demand formula shows that the option strategies commonly used in trading practice can indeed be rationalized in the equilibrium setting.[8] Due to its wide usage, the option strategy for trading on high volatility, known as the straddle, has also been incorporated in various empirical elements of the literature—see, for example, Coval and Shumway (2001) and Driessen et al. (2009). Our results provide microfoundation for such empirical considerations and extend their scope.

Our theory has empirical implications for options price discovery and the cross-section of option returns.[9] For options price discovery with respect to higher-order moments, our theory provides adverse selection measures that can be estimated from data—namely, the cross-market price impact between prescribed option contracts. The within-market price impact is a well-established adverse selection measure in the market microstructure literature.[10] It is generalized to higher-order moments by the cross price impact measure. This generalization leads to new empirical hypotheses regarding the predictability of higher moments of the underlying payoff by options cross price impact. For example, high cross price impact between the put-call pair in a straddle means that the (future) volatility of underlying payoff is currently underpriced. In turn, this predicts an increase in the volatility of underlying.

For the cross-section of option returns, our theory suggests candidate systematic factors reflecting the extent to which higher-order moment price discovery takes place across option markets. For example, if options price discovery takes place with respect to certain higher moments, the returns of option portfolios for trading on those moments should help explain the cross-sectional variation of option returns. In this paper, we specify the equivalent options formulation and indicate avenues for further analysis.[11]

Allowing arbitrary private information and arbitrary demand for derivatives means that the Bayesian trading game between the informed and uninformed traders is infinite-dimensional. The informed trader's demand and the uninformed trader's pricing kernel are both infinite-dimensional, and nonlinear, objects. Tractability in this general setting is made possible by appropriate rigorous modelling and methodology. We show that the game can be reduced to an isomorphic finite-dimensional game. The finite-dimensional game is canonical, in the sense that it is independent of the specification of payoff distributions. This finite-dimensional reduction yields a closed-form equilibrium solution characterized

[8] See Examples 6.2 and 6.3.

[9] See Section 7.2.

[10] See, for example, the price impact regressions employed in Glosten and Harris (1988), Lin et al. (1995), Huang and Stoll (1997), Goyenko et al. (2009), Hendershott et al. (2011), and Makarov and Schoar (2020).

[11] Applications to insider trading on high-order moments and option markets within our general framework are to be explored in future papers.

by *one* endogenous constant.

Our chosen setting is space-continuous—as in many models of asset pricing and derivative pricing—and time-static. Standard techniques from stochastic analysis used to solve continuous-time models of insider trading—e.g. Back (1993)—are not applicable. A rigorous formulation of the uninformed trader's infinite-dimensional Bayesian inference problem is not possible with the standard construction of stochastic integral, and we adapt tools from rough paths theory to construct a pathwise stochastic integral.[12]

The continuous-space assumption is made in the same spirit as the standard continuous-time assumption in dynamic models. It yields a simple intuitive equilibrium in a general model. In the options formulation, this assumption means there are option markets corresponding to all possible strikes. Actual option strike grids can be viewed as discretizations of this stylized setting. Indeed, certain widely traded synthetic derivatives are constructed using options by discretizing the same Breeden-Litzenberger formula used in our formulation.[13] Moreover, the option strategies for trading on higher moments prescribed by the informed demand formula can be implemented using finitely many options.

The empirical literature has documented that options orders and trades contain higher-order information regarding the underlying asset's payoff distribution—see, for example, Cao et al. (2005), Pan and Poteshman (2006), Ni et al. (2008), Goyal and Saretto (2009), Roll et al. (2009), and Augustin et al. (2019).[14] On the other hand, existing theory literature on informed options trading primarily compares aspects of the market with and without an option, with information asymmetry restricted to asset fundamental value in all cases—see, for example, Back (1993), Biais and Hillion (1994), Easley et al. (1998), and Collin-Dufresne et al. (2021).[15] None considers informed trading in more than one option or speaks to the higher-order information content of options orders and trades.[16] This paper

[12]Rough paths theory originates from probability. See Lyons et al. (2007), Hairer (2014), and Friz and Hairer (2020).

[13]One such synthetic derivative is the Cboe VIX. See `https://www.cboe.com/tradable_products/vix/`.

[14]Pan and Poteshman (2006) and Ni et al. (2008) show that option trading volume contains information on future stock returns and volatility. Goyal and Saretto (2009) show that mispricing in the second moment can be traded on using the straddle. Roll et al. (2009) find that option-to-stock volume is closely related to firm value and positively associated with absolute returns of post-earnings announcements, pointing to the informational role played by the options markets. Cao et al. (2005) and Augustin et al. (2019) document abnormal options trading volume and predictability of options order flow around specific corporate events.

[15]Back (1993) shows that information asymmetry can lead to market incompleteness by introducing one option in a continuous-time setting, where the market is already complete with no options in the absence of information asymmetry. Biais and Hillion (1994) consider the implication of introducing an option when liquidity trades are endogenous. Easley et al. (1998) consider informed option trading when the asset value has two possible realizations. Collin-Dufresne et al. (2021) study information flow from stock markets to option markets when informed trading occurs only in the stock market.

[16]Chabakauri et al. (2022) also consider informed trading in Arrow-Debreu markets. They impose a parametric assumption on the Arrow-Debreu payoff distributions in a rational expectations setting, where traders are price takers. We impose no parametric assumption on the Arrow-Debreu payoff distributions in a strategic setting, where traders have price impact. In our setting, the informed trader can have private information on arbitrary higher-order moments and trades a complete menu of options, while their setting restricts informed trading beyond the first moment to one

bridges the gap between the theory and empirical literatures on informed options trading.

In the option pricing and financial econometrics literatures, large classes of models are concerned with the modelling of volatility information contained in option prices, as expressed by their (Black-Scholes) implied volatilities, which are known to exhibit the *volatility smile* across strikes. The class of stochastic volatility models, first introduced by Heston (1993), is both assessed and estimated by relating option prices to the model parameters via the volatility smile.[17] The same approach drives the development of local volatility models.[18] Despite being the focus of an important (reduced-form) modelling framework, no economic mechanism has been put forth to explain how the volatility smile may arise. Within our setting, the volatility smile arises through options price discovery of volatility. As the informed trader trades on volatility information, the equilibrium implied volatilities form an *insider smile*. This microfounded explanation may help shed light on the misspecification issues that reduced-form models aim to address.

The rest of the paper is organized as follows. Section 2 summarizes our model setup. Section 3 formulates the uninformed trader's Bayesian inference problem. Section 4 analyzes the informed trader's portfolio choice problem. Section 5 establishes equilibrium. Section 6 provides basic characterizations of Arrow-Debreu price discovery—the informed demand, price impact, and information efficiency of prices. Section 7 discusses the insider smile, empirical implications of our theory, and the case of risk-averse agents. Section 8 concludes. The Appendix contains proofs.

2 Model

2.1 General Setup

There are two risk-neutral agents, the insider and the market maker. There is a risk-free asset in perfectly elastic supply with risk-free rate zero. At $t = 0$, a signal $s \in S$ realizes, which informs the observer of the probability distribution across possible $t = 1$ states of the world. Only the insider observes the signal.[19] The market maker has a Bayesian prior $\pi_0(ds)$ on the set S of possible signals. At $t = 0$, there is a complete market of Arrow-Debreu (AD) securities for possible $t = 1$ states. The structure of the economy is common knowledge.

After observing the signal at $t = 0$, the insider submits his demand, i.e. market orders, for AD

volatility derivative. Their setting does not cover, for example, insider trading on volatility using options.

[17]For calibration of stochastic volatility models by the volatility smile, see, for example, Bates (1996), Duffie et al. (2000), and Britten-Jones and Neuberger (2000). For estimation and model selection using option prices, see Bates (2000), Aït-Sahalia and Kimmel (2007), Christoffersen et al. (2010), and Aït-Sahalia et al. (2021).

[18]See, for example, Berestycki et al. (2002) and Carr and Cousot (2012).

[19]The assumption that the insider observes the entire probability distribution is without loss of generality. If the insider has only partial information, which may be heterogeneous across signals, conditioning on his information signal-by-signal returns to our setting.

securities to maximize his expected utility of $t = 1$ consumption. There are noise traders who trade for exogenous (e.g. liquidity) reasons. The insider and noise traders face no leverage or short-selling constraints. The market maker observes the aggregate order flow of the insider and noise traders across all AD markets, and sets the prices to be the ex-ante zero-profit prices, according to his belief regarding the insider's trading strategy.[20]

The set (S, π_0) of possible signals is a Borel probability space where the probability measure π_0 is the market maker's prior. The set of possible $t = 1$ states is indexed by an interval $[\underline{x}, \bar{x}]$.[21] Conditional on the signal s, the probability distribution across possible $t = 1$ states is therefore specified by a density $\eta(\,\cdot\,, s) \colon [\underline{x}, \bar{x}] \to \mathbb{R}$. After observing s, the insider chooses a portfolio $W(\,\cdot\,, s) \colon [\underline{x}, \bar{x}] \to \mathbb{R}$, where $W(x, s)$ is the insider's order for state x security.

Noise trader orders are normally distributed with mean zero in each market and independent between markets. Passing to the continuum, noise trader orders across an infinitesimal increment of states $[x, x+dx]$ are given by

$$dN_x = \sigma(x) dB_x$$

where (B_x) is a standard Brownian motion, with the index x interpreted as state (rather than time), and $\sigma(x)$ is the noise trading intensity at state x market.

In the event that the insider submits orders $W(\,\cdot\,, s)$ after observing signal s, the total cumulative order flow received by the market maker is therefore a sample path ω of the Itô process $(Y_x)_{\underline{x} \le x \le \bar{x}}$ specified by[22]

$$\underbrace{dY_x}_{\substack{\text{total order} \\ \text{over } [x, x+dx]}} = \underbrace{W(x, s) dx}_{\substack{\text{insider order} \\ \text{over } [x, x+dx]}} + \underbrace{dN_x}_{\substack{\text{noise order} \\ \text{over } [x, x+dx]}} = W(x, s) dx + \sigma(x) dB_x. \tag{1}$$

[20]Implicitly, there is Bertrand competition among market makers.

[21]The interval need not be finite. $[\underline{x}, \bar{x}]$ can be replaced by $(-\infty, \infty)$.

[22]**Discrete States Analog** Following standard notation, the stochastic process (1), indexed by x, is specified in differential form. We emphasize that the market maker observes a sample path ω of the specified process $(Y_x)_{\underline{x} \le x \le \bar{x}}$ *over the entire set* $[\underline{x}, \bar{x}]$ *of states*. This means that the market maker observes the aggregate demand of the insider and noise trader *for each AD market* x. In the analogous setting with discrete states, x_i, $i \ge 1$, the market maker would observe the discrete sample path $(Y_i)_{i \ge 1}$, i.e. total cumulative order flow

$$Y_0 = 0,\ Y_i = \sum_{j \le i} W(x_j, s) + N_j,\ i \ge 1,$$

where $W(x_j, s)$ and $N_j \overset{d}{\sim} \mathcal{N}(0, \sigma^2(x_j))$ are the insider order conditional on s and noise order, respectively, for state x_j. This is equivalent to observing the first differences, i.e. the discrete analog of dY_x of (1),

$$\Delta Y_i = W(x_i, s) + N_i,\ i \ge 1,$$

which is the aggregate demand of the insider and noise traders for each AD market x_i.

While the discrete states setting does not admit the same tractability or generality, it can be approximated in the continuous states setting by discretization. See Section 2.4.

After receiving total order flow ω, the market maker updates his prior $\pi_0(ds)$ according to his belief $\widetilde{W}(\cdot\,,\cdot)\colon [\underline{x}, \bar{x}] \times S \to \mathbb{R}$ regarding the insider's trading strategy. According to the market maker's posterior $\pi_1(ds, \omega; \widetilde{W})$ conditional on ω and given his belief $\widetilde{W}$, the state x price set by the market maker is

$$P(x, \omega; \widetilde{W}) = \int_S \eta(x, s)\pi_1(ds, \omega; \widetilde{W}).$$

Conditional on observing s and given market maker belief $\widetilde{W}(\cdot\,,\cdot)$, the insider maximizes his expected utility of consumption over possible $t = 1$ states:

$$\max_{W(\cdot)} \mathbb{E}^{\mathbb{P}_W}[\int_{\underline{x}}^{\bar{x}} (\eta(x, s) - P(x, \omega; \widetilde{W})) \cdot W(x)dx], \tag{2}$$

where the expectation $\mathbb{E}^{\mathbb{P}_W}[\cdot]$ is taken over possible realizations of total order flow ω under the probability law $\mathbb{P}_W$ of the Itô process (1) induced by insider portfolio $W(\cdot)$.

In equilibrium, the optimal trading strategy of the insider, given the market maker's pricing kernel $P(\cdot\,, \omega; W^*)$ based on his belief $W^*(\cdot\,,\cdot)$ regarding insider's trades, coincides with $W^*(\cdot\,,\cdot)$.[23] That is, the optimal portfolio of the insider after observing signal s is $W^*(\cdot\,, s)$, for each $s \in S$, thereby confirming the market maker's belief. In equilibrium, the market maker's prior and posterior probabilities form a two-period martingale. Therefore so do the $t = 0$ and $t = 1$ prices at each AD market x.

2.2 Specializing to Kyle (1985)

When the insider's private information is the asset fundamental value (equivalently, expected asset payoff) and there are no derivative markets, our model reduces to the static case of the Kyle (1985) single asset model.

Let the signal $s \in S$ be the asset fundamental value and $\pi_0(ds)$ be the market maker's prior distribution of s. A trading strategy of the insider is a map $W\colon S \to \mathbb{R}$ where, after observing s, the insider submits a market order $W(s)$ for the asset. If the insider submits order W, the market maker observes $\omega = W + \varepsilon$, where ε is the noise trader order. The market maker has a belief $\widetilde{W}\colon S \to \mathbb{R}$ regarding the insider's strategy, computes his posterior $\pi_1(ds, \omega; \widetilde{W})$ regarding the fundamental value conditional on ω given his belief $\widetilde{W}(\cdot)$, and sets the asset price to be the zero profit price

$$P(\omega; \widetilde{W}) = \int_S s \cdot \pi_1(ds, \omega; \widetilde{W}).$$

[23]The formal statement of equilibrium is given in Definition 3.

Conditional on observing s and given market maker belief $\widetilde{W}(\cdot)$, the insider maximizes expected profit

$$\max_{W} \mathbb{E}^{\mathbb{P}_W}[(s - P(\omega; \widetilde{W})) \cdot W]$$

where the expectation $\mathbb{E}^{\mathbb{P}_W}[\cdot]$ is taken with respect to the distribution $\mathbb{P}_W$ of total order flow ω. An equilibrium is a trading strategy $W^*(\cdot)$ such that

$$W^*(s) = \operatorname*{argmax}_{W} \mathbb{E}^{\mathbb{P}_W}[(s - P(\omega; W^*)) \cdot W], \forall s. \tag{3}$$

This is a special case of the equilibrium described in Section 2.1 when there are no derivative markets.

Assume $s \overset{d}{\sim} \mathcal{N}(v_0, \sigma_v^2)$ and $\varepsilon \overset{d}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2)$, where v_0 and σ_v^2 are the prior mean and variance, respectively, of asset fundamental value, and σ_ε is the noise trading intensity. Restricting to linear strategies,

$$W^*(s) = \beta(s - v_0), \text{ and } P(\omega; W^*) = v_0 + \lambda\omega, \tag{4}$$

we recover the equilibrium obtained in Theorem 1 of Kyle (1985):

$$\beta = \frac{1}{2\lambda} = \frac{\sigma_\varepsilon}{\sigma_v}, \text{ and } \lambda = \frac{\sigma_v}{2\sigma_\varepsilon}. \tag{5}$$

λ is referred to as *Kyle's lambda*.

In this single asset setting, the *price impact*—the marginal effect of quantity on price—is the slope of the market maker's linear supply curve $P(Q) = v_0 + \lambda Q$ specified by (4) and (5). Price impact λ is proportional to the variation σ_v of the asset fundamental value. Higher variation of the asset fundamental value implies higher variation of the insider's trades across signals, which makes order flow more informative for the market maker, leading to higher price impact. In other words, price impact is determined by the information intensity of the insider's private signal. This intuition will generalize to the contingent claims setting, both within and between markets.[24]

2.3 Options

The equivalent options formulation of our model is as follows. Interpret the states as possible $t = 1$ spot prices of an underlying asset with the consumption good as numeraire, and assume there are $t = 0$ markets for a complete menu of put and call options maturing at $t = 1$. Let K_0 be the market maker's

[24] $\frac{d}{dQ}P = \frac{d}{dW}\mathbb{E}[P] = \lambda$ generalizes to the price impact expression $\frac{\partial}{\partial W(y)}\overline{P}(x, W; \widetilde{W})$ of Corollary 4.3. The equilibrium characterization of price impact in the general setting is given in Section 6.2.

prior expected payoff of the asset,

$$K_0 = \int_S \int_{\underline{x}}^{\bar{x}} x\, \eta(x,s) dx\, \pi_0(ds),$$

i.e. K_0 is the $t = 0$ futures price set by the market maker. By the decomposition formula of Breeden and Litzenberger (1978),

$$W(x,s) = W(K_0,s) + W'(K_0,s)(x - K_0) + \int_{\underline{x}}^{K_0} W''(K,s)(x-K)_- dK \qquad (6)$$
$$+ \int_{K_0}^{\bar{x}} W''(K,s)(x-K)_+ dK,$$

a portfolio $W(\cdot, s)$ of AD securities is equivalent to a portfolio consisting of positions $W(K_0,s)$ in the risk-free asset, $W'(K_0,s)$ in futures, and $W''(K,s)$ in out-of-the-money put (resp. call) options at strike $K < K_0$ (resp. $K > K_0$).

The insider submits futures and options orders after observing his private signal. Noise trades for futures follows a normal distribution $\mathcal{N}(0, \sigma_f^2)$, where σ_f is noise trading intensity in the futures market. If the insider submits futures order $W'(K_0,s)$ after observing signal s, the total futures order flow received by the market maker is a realization ω^f from the distribution $\mathcal{N}(W'(K_0,s), \sigma_f^2)$.[25] Noise trades for put (resp. call) options across an infinitesimal increment $[x, x+dx]$ of strikes follows a Brownian increment $\sigma(x)dB_x^p$ (resp. $\sigma(x)dB_x^c$), where $\sigma(x)$ is the noise trading intensity at strike x. If the insider submits put and call options orders $((W''(x,s))_{x<K_0}, (W''(x,s))_{x>K_0})$, the total options order flow received by the market maker is a sample path (ω^p, ω^c) from the two-dimensional stochastic process

$$dY_x = \begin{pmatrix} W''(x,s)dx + \sigma(x)dB_x^p \\ W''(x,s)dx + \sigma(x)dB_x^c \end{pmatrix}, \qquad (7)$$

indexed by strike x. The rest of the model proceeds along the same lines as the AD formulation. Conversely, the results obtained in the AD formulation can be translated to options (more generally, derivatives) by expressing options as portfolios of AD securities.

Options in Kyle (1985) In the options formulation, the discussion of Section 2.2 on specializing to the Kyle (1985) single asset setting translates as follows.

Proposition 2.1.

(i) If the asset payoff distributions only differ in expected values across signals and there are no options markets, then our model specializes to a static Kyle model where only futures are traded.

[25] As slight abuse of language, by a futures position we simply mean a linear position that pays off in the future, i.e. an AD portfolio $W(x)$ that is linear with respect to state x. We do not model margining, central clearing, etc.

(ii) Conversely, if the expected asset payoff is the same across all signals, then the futures market is redundant.

In Proposition 2.1(i), the assumption of no options markets is necessary. If the insider in Kyle (1985) observes the distribution and options markets are introduced, the informed demand for options would be non-zero. Options are not redundant even when agents are risk-neutral, and the possible asset payoff distributions differ only in the expected value. At any given state the payoff varies nonlinearly between signals, and the insider's optimal portfolio is therefore also nonlinear. This will be demonstrated as a special parametric case of our setting (see Example 6.1).

2.4 Basic Intuition

Consider the following two-asset two-signal toy case of our setting. This case has a simple intuition which, to the best of our knowledge, has not been explicitly set out in the literature. Discretize $[\underline{x}, \bar{x}]$ into a two-element partition $\chi_1 \sqcup \chi_2$, where sub-interval χ_i corresponds to state x_i. For $i, j = 1, 2$, let the state x_i security payoff conditional on signal s_j be specified by $p(i, j)$, i.e. $\eta(x, s_j) = p(1, j)1_{\chi_1}(x) + p(2, j)1_{\chi_2}(x)$. Suppose $p(1, 1) = p(2, 2) > p(1, 2) = p(2, 1)$. In other words, signal s_i informs the insider that x_i (resp. $x_{j \neq i}$) has high (resp. low) payoff.

The insider's equilibrium trading strategy is

$$W^*(x, s_1) = a^* 1_{\chi_1}(x) - a^* 1_{\chi_2}(x),\ W^*(x, s_2) = -a^* 1_{\chi_1}(x) + a^* 1_{\chi_2}(x) \tag{8}$$

for some endogenous constant $a^* > 0$ (which we shall obtain in the general setting). Conditional on his signal s_i, the insider buys the high payoff asset x_i and sells the low payoff asset $x_{j \neq i}$.

From the market maker's perspective, the equilibrium price impact within the state x_i market is (proportional to) the variance of x_i payoff across signals, $(p(i, 1), p(i, 2))$. The cross price impact between the x_1 and x_2 markets is (proportional to) the covariance of their payoffs across signals.[26] Since the payoffs of x_1 and x_2 are negatively correlated across signals, the cross price impact is negative. The variances and covariance are taken with respect to the market maker's equilibrium posterior on signals. The market maker's equilibrium posterior is biased towards the insider's private signal, because the market maker infers information from order flow.

The characterization of cross price impact extends the intuition of Kyle's lambda to between-markets. The covariance of asset payoffs determines the covariation of the insider's trades for the two assets, which then determines the informativeness of x_i order flow regarding $x_{j \neq i}$ payoff and, in turn, the cross price impact. In this case, the insider's trades (8) for x_1 and x_2 are negatively correlated

[26] The market maker's two-asset supply function, from $\mathbb{R}^2$ to $\mathbb{R}^2$, is nonlinear. Therefore, in contrast to the single asset setting, price impact within and across markets is not constant. The variance-covariance characterization of price impact holds at the margin.

across signals. Since the market maker correctly anticipates the insider's equilibrium strategy, he infers from a buy order for x_i that the insider's private signal is likely s_i, which means $x_{j\neq i}$ likely has low payoff. This leads the market maker to adjust the fair price of $x_{j\neq i}$ downwards. Similarly, a sell order for x_i leads to upward price adjustment for $x_{j\neq i}$.

A contribution of this paper is showing that the intuition described in this section is robust with respect to the number of assets, the specification of payoffs, and the number of signals. Indeed, the basic intuition articulated by our results in the general many-asset setting is that *such two-asset relationships operate pairwise throughout the entire set of assets*. In particular, the order flow at each market can have global price impact on all markets.

2.5 Methodology

The methodology we use to formally establish the above intuition in the general setting can be summarized as follows.

(i) In the market maker's Bayesian inference problem, the data ω (aggregate order flow across markets) is generated by a family of Itô processes parameterized by their drifts (conjectured insider orders $\widetilde{W}(\cdot, s)$ conditional on the signal s). The market maker's posterior conditional on his data ω is, by definition, an ω-by-ω object, and standard results for Itô processes do not admit such notions. To make the market maker's posterior well-defined and his infinite-dimensional inference problem tractable, we formulate an ω-by-ω stochastic integral.

(ii) Unlike the single asset setting, the insider's demand cannot be characterized by simply inverting the pricing rule of the market maker.[27] In particular, the insider optimizes over an infinite-dimensional function space—he may choose an arbitrary portfolio of state-contingent claims. The insider's first-order condition for his risk-sharing and price impact trade-offs must therefore be expressed in terms of functional derivatives.

(iii) Appropriately addressing (i) and (ii) makes it possible to deduce a finite-dimensional game that is isomorphic to the original infinite-dimensional game. This finite-dimensional reduction follows from the observation that, by a no-arbitrage argument, portfolios orthogonal to the possible payoff distributions can be excluded from the market maker's equilibrium belief. This reduction yields an equilibrium solution characterized by one endogenous constant.

Sections 3 and 4 place the preceding outline on a rigorous footing, clarify the economic mechanisms in effect, and derive necessary conditions for equilibrium.

[27] In the single asset equilibrium (5), $\beta = \frac{1}{2\lambda}$.

3 The Market Maker's Inference

As discussed in Section 2.4, intuitively the market maker's inference takes place not only within-market but also cross-market. In general, for any two states x and y, the order flow at state x may reveal information about state y payoff, which in turn leads to state y price adjustment. A rigorous formulation of the market maker's Bayesian inference problem will make this intuition precise (see Remark 3.2 below).

3.1 Bayes Rule Heuristic

Formally, the sample space in the market maker's Bayesian inference problem is the infinite-dimensional space $\Omega = C([\underline{x}, \bar{x}], \mathbb{R})$ of continuous functions on $[\underline{x}, \bar{x}]$. Suppose the market maker's belief regarding the insider's trading strategy is $\widetilde{W}(\,\cdot\,,\,\cdot\,)\colon [\underline{x}, \bar{x}] \times S \to \mathbb{R}$. According to this belief, conditional on the insider observing signal s the total order flow across all states received by the market maker is a sample path $\omega \in \Omega$ of the Itô process specified by (1)

$$dY_x^s = \widetilde{W}(x, s)dx + \sigma(x)dB_x. \tag{9}$$

Consider Bayes Rule,

$$\textit{posterior probability} \propto \textit{conditional likelihood of data} \times \textit{prior probability of parameter} \tag{10}$$

for the market maker's problem, where the data is order flow ω and the parameter $\widetilde{W}(\cdot, s)$ has prior probability $\pi_0(ds)$. On the space Ω of sample paths, consider the reference probability measure μ that specifies the stochastic process $(\sigma(x)B_x)$ and the probability measure $\mathbb{P}_{\widetilde{W}(\,\cdot\,,s)}$ that specifies (Y_x^s) of (9). The reference measure μ specifies the probability law of noise-only order flow when insider order is zero for all markets, while the measure $\mathbb{P}_{\widetilde{W}(\,\cdot\,,s)}$ specifies the probability law of order flow when insider order is $\widetilde{W}(\cdot, s)$. By Girsanov's Theorem, the Radon-Nikodym derivative of $\mathbb{P}_{\widetilde{W}(\,\cdot\,,s)}$ with respect to μ is given by[28]

$$\frac{d\mathbb{P}_{\widetilde{W}(\,\cdot\,,s)}}{d\mu} = e^{\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)}{\sigma(x)} dB_x - \frac{1}{2}\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)^2}{\sigma^2(x)} dx}. \tag{11}$$

At first glance, the Radon-Nikodym derivative (11) should specify the likelihood ratio of data ω conditional on the parameter $\widetilde{W}(\,\cdot\,, s)$ over the reference measure μ. Therefore, heuristic application of Bayes Rule (10) suggests that, conditional on order flow ω, the market maker's Bayesian posterior over

[28] See Theorem 5.1 in Chapter 3 of Karatzas and Shreve (2012).

the signal space S should be given by the expression

$$\frac{d\mathbb{P}_{\widetilde{W}(\,\cdot\,,s)}}{d\mu}(\omega) \times \pi_0(ds) \Big/ C(\omega, \widetilde{W}) \;=\; \underbrace{e^{\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)}{\sigma(x)} dB_x - \frac{1}{2}\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)^2}{\sigma^2(x)} dx}}_{\textit{conditional likelihood of data}} \times \underbrace{\pi_0(ds)}_{\textit{prior}} \Big/ C(\omega, \widetilde{W})\;, \tag{12}$$

where $C(\omega, \widetilde{W})$ is a normalization constant. This is problematic in the standard Itô construction of the $\int \cdot\, dB$ integral used in Girsanov's Theorem. In the Itô sense, the integral $\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)}{\sigma(x)} dB_x$—therefore the candidate expression (12) for the posterior—has no meaning for any given ω.[29] Unlike the case of finite-dimensional sample space $\mathbb{R}^n$, Bayesian posteriors are not trivially well-defined once the conditional probability law of data and the prior distribution of the parameter are specified.

3.2 The Posterior and Pricing Kernel

To make the candidate expression (12) rigorous, we consider a pathwise formulation of the stochastic integral, using tools adapted from rough paths theory.[30] Here we summarize the construction. The precise statement of the assumptions and proofs are given in Section A.1 of the Appendix.

Assume both the insider's contingent claim portfolio $W(\,\cdot\,, s)$ and the (squared) noise trading intensity $\sigma^2(\,\cdot\,)$ lie in the Hölder space $C^\delta([\underline{x}, \bar{x}], \mathbb{R})$ for some $0 < \delta < 1$, and restrict possible realization ω of total order flow to an appropriate subset Ω_γ of Ω.[31] Hölder δ-continuity is (much) weaker than differentiability and imposes no practical constraint on the insider's choice set. The modified sample space $\Omega_\gamma \subset \Omega$ is chosen so that ω lies in Ω_γ with probability one. Therefore restricting to $\omega \in \Omega_\gamma$ is without loss of generality. The restriction of the probability measure μ to Ω_γ will be denoted by μ_γ.

It can then be shown that, for each $\omega \in \Omega_\gamma$, the limit of Riemann sums

$$\int_{\underline{x}}^{\bar{x}} \frac{W_x}{\sigma^2(x)}\, d\omega_x \;\equiv\; \lim_{\substack{\max_k |x_{k+1}-x_k| \to 0 \\ \underline{x}=x_0<\cdots<x_n=\bar{x}}} \sum_{k=0}^{n-1} \frac{W_{x_k}}{\sigma^2(x_k)} \cdot (\omega_{x_{k+1}} - \omega_{x_k}) \qquad (\textit{pathwise integral}) \tag{13}$$

exists and therefore defines an ω-by-ω Riemann integral.[32] This ω-by-ω integral gives the Radon-Nikodym derivative $\frac{d\mathbb{P}_{\widetilde{W}(\,\cdot\,,s)}}{d\mu}$ of (11) rigorous meaning as an ω-by-ω density. Moreover, under this pathwise formulation of the stochastic integral, one can show that the data ω and the parameter

[29] An Itô integral is an equivalence class defined only up to indistinguishability. In the Itô sense, the Radon-Nikodym derivative $\frac{d\mathbb{P}_{\widetilde{W}(\,\cdot\,,s)}}{d\mu}$ of (11) defines a probability measure on Ω, but not an ω-by-ω density, which is necessary in order to define the conditional likelihood of data ω.

[30] See Friz and Hairer (2020).

[31] The Hölder space $C^\delta([\underline{x}, \bar{x}], \mathbb{R})$ consists of Hölder δ-continuous elements of $C([\underline{x}, \bar{x}], \mathbb{R})$, equipped with the Hölder seminorm $[f]_\delta = \sup_{x \neq y} \frac{|f(x)-f(y)|}{|x-y|^\delta}$.

[32] See Lemma A.1(i).

$\widetilde{W}(\,\cdot\,,s)$ in the market maker's intended Bayesian problem are jointly measurable.[33] Joint measurability makes conditioning on ω, therefore a Bayesian posterior, well-defined. Lastly, one can show that the candidate expression (12), now properly defined as random probability measure, specifies the market maker's intended posterior distribution on S.[34] We have the following theorem.

Theorem 3.1. *Under appropriate assumptions, the following holds.*[35] *Let* $\widetilde{W}(\,\cdot\,,\,\cdot\,)\colon [\underline{x},\bar{x}]\times S\to\mathbb{R}$ *be the market maker's belief regarding the insider's trading strategy, and* $\omega=(\omega_x)_{\underline{x}\le x\le\bar{x}}\in\Omega_\gamma$ *be the aggregate order flow received by the market maker.*

*(i) (**Market Maker Posterior**) The market maker's Bayesian inference problem is well-defined, and his posterior probability measure on* S *conditional on* ω *is given by*

$$\pi_1(ds,\omega;\widetilde{W}) = e^{\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)}{\sigma^2(x)}d\omega_x - \frac{1}{2}\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)^2}{\sigma^2(x)}dx} \cdot \pi_0(ds) \Big/ C(\omega,\widetilde{W})\,, \tag{14}$$

where the $\int\cdot\, d\omega$ *integral is the pathwise integral defined in* (13)*, and* $C(\omega,\widetilde{W})$ *is the normalization constant.*[36]

*(ii) (**Arrow-Debreu Prices/Pricing Kernel**) Conditional on receiving order flow* ω *and according to his belief* $\widetilde{W}(\,\cdot\,,\,\cdot\,)$*, the zero-profit AD prices set by the market maker are*

$$\underbrace{P(x,\omega\,;\widetilde{W})}_{\text{price of state } x \text{ AD security}} = \int_S \eta(x,s)\pi_1(ds,\omega\,;\widetilde{W}),\quad x\in[\underline{x},\bar{x}],$$

where $\pi_1(ds,\omega\,;\widetilde{W})$ *is the posterior obtained in (i).*

The posterior $\pi_1(ds,\omega;\widetilde{W})$ obtained in (14) is the heuristic expression (12) made rigorous. The assumptions of Theorem 3.1 will be maintained in all that follows.

Remark 3.2. *(Cross-Market Inference) Through pathwise integration, Theorem 3.1 makes precise the economic intuition of how the market maker arrives at his posterior and pricing kernel. After receiving total cumulative order flow* $\omega=(\omega_x)_{\underline{x}\le x\le\bar{x}}$ *across states, for each signal* s *the market maker compares* ω *with his belief* $\widetilde{W}(\,\cdot\,,s)=(\widetilde{W}(x,s))_{\underline{x}\le x\le\bar{x}}$ *regarding what the insider's orders across states would have been if* s *were the actual observed signal. If the observed order flow* ω *overlaps significantly with the market maker's belief* $\widetilde{W}(\,\cdot\,,s)$ *for a given signal* s*—i.e.* $\int_{\underline{x}}^{\bar{x}}\frac{\widetilde{W}(x,s)}{\sigma^2(x)}d\omega_x$ *is high, it reinforces the market maker's belief about the order flow he expects to see if* s *were the actual observed signal. This leads the market maker to revise his posterior probability for* s *upwards, and vice versa. The scaling by* $\frac{1}{\sigma^2(x)}$

[33] See Lemma A.1(ii). Joint measurability of data and parameter is a basic requirement in formulating any Bayesian problem, without which conditioning on data has no meaning, and the posterior cannot be defined.

[34] See Lemma A.1(iii)

[35] See Assumption A.1.

[36] $C(\omega,\widetilde{W})=\int_S e^{\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s')}{\sigma^2(x)}d\omega_x - \frac{1}{2}\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s')^2}{\sigma^2(x)}dx}\cdot\pi_0(ds')$.

adjusts for noise trading intensity. Through the overlap measures $(\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)}{\sigma^2(x)} d\omega_x)_{s\in S}$*, the market maker infers information globally about all states from the local order flow at any given state. The market maker then forms his pricing kernel* $P(\,\cdot\,,\omega\,;\widetilde{W})$ *by aggregating the across-states information implied by the order flow at each state.*

3.3 Overlap Measures

Definition 1. *Let the market maker's belief* $\widetilde{W}(\,\cdot\,,\,\cdot\,)\colon [\underline{x},\bar{x}]\times S \to \mathbb{R}$ *be given.*

(i) For the market maker, the (noise-adjusted) overlap between aggregate order ω *and his belief* $\widetilde{W}(\,\cdot\,,s)$ *regarding insider order conditional on* s *is*

$$\mathcal{O}_{mm}(\omega,s;\widetilde{W}) = \int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)}{\sigma^2(x)} d\omega_x.$$

(ii) For the insider, the (noise-adjusted) overlap between his order (or marginal order) $W(\,\cdot\,)$ *and* $\widetilde{W}(\,\cdot\,,s)$ *is*

$$\mathcal{O}_{insider}(W,s;\widetilde{W}) = \int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)}{\sigma^2(x)} W(x)dx.$$

$\mathcal{O}_{mm}(\omega,s;\widetilde{W})$ is the overlap between aggregate order ω observed by the market maker and what the market maker believes the insider's demand would be conditional on s, $\widetilde{W}(\,\cdot\,,s)$, adjusted for noise trading intensity. The overlap profile

$$\mathcal{O}_{mm}(\omega,\,\cdot\,;\widetilde{W}) = (\mathcal{O}_{mm}(\omega,s;\widetilde{W}))_{s\in S}$$

compares ω with market maker belief $\widetilde{W}(\,\cdot\,,s)$ for each signal s and summarizes the price impact of ω. Similarly, $\mathcal{O}_{insider}(W,s;\widetilde{W})$ is the overlap between actual insider order $W(\,\cdot\,)$ and what the market maker believes the insider's order would be, $\widetilde{W}(\,\cdot\,,s)$. The overlap profile

$$\mathcal{O}_{insider}(W,\,\cdot\,;\widetilde{W}) = (\mathcal{O}_{insider}(W,s;\widetilde{W}))_{s\in S}$$

compares W with the market maker's belief signal-wise and summarizes the extent to which an insider order $W(\,\cdot\,)$, i.e. ω net of noise trades, affects prices. As discussed in Remark 3.2, $\mathcal{O}_{mm}$ occurs naturally in the market maker's problem. $\mathcal{O}_{insider}$ will play an analogous role for the insider's problem (see Sections 4.2 and 4.3).

The difference between the two overlap measures $\mathcal{O}_{mm}(\omega,s;\widetilde{W})$ and $\mathcal{O}_{insider}(W,s;\widetilde{W})$ is that, while the market maker only observes the aggregate order ω, the insider knows his own order $W(\,\cdot\,)$. The market maker updates the posterior probability of s conditional on ω by considering $\mathcal{O}_{mm}(\omega,s;\widetilde{W})$. On the other hand, the insider's portfolio choice $W(\,\cdot\,)$ influences the market maker's posterior probability of s through $\mathcal{O}_{insider}(W, s; \widetilde{W})$. Both overlap measures are non-local with respect to the state. In equilibrium, this will result in cross-market price impact.

4 The Insider's Portfolio Choice

This section derives two necessary conditions for the insider's optimal portfolio—the first-order condition and no-arbitrage. Under the same maintained assumptions as Theorem 3.1, given the market maker's belief $\widetilde{W}(\,\cdot\,,\,\cdot\,)$ regarding his trading strategy, the formal statement of the insider's portfolio choice problem (2) conditional on observing signal $s \in S$ is[37]

$$\max_{W(\,\cdot\,)\in C^{\delta}([\underline{x},\bar{x}],\mathbb{R})} \underbrace{\int_{\underline{x}}^{\bar{x}} W(x)\eta(x,s)dx}_{\text{expected payoff}} - \underbrace{\int_{\underline{x}}^{\bar{x}} W(x)\cdot \mathbb{E}^{\mathbb{P}_W}[P(x,\omega;\widetilde{W})]dx}_{\text{expected cost}} \tag{15}$$

$$= \max_{W(\,\cdot\,)\in C^{\delta}([\underline{x},\bar{x}],\mathbb{R})} \int_{\underline{x}}^{\bar{x}} W(x)\eta(x,s)dx - \int_{\underline{x}}^{\bar{x}} W(x)\cdot \overline{P}(x,W;\widetilde{W})dx \tag{16}$$

$$\equiv \max_{W(\cdot)\in C^{\delta}([\underline{x},\bar{x}],\mathbb{R})} J(W;\widetilde{W},s) \tag{17}$$

where $\mathbb{P}_W$ is the probability measure over aggregate order flow $\omega \in \Omega_\gamma$ induced by the insider's choice $W(\,\cdot\,)$, $P(\,\cdot\,,\omega;\widetilde{W})$ is the market maker's pricing kernel, and

$$\overline{P}(x,W;\widetilde{W}) = \mathbb{E}^{\mathbb{P}_W}[P(x,\omega;\widetilde{W})] \tag{18}$$

is the expected price of state x security under $\mathbb{P}_W$ and given $\widetilde{W}$. The expression $J(W;\widetilde{W},s)$ in (17) defines the insider's expected utility functional conditional on s and given market maker belief $\widetilde{W}$.

Clearly, any portfolio W with the property that $\int_{\underline{x}}^{\bar{x}} W(x)\eta(x,s)dx = 0$ for all s achieves zero expected payoff for the insider. In other words, any portfolio W that is orthogonal to the observed payoff distribution $\eta(\,\cdot\,,s)$ has zero expected payoff. Therefore the domain of J can be restricted to (the closure of) the linear span of $\{\eta(\,\cdot\,,s)\}_{s\in S}$. Also, any portfolio W that is constant across states achieves zero expected utility. In other words, risk-free portfolios are payoff irrelevant, because there is no information asymmetry regarding the risk-free asset. Therefore, by adding a risk-free portfolio if necessary, we can assume the insider chooses a long-short portfolio W with $\int_{\underline{x}}^{\bar{x}} W(x)dx = 0$. We record these facts as Proposition 4.1.

Proposition 4.1.

(i) (No Insider Demand Orthogonal to Payoffs) Without loss of generality, for all $\widetilde{W}$ and s, the

[37] Under the maintained Assumption A.1, the expression of the expected utility functional in (2) is equal to that in (15) by the Fubini-Tonelli Theorem, which allows the two integrals $\int_{\underline{x}}^{\bar{x}} \cdot\, dx$ and $\mathbb{E}^{\mathbb{P}_W}[\,\cdot\,]$ to commute.

domain of the insider's expected utility functional $J(\,\cdot\,;\widetilde{W},s)$ can be restricted to the closure of the linear span of $\{\eta(\,\cdot\,,s)\}_{s\in S}$ in $C^\delta([\underline{x},\bar{x}],\mathbb{R})$.

(ii) (Long-Short Portfolios Only) Without loss of generality, the domain of $J(\,\cdot\,;\widetilde{W},s)$ can be further restricted to portfolio W's with $\int_{\underline{x}}^{\bar{x}} W(x)dx = 0$.

4.1 First-Order Condition

As the insider trades to exploit his private information regarding differential payoffs across states conditional on his signal, his trades incur a marginal cost by revealing his private information to the market maker.[38] Cross-market inference by the market maker means this cost operates not only through the within-market channel but also between-markets. In general, at any given state, the insider trades off the marginal payoff of the state against the marginal cost consisting of, first, the current price of the state itself, conditional on no price impact, and second, the global price impact of marginal demand at the given state. This economic intuition is made rigorous by Theorem 4.2.

Theorem 4.2. *(Insider FOC)*

Conditional on observing s and given market maker belief $\widetilde{W}(\,\cdot\,,\,\cdot\,)$, the insider's expected utility functional $J(\,\cdot\,;\widetilde{W},s)\colon C^\delta([\underline{x},\bar{x}],\mathbb{R}) \to \mathbb{R}$ is Gâteaux-differentiable. Its Gâteaux derivative at W, i.e. marginal utility functional, $dJ(W;\widetilde{W},s)(\,\cdot\,)\colon C^\delta([\underline{x},\bar{x}],\mathbb{R}) \to \mathbb{R}$ decomposes into the following sum

$$\underbrace{dJ(W;\widetilde{W},s)(\,\cdot\,)}_{\textit{marginal utility}} = \underbrace{N_p(W)(\,\cdot\,)}_{\textit{marginal payoff}} - \underbrace{\big(\,N_{AD}(W)(\,\cdot\,) + N_K(W)(\,\cdot\,)\,\big)}_{\textit{marginal cost}}$$

where

$$\underbrace{N_p(W)(v)}_{\textit{marginal payoff of } v} = \int_{\underline{x}}^{\bar{x}} v(x)\eta(x,s)ds, \tag{19}$$

$$\underbrace{N_{AD}(W)(v)}_{\substack{\textit{marginal cost of } v,\\ \textit{AD term}}} = \int_{\underline{x}}^{\bar{x}} \overline{P}(x,W;\widetilde{W})v(x)dx \tag{20}$$

$$\underbrace{N_K(W)(v)}_{\substack{\textit{marginal cost of } v,\\ \textit{price impact term}}} = \int_{\underline{x}}^{\bar{x}} W(x)\mathbb{E}^{\mu_\gamma}\left[\mathbb{Cov}^{\pi_1(ds,\omega;\widetilde{W})}\Big(\eta(x,s),\mathcal{O}_{insider}(v,s;\widetilde{W})\Big)\right]dx. \tag{21}$$

[38] A risk-averse insider would, in addition to exploiting his private information regarding differential payoffs, trade to smooth consumption across states. See Section 7.3.

Therefore, for any optimal portfolio W conditional on signal s, the ***first-order condition***

$$\underbrace{\underbrace{N_p(W)(\,\cdot\,)}_{\textit{marginal payoff}} = \underbrace{\underbrace{N_{AD}(W)(\,\cdot\,)}_{\textit{AD term}} + \underbrace{N_K(W)(\,\cdot\,)}_{\textit{price impact term}}}_{\textit{marginal cost}}} \tag{22}$$

must hold.[39]

The Gâteaux derivative $dJ(W;\widetilde{W},s)(v)$ evaluated at $v \in C^{\delta}([\underline{x},\bar{x}],\mathbb{R})$ is the insider's marginal utility of marginal demand v at portfolio W, conditional on observing s and given market maker belief $\widetilde{W}(\,\cdot\,,\,\cdot\,)$. In the price impact term $N_K(W)$ of (21), the covariance quantity

$$\mathbb{Cov}^{\pi_1(ds,\omega;\widetilde{W})}\Big(\eta(x,s), \mathcal{O}_{insider}(v,s;\widetilde{W})\Big), \tag{23}$$

for a given state x, is the covariance between the state x payoff profile $\eta(x,\,\cdot\,)$ and the overlap measure profile $\mathcal{O}_{insider}(v,\,\cdot\,;\widetilde{W})$ across signals, under the market maker's posterior $\pi_1(ds,\omega;\widetilde{W})$ on S.[40] In the expression

$$\mathbb{E}^{\mu_\gamma}\Big[\mathbb{Cov}^{\pi_1(ds,\omega;\widetilde{W})}(\cdots,\cdots)\Big]$$

entering $N_K(W)$, the inner covariance operator $\mathbb{Cov}^{\pi_1(ds,\omega;\widetilde{W})}(\,\cdot\,,\,\cdot\,)$ is taken with respect to market maker's posterior $\pi_1(ds,\omega;\widetilde{W})$ conditional on ω, and the outer expectation operator $\mathbb{E}^{\mu_\gamma}[\,\cdot\,]$ is taken with respect to μ_γ over ω.

The first-order condition (22) equalizes the marginal payoff functional $N_p(W)$ with the marginal cost functional $N_{AD}(W) + N_K(W)$. The marginal payoff functional $N_p(W)$ can be identified with the payoff distribution $\eta(\,\cdot\,,s)$. That is, conditional on signal s, the payoff of marginal demand for state x security is the true payoff $\eta(x,s)$ of that security, which is known to the insider. The marginal cost functional decomposes into the risk-sharing *Arrow-Debreu term* $N_{AD}(W)$ and the *price impact term* $N_K(W)$. The Arrow-Debreu term $N_{AD}(W)$ can be identified with the current (expected) AD prices $\overline{P}(\,\cdot\,,W;\widetilde{W})$. In other words, the marginal cost for state x order would be the current state x price set by the market maker, if marginal demand at x does not cause the market maker to revise his belief $\widetilde{W}$. However, in general a marginal demand at x will cause the market maker to revise his belief and result in global price impact on all markets. This corresponds to the price impact term $N_K(W)$ of the marginal cost.

[39] The parametrization by $\widetilde{W}$ and s is suppressed in the notations for the Arrow-Debreu term N_{AD} and the price impact term N_K.

[40] Let us further explain the notation of the covariance quantity (23). The variable s in (23) is a dummy variable of integration and integrated away over the distribution $\pi_1(ds,\omega;\widetilde{W})$ after taking the covariance. Given $\widetilde{W}$, the covariance quantity is therefore a function of x, v, and ω.

4.2 Cross-Market Price Impact

By the characterization of the price impact term $N_K(W)$ of the marginal cost functional in Theorem 4.2, the global price impact of insider marginal demand v is given by the following linear operator $\Lambda\colon C^\delta([\underline{x},\bar{x}],\mathbb{R}) \to C^\delta([\underline{x},\bar{x}],\mathbb{R})$,

$$v(\,\cdot\,) \overset{\Lambda}{\mapsto} \mathbb{E}^{\mu_\gamma}\left[\mathbb{Cov}^{\pi_1(ds,\omega;\widetilde{W})}\left(\eta(\,\cdot\,,s), \mathcal{O}_{insider}(v,s;\widetilde{W})\right)\right] \qquad \textit{(global price impact operator)}. \tag{24}$$

Λ maps a marginal demand v of the insider to its price impact on each state. Λv evaluated at x, $(\Lambda v)(x)$, is the price impact of v on state x market. In the definition of Λ, for a given state x, the covariance quantity (23) is the covariance, under the market maker's posterior $\pi_1(ds,\omega;\widetilde{W})$, of two profiles across signals:

- The payoff profile of state x security, $\eta(x,\,\cdot\,)$.
- The overlap profile $\mathcal{O}_{insider}(v,\,\cdot\,;\widetilde{W})$ of Definition 1(ii), between marginal demand v and what the market maker believes the insider's orders would be if a signal were the actual observed signal.

The operator Λ therefore captures the following intuition. If the marginal demand v interacts with the market maker belief $\widetilde{W}$ signal-by-signal to produce an overlap profile $\mathcal{O}_{insider}(v,\,\cdot\,;\widetilde{W})$ that resembles the state x payoff $\eta(x,\,\cdot\,)$, then v causes the market maker to raise state x price, vice versa.

Of particular interest is the special case of *localized pairwise price impact*—the marginal effect of state y order on state x price. We have the following immediate corollary of Theorem 4.2.

Corollary 4.3. *Let $x,y \in [\underline{x},\bar{x}]$ be two states, Λ be the global price impact operator defined in* (24)*, and $v(\,\cdot\,)$ be a marginal demand that has support in a vanishingly small neighborhood of state y. Then the price impact of state y order on (expected) state x price under market maker belief $\widetilde{W}$ is given by*

$$\frac{\partial}{\partial W(y)}\overline{P}(x,W;\widetilde{W}) = (\Lambda v)(x) \qquad \textit{(price impact of } y \textit{ on } x\textit{)}. \tag{25}$$

Example 4.4. *The marginal demand v in Corollary 4.3 can be taken as the Dirac delta function δ_y at state y, so that*[41]

$$\mathcal{O}_{insider}(v,\,\cdot\,;\widetilde{W}) = \left(\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s')}{\sigma^2(x)}\delta_y(x)dx\right)_{s'\in S} = \left(\frac{\widetilde{W}(y,s')}{\sigma^2(y)}\right)_{s'\in S}.$$

Suppose y can only realize under signal s and never realize under signals $s' \neq s$, and the market maker believes that $\widetilde{W}(y,s) > 0$ and $\widetilde{W}(y,s') = 0$ for $s' \neq s$, i.e. insider demand at y is positive if the true signal is s and zero otherwise. Such belief by the market maker is plausible. This gives an overlap

[41] In the options formulation, δ_y is familiar to option traders as an (infinitesimally) tight butterfly around strike y.

profile $\mathcal{O}_{insider}(v,\,\cdot\,;\widetilde{W})$ *that concentrates at* s*—positive at* s *and zero for* $s' \neq s$*. Now consider the following two scenarios for a state* x*:*

- *Scenario (i). Suppose* x *can only realize under signal is* $s' \neq s$*, i.e.* $\eta(x,s) = 0$ *and* $\eta(x,s') > 0$ *for* $s' \neq s$*. Then state* y *order has negative price impact on state* x *market, because the overlap profile* $\mathcal{O}_{insider}(v,\,\cdot\,;\widetilde{W})$ *and payoff profile* $\eta(x,\,\cdot\,)$ *have negative correlation. The overlap profile* $\mathcal{O}_{insider}(v,\,\cdot\,;\widetilde{W})$ *concentrates at* s *while the payoff profile* $\eta(x,\,\cdot\,)$ *concentrates on the complement* $\{s' \colon s' \neq s\}$*. The negativity of price impact of state* y *order on state* x *price has a familiar intuition. An insider buy order at* y *will likely lead the market maker to believe that* x *has low probability of realizing, which lowers the state* x *price.*
- *Scenario (ii). Suppose* x*, like* y*, can only realize under* s*, i.e.* $\eta(x,s) > 0$ *and* $\eta(x,s') = 0$ *for* $s' \neq s$*. Then state* y *order has positive price impact on state* x *market, because the overlap profile* $\mathcal{O}_{insider}(v,\,\cdot\,;\widetilde{W})$ *and the payoff profile* $\eta(x,\,\cdot\,)$ *have positive correlation, both concentrating at* s*. The intuition is the flip side of Scenario (i).*

Example 6.5(i) considers Scenario (i) in equilibrium, when the market maker holds equilibrium belief.

4.3 No-Arbitrage

No-arbitrage is a general necessary condition for equilibrium. We now derive this condition for our setting, which is also the basis of the equilibrium solution.[42] Here, as in Section 4.2, the overlap measure $\mathcal{O}_{insider}(W,s;\widetilde{W})$ between the insider's trades and the market maker's belief naturally occurs, but with W being an insider portfolio instead of a marginal demand. Consider the bounded linear operator $K(\,\cdot\,;\widetilde{W})\colon C^{\delta}([\underline{x},\bar{x}],\mathbb{R}) \to C(S,\mathbb{R})$ defined by[43]

$$K(W;\widetilde{W}) = \mathcal{O}_{insider}(W,\,\cdot\,;\widetilde{W}) = \Big(\mathcal{O}_{insider}(W,s;\widetilde{W})\Big)_{s\in S}. \tag{26}$$

Definition 2. *The* ***zero price impact subspace*** $\mathcal{V}_0(\widetilde{W})$ ***under market maker belief*** $\widetilde{W}$ *is the null space of* $K(\,\cdot\,;\widetilde{W})$*, i.e.*

$$\mathcal{V}_0(\widetilde{W}) = K^{-1}(0;\widetilde{W}) \subset C^{\delta}([\underline{x},\bar{x}],\mathbb{R}).$$

Theorem 4.5. *(No-Arbitrage)*

For the insider's portfolio choice problem (15) *conditional on observing signal* s *and given market maker belief* $\widetilde{W}$*, an optimal portfolio exists only if his expected utility functional* $J(\,\cdot\,;\widetilde{W},s)$ *descends to the quotient space* $C^{\delta}([\underline{x},\bar{x}],\mathbb{R})\big/\mathcal{V}_0(\widetilde{W})$*. In other words, for all* $c \in C(S,\mathbb{R})$*, the insider's expected*

[42] In our setting, we interpret "no-arbitrage" to mean that the insider cannot achieve unbounded expected utility.

[43] That the range of $K(\,\cdot\,;\widetilde{W})$ lies in $C(S,\mathbb{R})$ and that $K(\,\cdot\,;\widetilde{W})$ is a bounded operator both follow from the Dominated Convergence Theorem.

utility functional $J(\,\cdot\,;\widetilde{W},s)$ *must be constant on the closed affine subspace* $\mathcal{V}_c(\widetilde{W}) = K^{-1}(c;\widetilde{W})$. *In particular,* $J(\,\cdot\,;\widetilde{W},s)$ *must be zero on the zero price impact subspace* $\mathcal{V}_0(\widetilde{W})$.

The proof of the theorem illustrates the underlying economic intuition.

Proof. Inspecting the expression (18) of $\overline{P}(x,W;\widetilde{W})$ shows that, if $K(W;\widetilde{W}) = c$ for a fixed $c \in C(S,\mathbb{R})$, then the (expected) state x price $\overline{P}(x,W;\widetilde{W})$ faced by the insider depends only on the state x and not on W.[44] That is, conditional on $c \in C(S,\mathbb{R})$, an insider order $W \in \mathcal{V}_c(\widetilde{W})$ does not change the (expected) state prices.

In particular, an insider portfolio W in the zero price impact subspace $\mathcal{V}_0(\widetilde{W})$ has an overlap profile

$$\mathcal{O}_{insider}(W,\,\cdot\,;\widetilde{W}) = 0 \in C(S,\mathbb{R})$$

that is identically zero on S. That is, a portfolio $W \in \mathcal{V}_0(\widetilde{W})$ causes zero price impact, because it is orthogonal to what the market maker believes the insider's demand would be regardless of s. For the insider's problem to be well-posed, such a portfolio must give zero expected utility. Otherwise, the insider can arbitrage by arbitrarily scaling up his portfolio. If $W \in \mathcal{V}_0(\widetilde{W})$ and $J(W;\widetilde{W},s) > 0$, then the insider's expected utility $J(\alpha W;\widetilde{W},s) \to \infty$ as $\alpha \to \infty$, because the scaled portfolio $\alpha W \in \mathcal{V}_0(\widetilde{W})$ causes no price impact as $\alpha \to \infty$. This cannot occur in equilibrium. Similarly, $J(W;\widetilde{W},s) < 0$ for some $W \in \mathcal{V}_0(\widetilde{W})$ would allow arbitrage.

More generally, for any two portfolios W_1 and W_2 such that $K(W_1;\widetilde{W}) = K(W_2;\widetilde{W}) = c \in C(S,\mathbb{R})$, i.e. $W_1, W_2 \in \mathcal{V}_c(\widetilde{W})$, no-arbitrage requires that $J(W_1;\widetilde{W},s) = J(W_2;\widetilde{W},s)$. Otherwise, a long-short portfolio, $W_1 - W_2$ or $W_2 - W_1 \in \mathcal{V}_0(\widetilde{W})$, would be an arbitrage opportunity. In other words, to preclude arbitrage, the insider's expected utility functional $J(\,\cdot\,;\widetilde{W},s)$ must be constant on the affine subspace $\mathcal{V}_c(\widetilde{W})$ for each $c \in C(S,\mathbb{R})$. This proves the theorem. $\square$

Restriction on Equilibrium Belief Since no-arbitrage is a necessary condition for equilibrium, Theorem 4.5 gives a restriction on the market maker's equilibrium belief. In equilibrium, the market maker would not believe there is any informed demand in the zero price impact subspace $\mathcal{V}_0(\widetilde{W})$.

[44] Let

$$\mathcal{I}(\omega,s;\widetilde{W}) = \int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)}{\sigma^2(x)} d\omega_x - \frac{1}{2}\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)^2}{\sigma^2(x)} dx$$

and $C'(\omega,W;\widetilde{W}) = \int_S e^{\mathcal{I}(\omega,s';\widetilde{W})} e^{\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x',s')W(x')}{\sigma^2(x')}dx'} \pi_0(ds')$. We have

$$\begin{aligned}\overline{P}(x,W;\widetilde{W}) &= \mathbb{E}^{\mu_\gamma}\Big[\int_S \eta(x,s) e^{\mathcal{I}(\omega,s;\widetilde{W})} e^{\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x',s)W(x')}{\sigma^2(x')}dx'} \pi_0(ds) / C'(\omega;\widetilde{W})\Big] \\ &= \mathbb{E}^{\mu_\gamma}\Big[\int_S \eta(x,s) e^{\mathcal{I}(\omega,s;\widetilde{W})+c(s)} \pi_0(ds) / C'(\omega;\widetilde{W})\Big]\end{aligned}$$

which does not depend on W.

Intuitively, we can therefore exclude the zero price impact subspace $\mathcal{V}_0(\widetilde{W})$ from the market maker's equilibrium belief and restrict the belief to the "orthogonal complement" of $\mathcal{V}_0(\widetilde{W})$—more precisely, the quotient space $C^\delta([\underline{x},\bar{x}],\mathbb{R}) \big/ \mathcal{V}_0(\widetilde{W})$ in Theorem 4.5. In particular, when S is finite, $K(\,\cdot\,;\widetilde{W})$ is a finite-rank operator, as the overlap profile $\mathcal{O}_{insider}(W,\,\cdot\,;\widetilde{W})$ is finite-dimensional in this case. This means the orthogonal complement of $\mathcal{V}_0(\widetilde{W})$ is also finite-dimensional.[45] This leads to a finite-dimensional reduction of the infinite-dimensional trading game that we will exploit in Section 5.

Remark 4.6. *More generally, when S is a compact subset of Euclidean space, $K(\,\cdot\,;\widetilde{W})$ is a compact linear operator.*[46] *The restriction on the market maker's possible equilibrium belief $\widetilde{W}$ given by Theorem 4.5 remains in effect. The null space of a compact operator—the zero price impact subspace $\mathcal{V}_0(\widetilde{W})$—is "large." This restricts the market maker's possible equilibrium belief to a "small" subspace that is "almost finite-dimensional." Our methodology for equilibrium solution remains applicable in the more general case of compact S, such as a closed bounded subset of $\mathbb{R}$, with no material change in the basic economic results. The degree of generality under which equilibrium obtains speaks to the robustness of the underlying economic intuition. To reduce the technicality of the exposition, we consider the finite S case in what follows.*

5 Equilibrium

In this section, we establish equilibrium for the case of finitely many signals.

Assumption 1. *(Finite S Case)*

(i) The signal space $S = \{s_i\}_{i=1,\cdots,I}$ is finite. The probability distribution over states $[\underline{x},\bar{x}]$ conditional on s_i is $\eta(\,\cdot\,,s_i)$ for $i = 1,\cdots,I$.

(ii) The prior π_0 is the uniform distribution on S, i.e. $\pi_0(s_i) = \frac{1}{I}$ for each i.

(iii) Noise trading intensity is constant $\sigma > 0$ across states, i.e. $\sigma(x) = \sigma > 0$ for all $x \in [\underline{x},\bar{x}]$.

(iv) The set of distributions $\{\eta(\,\cdot\,,s_i)\}_{i=1,\cdots,I}$ is linearly independent.

To isolate the effect of information asymmetry regarding the payoff distribution, Assumption 1(ii) and (iii) control for the Bayesian prior and noise trading intensity, respectively.[47] Assumption 1(iv) is an identification condition, from the perspective of the market maker.

[45] This follows from a general Fundamental Theorem of Linear Algebra.

[46] A linear operator K on a topological vector space is compact if the image of a bounded set under K is precompact. Finite-rank operators are compact.

[47] In fact, controlling for noise trading intensity is redundant in the case of finitely many signals. When S is finite, equilibrium is possible only when noise trading intensity $\sigma^2(x)$ is constant across states $x \in [\underline{x},\bar{x}]$. In other words, when there are only finitely many signals, varying noise trading intensity across infinitely many markets would allow arbitrage, and no equilibrium can exist.

Definition 3. *A **(Bayesian Nash) equilibrium** in our model is a map* $W^*(\cdot,\cdot):[\underline{x},\bar{x}]\times S\to\mathbb{R}$ *such that, for all* $s\in S$,

$$W^*(\cdot,s)\in\underset{W(\cdot)\in C^{\delta}([\underline{x},\bar{x}],\mathbb{R})}{\operatorname{argmax}} J(W;W^*,s),$$

where $J(\cdot;W^*,s)$ *is the insider's expected utility functional* (17)*, conditional on signal* s *and given market maker belief* $W^*(\cdot,\cdot)$.

In other words, in equilibrium the market maker holds belief $W^*(\cdot,\cdot)$, and the insider's optimal portfolio conditional on each s is $W^*(\cdot,s)$, thereby confirming the market maker's belief.

Definition 4. *The **information intensity matrix*** $\mathbf{M}(\eta)\in\mathbb{R}^{I\times I}$ *is*

$$\mathbf{M}(\eta)=[\mathbf{M}_{ij}(\eta)]_{1\le i,j,\le I}=[\int_{\underline{x}}^{\bar{x}}\eta(x,s_i)\eta(x,s_j)dx]_{1\le i,j,\le I}.$$

The information intensity matrix $\mathbf{M}(\eta)$ is positive definite under the identification condition of Assumption 1(iv). It has a clear interpretation. It is a measure of the correlation of asset payoff distributions across signals. Partially ordered by positive-semidefiniteness, a larger $\mathbf{M}(\eta)$ means a higher correlation of asset payoff distributions $\eta(\cdot,s)$ across signals s. There is a direct analogy between $\mathbf{M}(\eta)$ and the variance of asset fundamental value in the Kyle (1985) single asset setting of Section 2.2. They both reflect the informativeness—*information intensity*—of the insider's signal. A change of basis via $\mathbf{M}(\eta)$ will give a finite-dimensional reduction of the trading game that is invariant with respect to the specification of $\{\eta(\cdot,s_i)\}_{i=1,\cdots,I}$ (see Remark 5.2).

Corollary 5.1 below follows from Proposition 4.1(i) and strengthens the no-arbitrage condition of Theorem 4.5 under equilibrium.

Corollary 5.1. *Under Assumption 1 (i), for all market maker belief* $\widetilde{W}(\cdot,\cdot)$ *and signal* s_j*, the domain of insider's expected utility functional* $J(\cdot;\widetilde{W},s_j)$ *can be restricted to the contingent claim portfolios in the linear span of* $\{\eta(\cdot,s_i)\}_{i=1,\cdots,I}$*. Therefore, in equilibrium* $W^*(\cdot,s_j)$ *must also lie in the linear span of* $\{\eta(\cdot,s_i)\}_{i=1,\cdots,I}$*, for all* j.

Corollary 5.1 is very intuitive. By Proposition 4.1(i), the portfolios orthogonal to the linear span of $\{\eta(\cdot,s_i)\}_{i=1,\cdots,I}$ give zero payoff to the insider. Therefore, by the no-arbitrage condition of Theorem 4.5, the market maker's equilibrium belief can be restricted to the same linear span. In an equilibrium W^*, the orthogonal complement of the linear span is the zero price impact subspace $\mathcal{V}_0(W^*)$. The finite-dimensional reduction of the infinite-dimensional game is achieved by excluding, without loss of generality, the "large" subspace $\mathcal{V}_0(W^*)$ so that only a "small" residual subspace remains.[48]

[48]More precisely, the excluded zero price impact subspace $\mathcal{V}_0(W^*)$ has finite codimension in the infinite-dimensional space $C^{\delta}([\underline{x},\bar{x}],\mathbb{R})$. The quotient space $C^{\delta}([\underline{x},\bar{x}],\mathbb{R})\big/\mathcal{V}_0(W^*)$ in Theorem 4.5 is isomorphic to the linear span of $\{\eta(\cdot,s_i)\}_{i=1,\cdots,I}$.

5.1 Finite-Dimensional Reduction

In view of Corollary 5.1, the Bayesian trading game between the insider and the market maker admits a finite-dimensional reduction. We now make this explicit by considering the linear expansion coefficients with respect to $\{\eta(\,\cdot\,, s_i)\}_{i=1,\cdots,I}$. Characterizing equilibrium is then equivalent to characterizing the fixed points of a map on $\mathbb{R}^{I\times I}$.

For a portfolio

$$W(\,\cdot\,) = \sum_{i=1}^{I} d_i \eta(\,\cdot\,, s_i)$$

in the linear span of $\{\eta(\,\cdot\,, s_i)\}_{i=1,\cdots,I}$, collect the expansion coefficients in the vector

$$d = (d_i) \in \mathbb{R}^I. \tag{27}$$

Similarly, for market maker belief $\widetilde{W}(\,\cdot\,,\,\cdot\,)$ in the same linear span,

$$\widetilde{W}(\,\cdot\,, s_i) = \sum_{j=1}^{I} \tilde{d}_j^{(i)} \eta(\,\cdot\,, s_j), \text{ where } \tilde{d}^{(1)}, \cdots, \tilde{d}^{(I)} \in \mathbb{R}^I\ ,$$

form the square matrix

$$\widetilde{D} = [\tilde{d}^{(1)}\ \cdots\ \tilde{d}^{(I)}] \in \mathbb{R}^{I\times I},$$

where $\tilde{d}^{(i)}$ is the i-th column of $\widetilde{D}$. Furthermore, impose the change of basis on $\mathbb{R}^I$ given by[49]

$$\hat{d} = \frac{1}{\sigma}\mathbf{M}^{\frac{1}{2}}(\eta) d. \tag{28}$$

Let $\widehat{D}$ denote the market maker's belief $\widetilde{D}$ of (29) under this change of basis transformation, i.e.

$$\widehat{D} = \frac{1}{\sigma}\mathbf{M}^{\frac{1}{2}}(\eta)\widetilde{D} = [\hat{d}^{(1)} \cdots \hat{d}^{(I)}]. \tag{29}$$

Let e_i denote the i-th standard basis vector of $\mathbb{R}^I$ and $\mathbf{I} \in \mathbb{R}^{I\times I}$ be the identity matrix.

Remark 5.2. *The change of basis* (28) *normalizes the information intensity across signals. For example, if the payoffs* $\{\eta(\,\cdot\,, s_i)\}_{i=1,\cdots,I}$ *have disjoint supports on* $[\underline{x}, \bar{x}]$*, then the change of basis is equivalent to simply replacing each* $\eta(\,\cdot\,, s_i)$ *by*

$$\sigma \cdot \frac{1}{(\int_{\underline{x}}^{\bar{x}} \eta(x, s_i)^2 dx)^{\frac{1}{2}}} \cdot \eta(\,\cdot\,, s_i)$$

[49] $\mathbf{M}^{\frac{1}{2}}(\eta)$ denotes the unique positive-definite square root of the information intensity matrix $\mathbf{M}(\eta)$, which exists by Assumption 1(iv).

in the original game, which normalizes the across-state variations of the payoffs.[50]

Proposition 5.3. *Under Assumption 1, the Bayesian trading game between the insider and the market maker is isomorphic to the following finite-dimensional* ***canonical game****:*

- *The market maker has uniform prior on signals $s_1, \cdots, s_I$. The insider observes the realization of the signal.*
- *If the insider chooses $\hat{d} \in \mathbb{R}^I$, the market maker observes $\hat{d} + \widehat{X}$, where $\widehat{X} = (\widehat{X}_i)_{i=1,\cdots,I}$, $\widehat{X}_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$.*
- *After observing $\hat{d} + \widehat{X}$ and given his belief $\widehat{D} = [\hat{d}^{(1)} \cdots \hat{d}^{(I)}]$ regarding the insider's strategy, the market maker updates his prior to the posterior probability mass function $(\hat{\pi}_1(i, \hat{d}; \widehat{D}))_{i=1,\cdots,I}$ on S specified by*

$$\hat{\pi}_1(i, \hat{d}, \widehat{X}; \widehat{D}) = \frac{e^{(\hat{d}^{(i)})^T \hat{d} + (\hat{d}^{(i)})^T \widehat{X} - \frac{1}{2}(\hat{d}^{(i)})^T \hat{d}^{(i)}}}{\sum_{j=1}^{I} e^{(\hat{d}^{(j)})^T d + (\hat{d}^{(j)})^T \widehat{X} - \frac{1}{2}(\hat{d}^{(j)})^T \hat{d}^{(j)}}}, \quad i = 1, \cdots, I. \tag{30}$$

- *Conditional on observing s_i, the insider's maximization problem is*

$$\max_{\hat{d} \in \mathbb{R}^I} e_i^T \hat{d} - \hat{\pi}_1(\hat{d}; \widehat{D})^T \hat{d} \equiv \max_{\hat{d} \in \mathbb{R}^I} J(\hat{d}; \widehat{D}, i), \tag{31}$$

where $\hat{\pi}_1(\hat{d}; \widehat{D})$ is the expectation of

$$\left(\hat{\pi}_1(1, \hat{d}, \widehat{X}; \widehat{D}), \cdots, \hat{\pi}_1(I, \hat{d}, \widehat{X}; \widehat{D})\right)^T \tag{32}$$

over possible realizations of $\widehat{X}$.

The canonical game is invariant with respect to the specification of possible payoff distributions $\{\eta(\,\cdot\,, s_i)\}_{i=1,\cdots,I}$. The isomorphism between the original game and the canonical game derives from the explicit finite-dimensional reduction.

Interpretation of Canonical Game The canonical game admits the following natural interpretation as a "trading game on signals." There are pseudo-"Arrow-Debreu markets" corresponding to pseudo-"Arrow-Debreu states" $i = 1, \cdots, I$. Conditional on signal s_i, the i-th "state" will realize with probability one and the other "states" have probability zero. Conditional on observing s_i, the insider submits demand $\hat{d} \in \mathbb{R}^I$ for the "Arrow-Debreu securities" to maximize the "expected profit" $J(\hat{d}; \widehat{D}, i)$ defined in (31). The "noise trader order" for the i-th "state" is $\widehat{X}_i$. The market maker has uniform prior on $\{s_i\}_{i=1,\cdots,I}$ and, after observing total order flow $\hat{d} + \widehat{X}$, sets the "Arrow-Debreu prices" according to (30). This pricing rule encodes the market maker's zero profit condition in the original game.

[50]More precisely, this normalizes the L^2-norms of payoffs over the state space.

An equilibrium in the canonical game is a set of trading strategies $\delta^{(i)} \in \mathbb{R}^I$, $i = 1, \cdots, I$, such that, for all i,

$$\delta^{(i)} \in \underset{\hat{d} \in \mathbb{R}^I}{\operatorname{argmax}}\, J(\hat{d}; D^*, i), \tag{33}$$

where $D^* = [\delta^{(1)} \; \cdots \; \delta^{(I)}] \in \mathbb{R}^{I \times I}$ and $J(\hat{d}; D^*, i)$ is defined in (31).

Proposition 5.4. *Under the isomorphism between the canonical game and the original trading game, an equilibrium D^* of the canonical game gives an equilibrium of the original game (as defined in Definition 3) specified by*

$$W^*(\,\cdot\,, s_i) = \sum_{j=1}^{I} \beta_j^{(i)} \eta(\,\cdot\,, s_j), \quad \textit{where } \beta^{(i)} = \sigma D^* \mathbf{M}^{-\frac{1}{2}}(\eta) e_i \in \mathbb{R}^I \tag{34}$$

for each signal s_i. The matrix $\sigma D^* \mathbf{M}^{-\frac{1}{2}}(\eta)$ will be referred to as equilibrium informed demand in ***canonical form****.*

To express the equation to follow more compactly, let $(p_1, \cdots, p_I)^T$ denote the random vector (32) of the market maker's posterior probabilities in the canonical game, with the dependence on $\hat{d}$, $\widehat{X}$, and $\widehat{D}$ understood. Then, conditional on observing s_i and given market maker belief $\widehat{D}$, the first-order condition for the insider's maximization problem (31) in the canonical game is[51]

$$e_i - \underbrace{\mathbb{E}^{(\hat{d};\widehat{D})} \begin{bmatrix} p_1 \\ \ddots \\ p_I \end{bmatrix}}_{\text{AD term } N_{AD}(\hat{d})} \underbrace{- \widehat{D} \cdot \left(\mathbb{E}^{(\hat{d};\widehat{D})} \left[\begin{bmatrix} p_1 & & \\ & \ddots & \\ & & p_I \end{bmatrix} - \begin{bmatrix} p_1 \\ \vdots \\ p_I \end{bmatrix} \begin{bmatrix} p_1 & \cdots & p_I \end{bmatrix} \right] \right) \cdot \hat{d}}_{\text{price impact term } N_K(\hat{d})} = 0, \tag{35}$$

where $\mathbb{E}^{(\hat{d};\widehat{D})}[\,\cdot\,]$ is the expectation taken over possible pseudo-"total order flow" $\hat{d} + \widehat{X}$, given insider choice $\hat{d}$ and market maker belief $\widehat{D}$.

[51] Under the isomorphism of Proposition 5.3, the marginal cost in the insider's first-order condition (35) in the canonical game retains the same Arrow-Debreu and price impact decomposition obtained in Theorem 4.2 for the original game, but with respect to the pseudo-Arrow-Debreu states. The marginal cost at current demand $\hat{d}$ is the sum

$$N_{AD}(\hat{d}) + N_K(\hat{d}).$$

With respect to the pseudo-Arrow-Debreu markets, the Arrow-Debreu term $N_{AD}(\hat{d})$ is the cost of marginal demand at $\hat{d}$, conditional on no price impact. The price impact term $N_K(\hat{d})$ reflects the price impact of marginal demand on all pseudo-Arrow-Debreu markets.

5.2 Symmetric Equilibrium

Define

$$\mathbf{Q} = \mathbf{I} - \frac{1}{I}\bar{e}\bar{e}^T \in \mathbb{R}^{I\times I} \tag{36}$$

where $\bar{e} = \sum_{i=1}^{I} e_i$.

Equilibrium Ansatz The canonical game is symmetric across signals—*up to permutation on i, the insider's problem is identical across signals.* This symmetry owes to the normalization discussed in Remark 5.2. Given the symmetry of the game and Proposition 4.1(ii), a natural conjecture for equilibrium is

$$D^* = a\mathbf{Q}, \text{ for some } a > 0. \tag{37}$$

That is, conditional on observing s_i, the insider demand is $a(1-\frac{1}{I})$ for the i-th pseudo-AD state and $-\frac{a}{I}$ for the other states. This portfolio of pseudo-AD securities corresponds to the i-th column D^*e_i of the equilibrium ansatz $D^* = a\mathbf{Q}$. Substituting the equilibrium ansatz into the first-order condition (35) for each i and stacking the resulting vector equations side-by-side yield a matrix equation

$$\Phi(a)\mathbf{Q} = 0, \tag{38}$$

where $\Phi(a)$ is a scalar-valued function of a. Provided that, in such a fixed point (38) of the joint first-order conditions, the first-order condition is sufficient for optimality conditional on each signal, an equilibrium is given by a solution a^* to the *equilibrium equation*

$$\Phi(a) = 0. \tag{39}$$

Theorem 5.5. *(Existence of Equilibrium)*

(i) Under Assumption 1, there exists $a^* > 0$ such that $D^* = a^*\mathbf{Q}$ is an equilibrium of the canonical game.

(ii) Therefore, by Proposition 5.4, the original trading game has an equilibrium where the canonical form informed demand is $a^*\sigma\mathbf{Q}\mathbf{M}^{-\frac{1}{2}}(\eta)$.

In the equilibrium obtained in Theorem 5.5, the informed demand is positive in the observed signal and negative in the other signals. The insider's trading strategy conditional on the observed signal therefore shifts the market maker's posterior toward the observed signal. The market maker's posterior on signals, conditional on the observed signal, is biased toward the observed signal. Examples 5.6 and 5.7 demonstrate the special case of Theorem 5.5 where the signal is binary.

Example 5.6. *(Binary Signal)*

Consider the binary signal case, $S = \{s_1, s_2\}$.

(i) <u>*Derivation of Equilibrium Equation* $\Phi(a) = 0$</u>.

The equilibrium ansatz (37) *is*[52]

$$D^* = \begin{bmatrix} a & -a \\ -a & a \end{bmatrix}.$$

Substituting D^* *into the first-order condition of* (35) *conditional on* s_1 *gives*

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} - \begin{bmatrix} \mathbb{E}[p_1] \\ 1 - \mathbb{E}[p_1] \end{bmatrix} - 4a^2 \begin{bmatrix} \mathbb{E}[p_1 p_2] \\ -\mathbb{E}[p_1 p_2] \end{bmatrix} = (1 - \mathbb{E}[p_1] - 4a^2 \mathbb{E}[p_1 p_2]) \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 0 \tag{40}$$

where $\mathbb{E}[\,\cdot\,]$ *is the expectation taken under the conjectured equilibrium* D^*, *and* p_i *is the posterior probability of* s_i. *Substituting* D^* *into the market maker's posterior* (30) *gives*

$$(p_1, p_2) \overset{d}{\sim} \left(\frac{e^Z}{e^Z + 1}, \frac{1}{e^Z + 1}\right), \qquad \text{where } Z \overset{d}{\sim} \mathcal{N}(2a^2, 8a^2). \tag{41}$$

By (41), *the quantities* $\mathbb{E}[p_1]$ *and* $\mathbb{E}[p_1 p_2]$ *in* (40) *are moments of a logit normal distribution and functions of* a. *Write*

$$\mathbb{E}[p_1] = \phi_1(a), \text{ and } \mathbb{E}[p_1 p_2] = \phi_2(a).$$

The first-order condition (40) *conditional on* s_1 *can then be written as*

$$\Phi(a) \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 0 \tag{42}$$

where $\Phi(a) = 1 - \phi_1(a) - 4a^2 \phi_2(a)$. *The first-order condition conditional on* s_2

$$\Phi(a) \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 0 \tag{43}$$

is identical with (42), *up to permuting* s_1 *with* s_2. *Stacking the two symmetric equations* (42) *and* (43) *side-by-side gives the matrix equation* $\Phi(a)\mathbf{Q} = 0$ *of* (38) *in the special case* $I = 2$ *(up to scalar multiple*

[52] To be exact, in the binary signal case, we have

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

as defined by (36), and the equilibrium ansatz of (37) is

$$D^* = a\mathbf{Q} = \frac{1}{2} \begin{bmatrix} a & -a \\ -a & a \end{bmatrix}.$$

In this calculation for the binary signal case, we parameterize by $\frac{1}{2}a$ instead of a.

$\frac{1}{2}$). *An equilibrium is given by a solution* $a^* > 0$ *to* $\Phi(a) = 0$, *which is the equilibrium equation* (39) *in the special case* $I = 2$.

(ii) *Existence of Equilibrium.*

Since

$$\Phi(0) = \frac{1}{2} \quad \text{and} \quad \lim_{a\to\infty} \Phi(a) \uparrow 0,$$

$\Phi(a^*) = 0$ *for some* $a^* > 0$ *by the Intermediate Value Theorem, and the existence of equilibrium follows.*[53] *This proves Theorem 5.5 in the case of binary signal.*

The endogenous constant a^* in Theorem 5.5 solves the equilibrium equation $\Phi(a) = 0$ of (39). It enters the insider's demand by scaling his contingent claim portfolio uniformly across all states and all signals, in both the canonical game and the original game. Example 5.7 further illustrates the existence of equilibrium by considering how the market maker's posterior varies with respect to a under equilibrium belief, in the binary signal case.

Example 5.7. *(Binary Signal—MM Equilibrium Posterior)*

Consider the probability law (41), *obtained Example 5.6, of the market maker's equilibrium posterior conditional on* s_1,

$$(p_1, p_2) \overset{d}{\sim} (\frac{e^Z}{e^Z+1}, \frac{1}{e^Z+1}), \qquad \text{where } Z \overset{d}{\sim} \mathcal{N}(2a^2, 8a^2).$$

(i) *As* $a \to 0$, *the market maker's posterior approaches the uniform prior* $(\frac{1}{2}, \frac{1}{2})$ *with probability one, In other words, if the insider demand is zero for all markets, order flow is pure noise across all markets and there is no updating by the market maker.*

(ii) *As* $a \to \infty$, *the market maker's posterior approaches the point mass* $(1, 0)$ *on the observed signal* s_1 *with probability one. In other words, as the insider scales up his orders uniformly across all markets and signals, eventually he reveals his private signal.*

The insider faces increasing marginal cost.[54] *The marginal cost is the sum of the Arrow-Debreu and price impact terms characterized in Theorem 4.2. Equilibrium* a^* *attains when the marginal cost exactly offsets the marginal payoff. The market maker's equilibrium posterior is biased toward the observed signal* s_1, *as expected.*

[53] It is clear that $\Phi(0) = \frac{1}{2}$. That $\lim_{a\to\infty} \Phi(a) \uparrow 0$ can be seen as follows. Write

$$\Phi(a) = 1 - E[p_1 + 4a^2 p_1 - 4a^2 p_1^2]$$

where

$$p_1 = \frac{e^{2a^2+2aY}}{e^{2a^2+2aY}+1}, \quad Y \overset{d}{\sim} \mathcal{N}(0, 2).$$

We have $p_1 + 4a^2 p_1 - 4a^2 p_1^2 \to 1$ almost surely as $a \to \infty$. Moreover, $p_1 + 4a^2 p_1 - 4a^2 p_1^2 > 1$ eventually almost surely as $a \to \infty$. By Fatou's Lemma, $E[p_1 + 4a^2 p_1 - 4a^2 p_1^2] \to 1$ from above as $a \to \infty$. This implies $\Phi(a) \to 0$ from below as $a \to \infty$.

[54] See Lemma A.2.

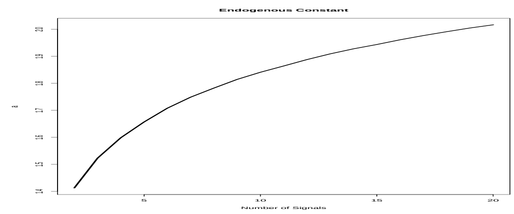


Figure 1: **Comparative Statics of Endogenous Constant** a^* The endogenous constant a^* as a function of number signals I, plotted over the range $2 \leq I \leq 20$.

Comparative Statics of a^* The endogenous constant a^* characterizing equilibrium only depends on the number of signals I. In particular, it is independent of the specification of possible payoff distributions $\{\eta(\,\cdot\,, s_i)\}_{i=1,\cdots,I}$ over states. $a^*(I)$ is an increasing concave function of I, as shown in Figure 1. As the market belief becomes more dispersed, the insider scales up his demand, uniformly across states and signals, to exploit his private information. The marginal gain of scaling up diminishes as the increased variation of his trades across signals also makes it easier for the market maker to infer his private signal.

6 Price Discovery

6.1 Informed Demand

In the equilibrium of Theorem 5.5 characterized by $a^* > 0$, the informed demand in the canonical form of Proposition 5.4 is

$$a^* \sigma \mathbf{Q} \mathbf{M}^{-\frac{1}{2}}(\eta) = [\beta^{(1)} \;\cdots\; \beta^{(I)}] \in \mathbb{R}^{I \times I} \tag{44}$$

where $\beta^{(i)} = a^* \sigma \mathbf{Q} \mathbf{M}^{-\frac{1}{2}}(\eta) e_i$. Explicitly, the *equilibrium informed demand $W^*(\cdot, s_i)$ conditional on signal s_i* is

$$W^*(\,\cdot\,, s_i) = \sum_{j=1}^{I} \beta_j^{(i)} \eta(\,\cdot\,, s_j), \tag{45}$$

which is a long-short portfolio of the type specified in Proposition 4.1(ii). Applying the Breeden-Litzenberger formula (6) to $W^*(\,\cdot\,, s_i)$ gives the equivalent options portfolio.

Informed Demand Portfolio Construction The intuition underlying the informed demand formula (45) can be seen in the following portfolio construction procedure:

- Step 1. Form an initial long-short portfolio where the observed payoff distribution $\eta(\,\cdot\,, s_i)$ receives positive weight $1 - \frac{1}{I}$ and the other distributions $\eta(\,\cdot\,, s_j)$, $j \neq i$, receive equal negative weights $-\frac{1}{I}$.

- Step 2. Adjust the initial portfolio weights by applying the linear transformation $\mathbf{M}^{-\frac{1}{2}}(\eta)$.
- Step 3. Scale the overall portfolio by the endogenous constant a^*.
- Step 4. Finally, further scale the overall portfolio by noise trading intensity σ.

Step 1 constructs an initial portfolio by buying the observed signal and selling the other signals. Step 2 adjusts the initial portfolio for the information intensity of the insider's signal, as measured by $\mathbf{M}(\eta)$. Higher information intensity of signals forces the insider to scale down his demand to reduce the variation of his trades across signals.[55] Step 3 scales the portfolio by the endogenous constant a^* to optimize with respect to dispersion of market belief. Higher dispersion of market belief allows the insider to make larger trades.[56] Step 4 scales the portfolio by σ to optimize with respect to noise trading intensity. Higher noise trading intensity allows the insider to better conceal larger trades.

The informed demand in canonical form

$$a^* \sigma \mathbf{Q} \mathbf{M}^{-\frac{1}{2}}(\eta)$$

and the single asset informed demand from (5)

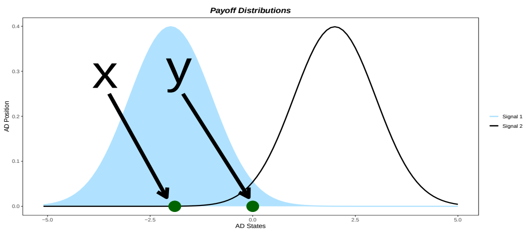
$$\beta(s - v_0) = \sigma_\varepsilon \cdot \frac{1}{\sigma_v}(s - v_0), \quad \beta = \sigma_\varepsilon \cdot \frac{1}{\sigma_v}$$

share the same intuition regarding the information intensity of signal and noise trading intensity. The quantity $\frac{1}{\sigma_v}$ entering β has its contingent claims counterpart in the matrix $\mathbf{M}^{-\frac{1}{2}}(\eta)$. Both σ_v and the (square root of) information intensity matrix $\mathbf{M}^{\frac{1}{2}}(\eta)$ determine the informativeness of order flow. In the single asset setting, higher σ_v means the market maker can attribute more of the variation of order flow to that of the asset fundamental value. In the general contingent claims setting, a larger $\mathbf{M}^{\frac{1}{2}}(\eta)$, in the sense of positive-semidefiniteness, means the market maker can infer more information regarding the payoff distribution from total order flow across AD markets. Consequently, the informed demand is decreasing with respect to σ_v and $\mathbf{M}^{\frac{1}{2}}(\eta)$, in their respective settings. Noise trading intensity σ_ε in β plays exactly the same role as σ in D^*. The informed demand is increasing in noise trading intensity, because more noise trades allows the insider to better hide his trades, whether in a single asset market or across contingent claim markets.

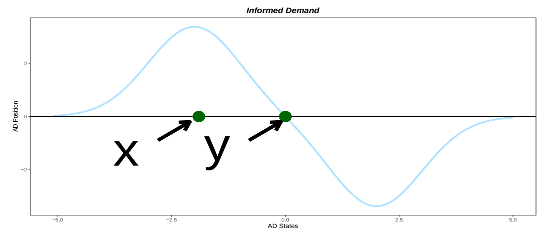
Next, we demonstrate the informed demand for special parametric cases where private information is restricted to the first, second, and third moments of the underlying asset payoff. In Examples 6.1, 6.2, and 6.3, the signal is binary, $S = \{s_1, s_2\}$ and the state space is $(-\infty, \infty)$.

[55] Step 2 reverses the normalization of information intensity across signals applied to derive the canonical game, discussed in Remark 5.2.

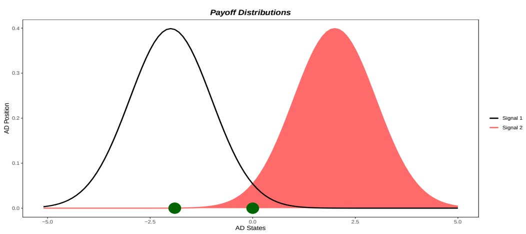
[56] As shown in Figure 1, a^* is increasing with respect to I.



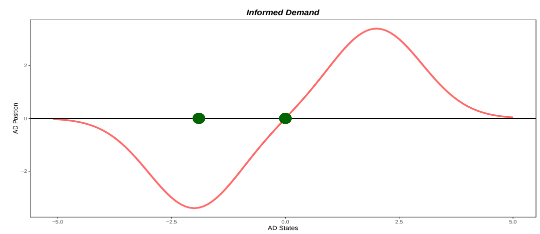
(a) Low expected payoff signal s_1



(b) Informed Demand $W^*(\cdot, s_1)$ Conditional on s_1



(c) High expected payoff signal s_2



(d) Informed Demand $W^*(\cdot, s_2)$ Conditional on s_2

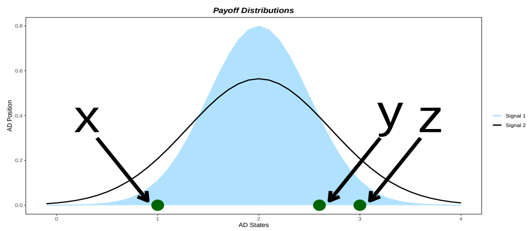
Figure 2:
Insider Trading on Mean (Kyle (1985) with Option Markets) The shaded distributions in the left column indicate the private signal of the insider. The figures in the right column show the informed demand W^*, conditional on the corresponding signals.
Within-Market Price Impact Two states x and y are indicated for illustration of within-market price impact. We have $\Lambda_{x,x} > \Lambda_{y,y}$. State x has a higher within-market price impact than state y because the variation of the insider's signal at x is higher. In fact, state y payoff has zero variation across signals.

Example 6.1. (Insider Trading on Mean)

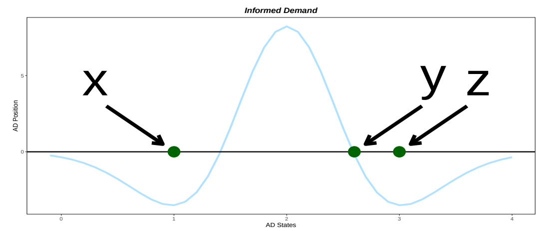
Let signal s_1 (resp. s_2) specify a payoff distribution that is normal $\mathcal{N}(\mu_1, \sigma)$ (resp. $\mathcal{N}(\mu_2, \sigma)$), with $\mu_1 < \mu_2$. This is a Kyle (1985)-type setting, where information asymmetry is restricted to the expected asset payoff, but extended to incorporate derivative markets. Observing s_1 (resp. s_2) informs the insider of low (resp. high) expected payoff.

Figure 2a shows the two possible distributions $\eta(\cdot, s_1)$ and $\eta(\cdot, s_2)$ with the shaded distribution indicating the low expected payoff signal s_1. Figure 2b shows the informed demand $W^*(\cdot, s_1)$ conditional on s_1. When the insider knows the expected payoff is lower than the current market belief, he longs the low payoff states and shorts the other states. The informed demand is nonlinear. In terms of futures and options, the linear short futures-only position of Kyle (1985) is sub-optimal, although it approximates the optimal portfolio on strikes close to the $t = 0$ futures price.

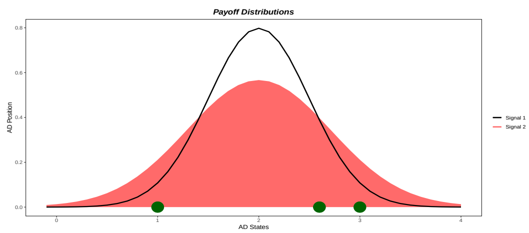
The other case where the insider observes the high expected payoff signal s_2 is symmetric. In Figure 2c, the shaded distribution indicates the signal s_2. Figure 2d shows the informed demand $W^*(\cdot, s_2)$ conditional on s_2, To trade on an expected payoff which is higher than the current market belief, the insider longs the high payoff states and shorts the other states. In terms of futures and options, a linear *long futures position is again sub-optimal.*



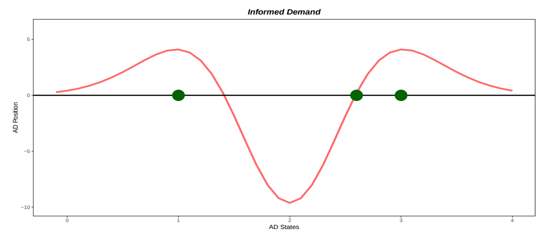
(a) Low volatility signal s_1



(b) Informed Demand $W^*(\cdot, s_1)$—Option Butterfly (short volatility)



(c) High volatility signal s_2



(d) Informed Demand $W^*(\cdot, s_2)$—Straddle (long volatility)

Figure 3:
Insider Trading on Volatility The shaded distributions in the left column indicate the private signal of the insider. The figures in the right column show the informed demand W^* conditional on the corresponding signals. The insider's positions are approximated by option butterfly/straddle positions for shorting/longing volatility.
Cross-Market Price Impact Three states—x, y, and z—are indicated for illustration of cross-market price impact. State y has zero price impact on any state y', i.e. $\Lambda_{y,y'} = 0$. This is because state y payoff has zero variation across signals, $\eta(y, s_1) = \eta(y, s_2)$, which makes order flow at y pure noise—see corresponding *zero informed demand at* y shown in Figures 3b and 3d, $W^*(y, s_1) = W^*(y, s_2) = 0$. The cross price impact between x and z is positive, $\Lambda_{x,z} = \Lambda_{z,x} > 0$. This is because state x and state z payoffs ($\eta(x, \cdot)$ and $\eta(z, \cdot)$, respectively) are positively correlated across signals $s = s_1, s_2$.

Example 6.2. (Insider Trading on Volatility)

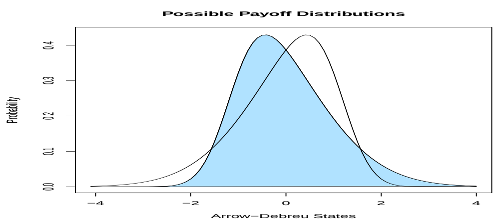
Let signal s_1 (resp. s_2) specify a payoff distribution that is normal $\mathcal{N}(\mu, \sigma_1)$ (resp. $\mathcal{N}(\mu, \sigma_2)$), with $\sigma_1 < \sigma_2$.[57] *The insider therefore has private information regarding the volatility of asset payoff. Observing s_1 (resp. s_2) informs the insider of low (resp. high) volatility.*

Figure 3a shows the two possible distributions $\eta(\cdot, s_1)$ and $\eta(\cdot, s_2)$ with the shaded distribution indicating the low volatility signal s_1 of the insider. Figure 3b shows the corresponding informed demand $W^*(\cdot, s_1)$ conditional on s_1. In terms of options, to trade on the low volatility signal, the insider takes a position that is well approximated by an ***option butterfly****, which is used to trade on low volatility*

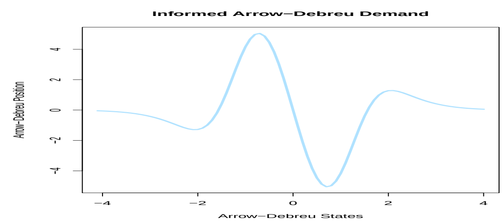
[57] $\mu = 2$, $\sigma_1 = 0.25$, and $\sigma_2 = 0.5$.

signals in practice.[58]

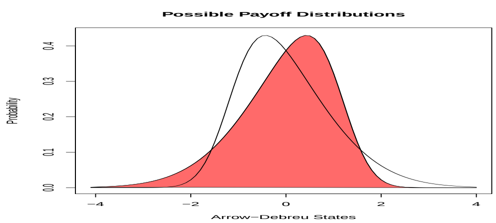
Figure 3d shows the informed demand $W^*(\cdot, s_2)$ conditional on the high volatility signal s_2, which is shaded in Figure 3c. In terms of options, to trade on the high volatility signal, the insider takes a position which is approximated by a ***straddle****.*[59] *The straddle is both used to trade on high-volatility signals in practice and documented in the empirical literature to contain information on expected increase in volatility—see, e.g. Pan and Poteshman (2006), Ni et al. (2008), and Goyal and Saretto (2009). An alternative implementation of the informed demand $W^*(\cdot, s_2)$ is a short butterfly.*



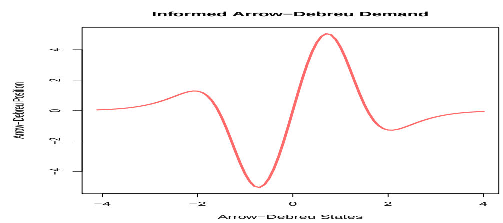
(a) Left-skewed signal s_1



(b) Informed Demand $W^*(\cdot, s_1)$



(c) Right-skewed signal s_2



(d) Informed Demand $W^*(\cdot, s_2)$

Figure 4: **Insider Trading on Skewness** The shaded distributions in the left column indicate the private signal of the insider. The figures in the right column show the informed demand W^* conditional on the corresponding signals.

Example 6.3. (Insider Trading on Skewness)

This example considers payoff distributions with different skewness while controlling for all other moments. Let signal s_1(resp. s_2) specify a skew normal payoff distribution $\eta(\cdot, s_1)$(resp. $\eta(\cdot, s_2)$ with skewness parameter α_1(resp. α_2), where $\alpha_1 < \alpha_2$.[60] *The mean and variance are the same across both signals. Figure 4a shows the two distributions $\eta(\cdot, s_1)$ and $\eta(\cdot, s_2)$ with the shaded distribution indicating the left-skewed signal s_1 of the insider. Figure 4b shows the corresponding informed demand $W^*(\cdot, s_1)$. Figure 4c and Figure 4d show the other case when the insider observes the right-skewed signal s_2. In*

[58] For a practitioner's perspective on option butterfly, see e.g. `https://www.cmegroup.com/education/courses/option-strategies/option-butterfly.html`.

[59] A straddle consists of a call option and a put option on the same underlying with the same strike price. For a practitioner's perspective, see e.g. `https://www.cmegroup.com/education/courses/option-strategies/straddles.html`.

[60] $\alpha_1 = -3$ and $\alpha_2 = 3$.

practitioner parlance, the informed demand can be implemented by long-short portfolios of so-called "Christmas tree" options positions, with the long position aligned with the direction of skewness.

6.2 Price Impact

We now characterize the equilibrium price impact across AD markets. Let $W^*(\,\cdot\,,\,\cdot\,)$ be the insider's equilibrium strategy (45). Substituting $W^* = W = \widetilde{W}$ into the price impact expression (25) of Corollary 4.3 yields the *equilibrium price impact* $\Lambda_{x,y}$ *between state* x *and* y *markets conditional on the insider observing* s_i,

$$\Lambda_{x,y} = (1 - \frac{1}{I}) \cdot a^* \cdot \frac{1}{\sigma} \cdot \mathbb{E}[\mathbb{C}\mathbf{ov}^{\pi_1^*(ds,\omega;s_i)}(\eta(x,\,\cdot\,), \eta(y,\,\cdot\,))], \tag{46}$$

where $\mathbb{C}\mathbf{ov}^{\pi_1^*(ds,\omega;s_i)}(\eta(x,\,\cdot\,), \eta(y,\,\cdot\,))$ is the covariance of the two states' payoff profiles $\eta(x,\,\cdot\,)$ and $\eta(y,\,\cdot\,)$ under the market maker's equilibrium posterior $\pi_1^*(ds, \omega; s_i)$ conditional on order flow ω, and $\mathbb{E}[\,\cdot\,]$ is taken with respect to its equilibrium probability law over ω. The equilibrium probability law of $\pi_1^*(ds, \omega; s_i)$ over ω is characterized in Proposition 6.9. The following corollary is immediate.

Corollary 6.4. *Cross price impact is symmetric with respect to states,* $\Lambda_{x,y} = \Lambda_{y,x}$ *for all* x *and* y.

Cross price impact is symmetric with respect to states because in equilibrium the market maker infers equal amount of information from the marginal demands across any given pair of states.

Price Impact as Complementary Product The intuition underlying the cross price impact between two states x and y can be seen in each of the quantities entering the complementary product defining $\Lambda_{x,y}$:

- $(1 - \frac{1}{I})$—the portfolio weight of the true payoff distribution in the informed demand (44).
- a^*—the endogenous constant. As the informed demand (44) scales up by a^*, this results in a one-to-one response for price impact.
- $\frac{1}{\sigma}$—the reciprocal of noise trading intensity. Price impact moves inversely with respect to noise trading intensity σ. Higher noise trading intensity means more noisy order flow for the market maker and lowers price impact.
- $\mathbb{E}[\mathbb{C}\mathbf{ov}^{\pi_1^*(ds,\omega;s_i)}(\eta(x,\,\cdot\,), \eta(y,\,\cdot\,))]$—the (expected) covariance of state x and y payoff profiles under the market maker's equilibrium posterior. Because of the information spillover between markets, the covariances of payoff profiles across signals are endogenous, in contrast to the single asset setting. When the state x and y payoffs are highly correlated across signals, the insider's demands for x and y securities are also highly correlated across signals. For the market maker, order flow

for x is therefore informative regarding the true payoff of y observed by the insider. This leads to high price impact between x and y markets. Example 6.5 further illustrates this intuition.

Example 6.5.

(i) (Negative Cross Price Impact.) Let x and y be two states. If x can only realize under signal s and y can only realize under signal $s' \neq s$, then

$$\Lambda_{x,y} = -(1 - \frac{1}{I}) \cdot a^* \cdot \frac{1}{\sigma} \cdot \mathbb{E}[\eta(x, s)\pi_1^*(s, \omega; s_i) \cdot \eta(y, s')\pi_1^*(s', \omega; s_i)] < 0.$$

Cross price impact is therefore negative, regardless of the insider's signal s_i. This is similar to Scenario (i) of Example 4.4, but in equilibrium. If the market maker receives a buy order at x, he infers that state y security likely has zero payoff and lowers its price.

(ii) (Zero Cross Price Impact.) If state y payoff $\eta(y, \cdot)$ has zero variation across signals, then $\Lambda_{y,y'} = 0$ for all y'. That is, state y has zero price impact on all states regardless of the signal. This means that the equilibrium order flow at y must contain no informed demand regardless of the signal—otherwise there is arbitrage. This is the case for the state y indicated in Figure 3, where

$$\eta(y, \cdot) = (\eta(y, s_1), \eta(y, s_2))$$

has zero variation across signals $\{s_1, s_2\}$, i.e. $\eta(y, s_1) = \eta(y, s_2)$. Figures 3b and 3d show that the corresponding informed demand $W^*(y, \cdot) = 0$ at y is indeed zero regardless of signal, making order flow at y pure noise with no information content—hence no price impact.

(iii) (Positive Cross Price Impact.) For the two states x and z indicated in Figure 3, the cross price impact $\Lambda_{x,z} = \Lambda_{z,x} > 0$ is positive, because the payoffs of x and z,

$$\eta(x, \cdot) = (\eta(x, s_1), \eta(x, s_2)) \;\; \textit{and} \;\; \eta(z, \cdot) = (\eta(z, s_1), \eta(z, s_2)),$$

are positively correlated across signals $\{s_1, s_2\}$.

The price impact between general derivative markets is given by the following formula.

Corollary 6.6. *(Price Impact Between General Derivatives)*

Let $\varphi_1, \varphi_2 : [\underline{x}, \bar{x}] \to \mathbb{R}$ be state-contingent claims (e.g. options). Then the cross price impact between the φ_1 and φ_2 markets is given by

$$\int_{\underline{x}}^{\bar{x}} \int_{\underline{x}}^{\bar{x}} \varphi_1(x)\varphi_2(y)\Lambda_{x,y} dx dy. \tag{47}$$

For example, high cross price impact between the put and call pair in a straddle predicts high volatility, as recorded in Proposition 6.7 below.

Proposition 6.7. *Under the parametric specification of Example 6.2, the cross price impact between the constituent put and call pair in a straddle is higher conditional on the high-volatility signal s_2 than on the low-volatility signal s_1.*

Intuitively, high cross price impact between the put and call pair in a long-volatility straddle means that volatility is underpriced and therefore should correct to a higher price level as price discovery takes place. Such considerations can be extended to higher moments beyond volatility and yield testable empirical predictions—see Section 7.2.

$\mathbf{\Lambda_{x,x}}$ and Kyle's lambda Specializing $\Lambda_{x,y}$ to $x = y$, we have the *equilibrium price impact* $\Lambda_{x,x}$ *within state x market conditional on s_i*,

$$\Lambda_{x,x} = (1 - \frac{1}{I}) \cdot a^* \cdot \frac{1}{\sigma} \cdot \mathbb{E}[\mathbb{V}\mathrm{ar}^{\pi_1^*(ds,\omega;s_i)}(\eta(x,\,\cdot\,))].$$

The intuition for Kyle's λ discussed in Section 2.2 applies identically to $\Lambda_{x,x}$. $\Lambda_{x,x}$ is proportional to the (expected) posterior variance $\mathbb{E}[\mathbb{V}\mathrm{ar}^{\pi_1^*(ds,\omega;s_i)}(\eta(x,\,\cdot\,))]$ of the state x security payoff, while Kyle's λ of (5) is proportional to the variation σ_v of the asset payoff. Price impact within an asset market is high when the asset payoff has high variation across signals, whether the asset is the only asset being traded or not.

Example 6.8. *For the two states x and y indicated in Figure 2, we have $\Lambda_{x,x} > \Lambda_{y,y} = 0$. State x has strictly positive within-market price impact because the state x payoff*

$$\eta(x,\,\cdot\,) = (\eta(x, s_1), \eta(x, s_2))$$

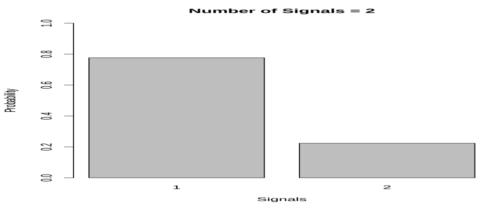
has strictly positive variation across signals $\{s_1, s_2\}$. On the other hand, the state y payoff has zero variation across signals. This can be seen in Figure 2a, which shows that $\eta(y, s_1) = \eta(y, s_2)$, with the s_1 and s_2 densities intersecting at y. Therefore, as in Example 6.5(ii), the equilibrium order flow at y contains zero informed demand (see Figures 2b and 2d) and $\Lambda_{y,y'} = 0$ for all y'.

6.3 Information Efficiency of Prices

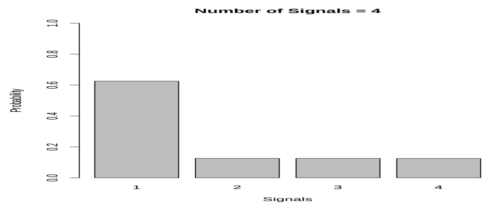
The equilibrium probability law over ω of the market maker's posterior probability mass function $\pi_1^*(ds, \omega; s_i)$ conditional on the insider observing s_i admits the following canonical characterization.

Proposition 6.9. *(Canonical Logistic Posterior)*

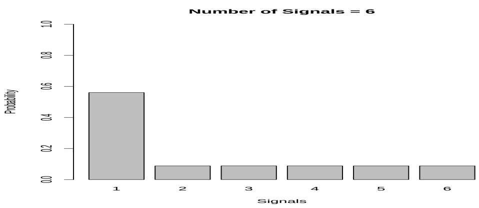
Let a^* be the endogenous constant characterizing equilibrium in Theorem 5.5, and $X \overset{d}{\sim} \mathcal{N}(0, (a^*)^2\,\mathbf{Q})$ be a multivariate normal random vector in $\mathbb{R}^I$ with mean 0 and covariance matrix $(a^*)^2\,\mathbf{Q}$. The market



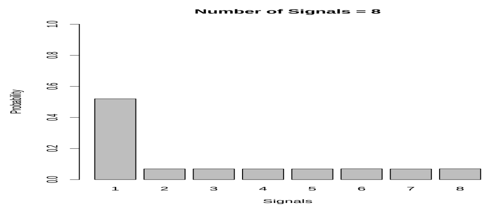
(a) $I = 2$



(b) $I = 4$

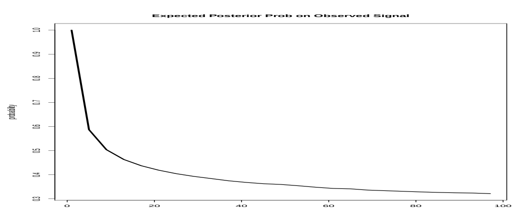


(c) $I = 6$

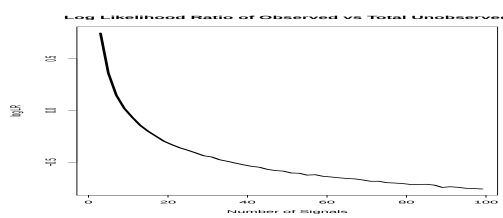


(d) $I = 8$

Figure 5: **Market Maker's Expected Posterior** The figures show the market maker's expected posterior probability mass functions on the signal space $S = \{s_1, \cdots, s_I\}$, conditional on the insider observing signal s_1, as the number I of signals varies. The cases shown are for $I = 2, 4, 6,$ and 8. In expectation, the probability assigned by the market maker's posterior to the observed signal s_1 decreases as I increases.



(a) Information Efficiency $\mathbb{E}[q_i^{(i)}]$



(b) Expected Posterior Log-likelihood Ratio of Observed vs All Other Signals

Figure 6: **Information Efficiency of Prices** Figure 6a shows the information efficiency of prices, as measured by the expected value $\mathbb{E}[q_i^{(i)}]$ of probability q_i assigned by market maker's posterior to the observed signal s_i, as a function of I over $2 \leq I \leq 100$. Figure 6b shows the expected posterior log-likelihood ratio $\log(\frac{\mathbb{E}[q_i^{(i)}]}{(I-1)\mathbb{E}[q_{-i}^{(i)}]})$ of observed signal s_i vs all other signals s_{-i}, over $2 \leq I \leq 100$.

maker's equilibrium posterior conditional on s_i has the same distribution as the logistic transform $q^{(i)}$ of X defined by[61]

$$q^{(i)} = \left(\frac{e^{X_1}}{\sum_j \cdots}, \cdots, \frac{e^{(a^*)^2 + X_i}}{\sum_j \cdots}, \cdots \frac{e^{X_I}}{\sum_j \cdots} \right)^T. \tag{48}$$

Figure 5 shows the (expected) posterior probability mass function $\mathbb{E}[q_j^{(1)}]$, $j = 1, \cdots, I$, conditional

[61] "$\sum_j \cdots$" is the random normalization constant so that $\sum_j q_j^{(i)} = 1$.

on the signal s_1, as I increases. By Proposition 6.9, the equilibrium probability law of the market maker's posterior on signals is *independent* of the possible payoff distributions $\{\eta(\,\cdot\,, s_i)\}_{i=1,\cdots,I}$, as well as the noise trading intensity σ. In equilibrium, the market maker correctly anticipates the insider's (ex ante) optimal trade-off of price impact along two dimensions—*between states conditional on signal* and *across signals*. Therefore, these two dimensions are informationally independent in equilibrium. The market maker's posterior on signals (48) summarizes information *across signals* and is therefore invariant with respect to the variation of asset payoffs *between states conditional on signal*. The probability law of AD prices for a given specification of possible payoff distributions is now an immediate corollary.

Corollary 6.10. *(Equilibrium Arrow-Debreu Prices)*

Let $q^{(i)}$, $i = 1, \cdots, I$, be the logistic-normal random vector of Proposition 6.9. In equilibrium, the market maker's pricing kernel $P^*(x, \omega; s_i)$, conditional on s_i, is distributed as

$$P^*(x, \omega; s_i) \overset{d}{\sim} \sum_{j=1}^{I} q_j^{(i)} \eta(x, s_j),$$

and the expected state x AD price $\overline{P}(x; s_i)$ conditional on s_i is given by

$$\overline{P}(x; s_i) = \sum_{j=1}^{I} \mathbb{E}[q_j^{(i)}] \eta(x, s_j). \tag{49}$$

The equilibrium AD prices are given by a (random) convex sum (49) of possible payoff distributions. The **information efficiency** of AD prices can be measured by the (expected) weight $\mathbb{E}[q_i^{(i)}]$ assigned to the true payoff distribution $\eta(\,\cdot\,, s_i)$ in this convex sum. Corollary 6.10 has the following immediate implication.

Corollary 6.11. *The information efficiency measure $\mathbb{E}[q_i^{(i)}]$ only depends on the number of signals I. In particular, it is independent of the specification of payoff distributions $\{\eta(\,\cdot\,, s_i)\}_{i=1,\cdots,I}$ or the noise trading intensity σ.*

Remark 6.12. *Corollary 6.11 states a basic property of complete markets with respect to price discovery—the information efficiency of contingent claim prices depends only on the dispersion of market belief but not the payoffs specified by that belief. In other words, the extent to which information is impounded into prices is invariant with respect to the nature of information. As the insider's trades (co-)vary across states according to his information intensity, the market maker correctly anticipates this (co-)variation and equilibrates prices to incorporate the insider's private information to the extent characterized by $\mathbb{E}[q_i^{(i)}]$, regardless of the information being traded on.*[62] *This invariance result can be viewed as the price*

[62] While the amount of information inferred by the market is invariant, the nature of the information inferred depends

discovery counterpart of risk sharing in complete markets, where market prices equilibrate to allocate risk efficiently, regardless of the idiosyncratic risks being hedged by agents.

Figure 6a shows that $\mathbb{E}[q_i^{(i)}]$ is a decreasing and convex function of I. We record this as a proposition.

Proposition 6.13. *The information efficiency of prices decreases as I increases. The marginal effect of I on information efficiency is diminishing.*

Proposition 6.13 is very intuitive. Dispersion of market belief hinders the ability of the market to infer the insider's private information from trades and therefore reduces information efficiency. The marginal effect is diminishing because, as the informed demand scales up to exploit higher market uncertainty, the increased variation of informed demand across signals also reveals some of the insider's private information and partially offsets the loss in information efficiency. Taking the statement of Proposition 6.13 to the limit $I \to \infty$ gives a canonical lower bound on information efficiency.

Proposition 6.14. *As $I \to \infty$, the information efficiency measure $\mathbb{E}[q_i^{(i)}]$ approaches a canonical lower bound $\rho > 0$*

Figure 6b shows the (expected) posterior log-likelihood ratio $\log \frac{\mathbb{E}[q_i^{(i)}]}{(I-1)\cdot\mathbb{E}[q_{-i}^{(i)}]}$ of observed signal s_i versus all the other signals, as a function of I.[63] As $I \to \infty$,

$$\log \frac{\mathbb{E}[q_i^{(i)}]}{(I-1)\cdot\mathbb{E}[q_{-i}^{(i)}]} \to -\kappa$$

for some canonical constant $\kappa > 0$. This implies that

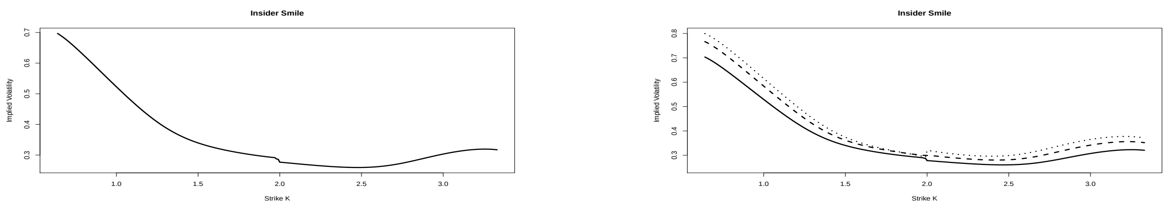
$$\mathbb{E}[q_i^{(i)}] \approx \frac{(I-1)e^{-\kappa}}{(I-1)e^{-\kappa} + (I-1)}$$

for large I. Therefore the information efficiency lower bound ρ of Proposition 6.14 is given by $\rho = \lim_{I\to\infty} \mathbb{E}[q_i^{(i)}] = \frac{e^{-\kappa}}{e^{-\kappa}+1}$.

The existence of a canonical lower bound on information efficiency is consistent with the comparative statics of the endogenous constant a^*. $a^*(I)$ is increasing and concave with respect to I and approaches a limit $\lim_{I\to\infty} a^*(I) = \alpha < \infty$ as $I \to \infty$ (as shown in Figure 1). As the dispersion of market belief increases, the marginal loss in information efficiency comoves with the insider's marginal gain as he scales up his trades. The insider's marginal gain from scaling up is decreasing, therefore so is the marginal loss in information efficiency. Eventually, the marginal gain reaches zero, the insider stops scaling up, and information efficiency reaches its lower bound.

on that of the insider's private information.

[63]The subscript $-i$ denotes any given j with $j \neq i$.



(a) (b)

Figure 7: **The Insider Smile (Smirk)** Figure 7a shows the strike structure of the implied volatility of the equilibrium option prices in the parametric setting of Example 6.2, which is computed as follows. For a realization ω of total order flow, use the pricing kernel $P^*(\,\cdot\,, \omega; s_2)$ to compute the out-of-the-money option prices for the entire spectrum of strikes $K \in [\underline{x}, \bar{x}]$. Given the option prices, the corresponding implied volatilities $(\sigma_{IV}(\omega, K))_{K \in [\underline{x}, \bar{x}]}$ are computed by inverting the Black-Scholes formula. This is done for $m = 1000$ realizations ω_k, $k = 1, \cdots, m$. The plot shows the resulting sample average of implied volatilities $\frac{1}{m} \sum_{k=1}^{m} \sigma_{IV}(\omega_k, K)$ over strike K. (The time-to-maturity parameter is set to 1. The risk-free rate and dividend yield are set to zero.) Figure 7b shows the insider smiles corresponding to $\sigma_2 = 0.5$ (same solid line as in Figure 7a), 0.6 (dashed line), and 0.7 (dotted line).

7 Discussion

7.1 The Insider Smile

The Black-Scholes implied volatility is a common transformation of option prices to express the volatility information they contain. The strike structures of implied volatilities are known to exhibit the *volatility smile (or smirk)*, which points to the misspecification of the Black-Scholes model. The option pricing and financial econometrics literatures continue to develop specifications to better capture this empirical regularity. A microfounded understanding provides a complementary perspective. Our framework shows that the volatility smile arises through the incorporation of, indeed, volatility information into option prices.

Consider the parametric setting of Example 6.2, where the insider is trading on the high volatility signal. The strike structure of the implied volatility of the equilibrium option prices is shown in Figure 7a. The implied volatility forms an *insider smile* that matches well qualitatively with a typical pattern of volatility smile, with skewness towards lower strikes. The intuition of the insider smile is simple. As the insider trades on high volatility, he buys options that pay off at deep in-the-money and out-of-the-money strikes, e.g. a straddle. The resulting price impact pushes up the prices of options at such strikes. In this case, the price impact is amplified cross-strike because the payoffs at in-the-money and out-of-the-money strikes are positively correlated.[64] Since the implied volatility is increasing with respect to option price, this leads to higher implied volatilities as strikes move away from being at-the-money and results in a smile.

[64]Consider, for example, the states x and z indicated in Figure 3.

Figure 7b shows the insider smiles for different values of the high volatility signal. The insider smile becomes more convex as the true volatility known to the insider increases. This is because a higher volatility signal means the covariance of payoffs between in-the-money and out-of-the-money strikes increases at a faster rate as the strikes move away from being at-the-money. Therefore the cross-strike price impact also increases at a faster rate, which makes the insider smile more convex.[65]

7.2 Empirical Considerations

Higher Moment Adverse Selection The empirical options literature has similarly resorted to using the implied volatility to extract volatility information from option prices (see, for example, Goyal and Saretto (2009)). This is an ad hoc measure because the Black-Scholes formula itself assumes that volatility is constant and options are redundant. Our theory provides an empirical measure for the information content of option prices that is based on cross price impact and can be adapted to moments beyond volatility. Namely, higher moment adverse selection in option markets can be measured by the *cross price impact between appropriately specified option contracts.* In other words, options cross price impact can be used to back out higher-order moment information. For instance, under the specification of Example 6.2 where private information concerns volatility, adverse selection with respect to volatility in option markets can be measured by the cross price impact between the put-call pair in a straddle.[66] When the price impact is high between the put-call pair in a long-volatility straddle, the market maker adverse-selects with respect to volatility, and volatility is underpriced.

Predictability of Return Higher Moments The adverse selection measures give rise to empirical hypotheses regarding the predictability of higher moments of the underlying return. In the case of volatility, this corresponds to the statement of Proposition 6.7:

Hypothesis 7.1. *Higher cross price impact between the option contracts in a straddle (resp. butterfly) portfolio predicts an increase (resp. decrease) in the volatility of the underlying return.*

Similarly, under the specification of Example 6.3, higher cross price impact between the long-short pair in the insider option portfolio conditional on the right-skewness signal predicts an increase in the skewness of the underlying return. Beyond the options strategies demonstrated in Examples 6.2 and 6.3, the informed demand formula (45) prescribes options strategies for trading on any give moment. Applying the informed demand formula and the cross price impact formula of Corollary 6.6 to each moment yields the following general hypothesis.

[65]Calibration of the insider smile with respect to data and quantitative implications are outside the scope of this paper and to be considered in future research.

[66]This is the same mechanism that explains the volatility smile, which embodies the inadequacy of the constant volatility assumption.

Hypothesis 7.2. *The cross price impact between appropriately specified option contracts predicts the corresponding moments of the underlying return.*[67]

To the extent that the cross price impact between the prescribed option contracts in Hypothesis 7.2 predicts the corresponding moments, the corresponding option portfolios should generate positive returns. For example, if the cross price impact between the put-call pair in a straddle predicts future volatility, the straddle should generate positive return. The same consideration applies to higher-order moments in general. If the prescribed options portfolios generate positive returns, it implies that mispricing of higher moments across option markets is corrected over time as price discovery takes place. A natural question is whether this holds only up to a certain number of moments or only up to certain predictive horizon. Or, Hypothesis 7.2 may be rejected for certain (e.g. lower-order) moments because option markets are informationally efficient with respect to those moments.

Question 7.3. *What are the moments for which the corresponding option portfolios in Hypothesis 7.2 generate positive returns? In other words, to what extent does higher-order moment price discovery take place in option markets?*

Cross-Section of Option Returns The cross-section of option returns remains a puzzle (see, for example, Bali and Murray (2013), Cao and Han (2013), Christoffersen et al. (2018), and Zhan et al. (2022)). Existing literature has proposed different underlying stock characteristics (Bali and Murray (2013), Cao and Han (2013), and Zhan et al. (2022)) or liquidity measures (Christoffersen et al. (2018)) as determinants of option returns. Considering the higher-moment predictability of Hypothesis 7.2 in the cross-sectional, instead of time series, sense suggests a distinct set of candidate *higher-order information factors* for the cross-section of option returns.

Conjecture 7.4. *Systematic factors for the cross-section of option returns may be constructed using those prescribed options portfolios which generate returns from trading on higher moments.*[68]

7.3 Risk Aversion

If the insider is risk-averse with utility function U, his portfolio choice problem (15) conditional on observing signal s becomes

$$\max_{W(\cdot)\in C^{\delta}[\underline{x},\bar{x}]} \mathbb{E}^{\mathbb{P}_W}[\int_{\underline{x}}^{\bar{x}} \eta(x,s)\cdot U\left(W(x)-P(x,\omega;\widetilde{W})\right)]dx. \tag{50}$$

[67]The hypothesis may be tested by regressing realized moments (e.g. squared returns in the case of volatility) on the lagged cross price impact between prescribed option contracts (e.g. the put-call pair in a straddle).

[68]The candidate factors may be constructed by sorting prescribed options portfolios according to the cross price impact between options in the portfolio and forming long-short portfolios as factors—for example, sorting straddles according to the cross price impact between the constituent put-call pair into quintiles and forming long-short $5-1$ portfolios.

In the expression of the marginal utility functional $dJ(W;\widetilde{W},s)(\cdot)$ of Theorem 4.2, for each state x the payoff $\eta(x,s)$ is replaced by the marginal utility-weighted payoff[69]

$$\mathbb{E}^{\mathbb{P}_W}[\eta(x,s)\cdot U'(W(x)-P(x,\omega;\widetilde{W}))].$$

As in Theorem 4.2, the insider's marginal disutility decomposes into an Arrow-Debreu term, which now reflects risk-sharing between the insider and the market maker, and a price impact term. More generally, when both agents are risk-averse, the insider's first-order condition equalizes the marginal utility-weighted payoff with (up to a Lagrange multiplier) the sum of the marginal disutility at current price—under a pricing kernel adjusted for the market maker's risk-aversion—conditional on no price impact and the marginal disutility from price impact. With risk-averse agents, the Arrow-Debreu prices aggregate risk as well as information. Further developing this case may shed light on the extent to which (higher-order) adverse selection distorts risk-sharing incentives. This is outside the scope of the current paper.

8 Conclusion

We have considered a basic framework of price discovery across contingent claim markets and characterized the economic mechanism through which contingent claim prices jointly incorporate information. Price discovery takes place not only within markets but also between markets. Our setting extends the theory of informed trading to derivatives and higher-order information. In doing so, our results explain empirical practice and stylized facts, bridge the gap with the empirical literature, and yield new empirical questions. Within the general framework, special cases of interest and possible extensions can be analyzed further. Our methodology provides tools to explore these avenues of future research.

A Appendix

A.1 Mathematical Assumptions

The rigorous formulations of both the market maker's and insider's problems are obtained under the following assumptions. These assumptions are maintained for all mathematical statements made in the paper.

For $\delta > 0$, recall the Hölder space $C^{\delta}([\underline{x},\bar{x}],\mathbb{R})$ consists of δ-Hölder continuous elements of $C([\underline{x},\bar{x}],\mathbb{R})$, equipped with the Hölder seminorm $[f]_\delta = \sup_{x\neq y} \frac{|f(x)-f(y)|}{|x-y|^\delta}$.

[69] The explicit expression of $dJ(W;\widetilde{W},s)(\cdot)$ is given in Section A.3 of the Appendix.

Assumption A.1.

Let $\mathcal{F}$ be the Borel σ-field generated by the uniform norm $\|\cdot\|_\infty$ on $\Omega = C([\underline{x},\bar{x}],\mathbb{R})$, μ be the probability measure on the measurable space $(\Omega,\mathcal{F})$ such that the canonical process $x \mapsto \omega_x$ has the same probability law as $(\sigma(x)B_x)$.

Fix $\gamma \in (\frac{1}{3},\frac{1}{2})$ and $\delta \in (0,1]$ with $\delta+\gamma > 1$. Let $C^\delta([\underline{x},\bar{x}],\mathbb{R})$ and $C^\gamma([\underline{x},\bar{x}],\mathbb{R})$ be the corresponding Hölder spaces, and $\Omega_\gamma = \{\omega \in \Omega : \omega(0) = 0$ and $\omega \in C^\gamma([\underline{x},\bar{x}],\mathbb{R})\}$.

(i) For every $s \in S$, the insider's portfolio of contingent claims $W(\cdot, s)\colon [\underline{x},\bar{x}] \to \mathbb{R}$ lies in $C^\delta([\underline{x},\bar{x}],\mathbb{R})$.

(ii) Possible realization ω of total order flow across states lies in the measurable space $(\Omega_\gamma,\mathcal{F}_\gamma)$, where $\mathcal{F}_\gamma = \mathcal{F}\cap\Omega_\gamma$, and μ_γ denotes the measure μ restricted to Ω_γ.

(iii) The insider's strategy $s \mapsto W(\cdot, s)$, $S \to (C^\delta[\underline{x},\bar{x}],\mathbb{R})$, is continuous.[70]

(iv) The (squared) noise trading intensity $\sigma^2(\cdot)$ across Arrow-Debreu markets lies in $C^\delta([\underline{x},\bar{x}],\mathbb{R})$ and $\sigma^2(x) > 0$ for all x.

The condition $\delta+\gamma > 1$ in Assumption A.1 ensures that the integrand $W(\cdot, s)$ and integrator ω together have sufficient pathwise (Hölder) regularity to define a pathwise integral. This condition means that $W(\cdot, s)$ and ω jointly satisfy a regularity condition that is infinitesimally stronger than the Lipschitz condition.[71] Assumption A.1(i) imposes Hölder δ-continuity in x on $W(x, s)$. Assumption A.1(ii) replaces the underlying probability space $(\Omega,\mathcal{F},\mu)$ by $(\Omega_\gamma,\mathcal{F}_\gamma,\mu_\gamma)$, where Hölder γ-continuity holds for the sample paths. This is without loss of generality because $\Omega_\gamma \in \mathcal{F}$ and $\mu(\Omega_\gamma) = 1$.[72]

Assumption A.1(iii) assumes continuity of insider's strategy with respect to signal s. This imposes no constraint on the insider's strategy when $\eta(\cdot, s)$ is continuous in s. Assumption A.1(iv) assumes that noise traders are present in all AD markets. If noise traders are absent in a neighborhood of states, there would be no trade in that neighborhood of states.

Lemma A.1.

Under Assumption A.1, the following holds.

(i) (Pathwise Integral) For all $\omega \in \Omega_\gamma$, $W \in C^\delta([\underline{x},\bar{x}],\mathbb{R})$, and $x \in [\underline{x},\bar{x}]$, the limit of Riemann sums

$$\int_{\underline{x}}^{x} \frac{W_y}{\sigma^2(y)}\,d\omega_y \equiv \lim_{\substack{\max_k |x_{k+1}-x_k|\to 0\\ \underline{x}=x_0<\cdots<x_n=\bar{x}}} \sum_{k=0}^{n-1} \frac{W_{x_k}}{\sigma^2(x_k)}\cdot[\omega_{x_{k+1}\wedge x} - \omega_{x_k\wedge x}]$$

exists and therefore defines an ω-by-ω Riemann integral.

[70]The topology on S is the same topology that generates the Borel σ-field on S.

[71]Having a Hölder exponent strictly larger than one can be interpreted as infinitesimal stronger than being Lipschitz.

[72]$\mu(\Omega_\gamma) = 1$ because Brownian paths lie in $C^{\frac{1}{2}-\epsilon}([\underline{x},\bar{x}],\mathbb{R})$ for any $\epsilon \in (0,\frac{1}{2})$.

(ii) (Joint Measurability of Data and Parameter) The map

$$\underbrace{(\omega, W)}_{\text{(data, parameter)}} \quad \mapsto \int_{\underline{x}}^{\bar{x}} \frac{W_x}{\sigma^2(x)}\, d\omega_x, \quad (\Omega_\gamma, [\,\cdot\,]_\gamma) \times C^\delta([\underline{x}, \bar{x}], \mathbb{R}) \to \mathbb{R},$$

is continuous—in particular, measurable.

(iii) (Conditional Likelihood of Data) For all $x \in [\underline{x}, \bar{x}]$ and $W \in C^\delta([\underline{x}, \bar{x}], \mathbb{R})$, there exists a μ_γ-null set N, which may depend on x and W, such that, for all $\omega \in \Omega_\gamma \setminus N$,

$$\int_{\underline{x}}^{x} \frac{W_y}{\sigma^2(y)}\, d\omega_y = \left[\int_{\underline{x}}^{x} \frac{W_y}{\sigma^2(y)}\, dB_y \right](\omega),$$

where the integral on the left-hand side is the pathwise integral defined in (i) and the integral on the right-hand side is a version of the Itô integral.

Proof. (i) We note first that Hölder continuity is preserved by taking quotients when the denominator is bounded away from zero. Let W, $\sigma \in C^\delta([\underline{x}, \bar{x}], \mathbb{R})$ with $\min_{x \in [\underline{x}, \bar{x}]} \sigma(x) > \alpha > 0$ for some $\alpha > 0$. Then

$$\begin{aligned} \left| \frac{W(y)}{\sigma(y)} - \frac{W(x)}{\sigma(x)} \right| &\leq \frac{|\sigma(x)|\ |W(y) - W(x)| + |W(x)|\ |\sigma(y) - W(x)|}{|\sigma(y)|\ |\sigma(x)|} \\ &\leq \frac{\|W\|_\infty\, [W]_\gamma + \|\sigma\|_\infty\, [\sigma]_\gamma}{\alpha^2} \cdot |y - x|^\delta, \end{aligned}$$

which implies $\frac{W}{\sigma} \in C^\delta([\underline{x}, \bar{x}], \mathbb{R})$. Therefore it suffices to prove the claim with W in place of $\frac{W}{\sigma^2}$.

Given α, $\beta > 0$, we shall consider the space $C_2^{\alpha,\beta}([\underline{x}, \bar{x}], \mathbb{R})$ of all functions Ξ from $\{(y, x) : \underline{x} \leq y \leq x \leq \bar{x}\}$ to $\mathbb{R}$ with seminorm

$$[\Xi]_{\alpha,\beta} \;\equiv\; \underbrace{[\Xi]_\alpha}_{\equiv \sup_{y<x} \frac{|\Xi_{y,x}|}{(x-y)^\alpha}} + \sup_{y<r<x} \frac{|\Xi_{y,x} - \Xi_{y,r} - \Xi_{r,x}|}{|x - y|^\beta} < \infty.$$

For $\omega \in \Omega_\gamma$ and $W \in C^\delta([\underline{x}, \bar{x}], \mathbb{R})$, define $\Xi^{\omega,W}$ by

$$\Xi^{\omega,W}_{y,x} \;\equiv\; W_y \cdot (\omega_x - \omega_y).$$

Since $\left|\Xi^{\omega,W}_{y,x}\right| \leq [\omega]_\gamma\ |x - y|^\gamma$ and, for $\underline{x} \leq y \leq r \leq x \leq \bar{x}$,

$$\left| \Xi^{\omega,W}_{y,x} - \Xi^{\omega,W}_{y,r} - \Xi^{\omega,W}_{r,x} \right| = |(W_y - W_r) \cdot (\omega_x - \omega_r)| \leq [W]_\delta\, [\omega_\gamma\ |x - y|^{\gamma+\delta},$$

we have $\Xi^{\omega,W} \in C_2^{\gamma,\gamma+\delta}([\underline{x}, \bar{x}], \mathbb{R})$. Thus, because $\gamma \leq 1 < \gamma + \delta$, it follows from the Sewing Lemma of

Friz and Hairer (2020) that the limit of Riemann sums

$$\lim_{\substack{\max_k |x_{k+1}-x_k| \to 0 \\ \underline{x}=x_0<\cdots<x_n=\bar{x}}} \sum_{k=0}^{n-1} W_{x_k} \cdot [\omega_{x_{k+1}\wedge x} - \omega_{x_k \wedge x}]$$

exists. This proves the claim.

(ii) We will prove the continuity of the map[73]

$$(\omega, W) \mapsto \int_0^\bullet W_t \, d\omega_t, \quad (\Omega_\gamma, [\,\cdot\,]_\gamma) \times C^\delta([\underline{x}, \bar{x}]\mathbb{R}) \to C^\gamma([\underline{x}, \bar{x}], \mathbb{R}).$$

By the Sewing Lemma quoted in (i), it suffices to check the continuity of

$$(\omega, W) \mapsto \Xi^{\omega, W}, \quad (\Omega_\gamma, [\,\cdot\,]_\gamma) \times C^\delta([\underline{x}, \bar{x}], \mathbb{R}) \to C_2^{\gamma, \gamma+\delta}([\underline{x}, \bar{x}], \mathbb{R}).$$

Given ω, $\tilde{\omega} \in \Omega_\gamma$, W, $\tilde{W} \in C^\delta([\underline{x}, \bar{x}], \mathbb{R})$,

$$\begin{aligned}
&\left|(\Xi^{\omega,W}_{y,x} - \Xi^{\omega,W}_{y,r} - \Xi^{\omega,W}_{r,x}) - (\Xi^{\tilde{\omega},\tilde{W}}_{y,x} - \Xi^{\tilde{\omega},\tilde{W}}_{y,r} - \Xi^{\tilde{\omega},\tilde{W}}_{r,x})\right| \\
&\quad = \left|(W_y - W_r) \cdot (\omega_x - \omega_r) - (\tilde{W}_y - \tilde{W}_r) \cdot (\tilde{\omega}_x - \tilde{\omega}_r)\right| \\
&\quad \le [W - \tilde{W}]_\delta \, [\omega]_\gamma \; |x-y|^{\gamma+\delta} + [\tilde{W}]_\delta \, [\omega - \tilde{\omega}]_\gamma \; |x-y|^{\gamma+\delta}.
\end{aligned}$$

Therefore (below we use the Hölder norm $\|\cdot\|_W$ instead of the seminorm $[\,\cdot\,]_W$)

$$\begin{aligned}
\left|\Xi^{\omega,W}_{y,x} - \Xi^{\tilde{\omega},\tilde{W}}_{y,x}\right| &= \left|W_y \cdot (\omega_x - \omega_y) - \tilde{W}_y \cdot (\tilde{\omega}_x - \tilde{\omega}_y)\right| \\
&\le \|W\|_\infty \, [\omega]_\gamma \; |x-y|^\gamma + \|W - \tilde{W}\|_\delta \cdot (b^\delta \vee 1) \, [\tilde{\omega}_\gamma \; |x-y|^\gamma.
\end{aligned}$$

This proves the claim.

(iii) Let $\int_{\underline{x}}^x W_y \, dB_y$ be a given version of the Itô integral. By the continuity of W, the chosen version is a probability limit[74]

$$\int_{\underline{x}}^x W_y \, dB_y = \lim_{\substack{\max_k |x_{k+1}-x_k| \to 0 \\ a=x_0<\cdots<x_n=x}} \sum_{k=0}^{n-1} W_{x_k} \cdot [B_{x_{k+1}\wedge x} - B_{x_k \wedge x}] \quad \text{in } \mu_\gamma\text{-probability}.$$

Then one can pass to a subsequence of the (implicitly given) sequence of partitions such that, for

[73] This is a stronger property than that stated in Lemma A.1(ii).
[74] See Revuz and Yor (2013).

μ_γ-a.a. $\omega \in \Omega_\gamma$(where we use the same notation for the subsequence),

$$\left[\int_{\underline{x}}^{x} W_y \, dB_y\right](\omega) \equiv \lim_{\substack{\max_k |x_{k+1}-x_k| \to 0 \\ \underline{x}=x_0<\cdots<x_n=x}} \sum_{k=0}^{n-1} W_{x_k} \cdot [\omega_{x_{k+1}\wedge x} - \omega_{x_k \wedge x}] = \int_{\underline{x}}^{x} W_y \, d\omega_y.$$

This proves the claim.

□

A.2 Proof of Theorem 3.1

Claim A.1. *Under Assumption A.1, for each $W \in C^\delta([\underline{x}, \bar{x}], \mathbb{R})$, define a probability measure $\mathbb{P}_W$ on the measurable space $(\Omega_\gamma, \mathcal{F}_\gamma)$ via the Radon-Nikodym density*

$$\frac{d\mathbb{P}_W}{d\mu_\gamma} = e^{\int_{\underline{x}}^{\bar{x}} \frac{W_x}{\sigma^2(x)} \, d\omega_x - \frac{1}{2}\int_{\underline{x}}^{\bar{x}} \frac{W_x^2}{\sigma^2(x)} \, dx}, \tag{A.1}$$

where μ_γ is the canonical Wiener measure μ restricted to Ω_γ, and $\int_{\underline{x}}^{\bar{x}} \frac{W_x}{\sigma^2(x)} d\omega_x$ is the pathwise integral defined in Lemma A.1(i). Then the canonical process on $(\Omega_\gamma, \mathcal{F}_\gamma, \mathbb{P}_W)$ has the same law as

$$W_x dx + \sigma(x) dB_x$$

where (B_x) is a standard Brownian motion.

Proof. This follows immediately from Lemma A.1(iii). □

Proof of Theorem It suffices to prove (i), after which (ii) follows trivially. Consider the probability space $(C^\gamma([\underline{x}, \bar{x}], \mathbb{R}), \mathcal{G}, \mu_\gamma)$, where the σ-field $\mathcal{G}$ is the Borel σ-field given by the uniform norm, and μ_γ is such that the canonical process has the same law as $(\sigma(x)B_x)$ where (B_x) is a standard Brownian motion.

Under Assumption A.1(i), (iii), and (iv), we have by Lemma A.1(i) that the expression $\pi_1(ds, \omega; \widetilde{W})$ is well-defined for each $\omega \in C^\gamma([\underline{x}, \bar{x}], \mathbb{R})$. By Claim A.1, under the probability measure $\mathbb{P}_{\widetilde{W}(\cdot,s)}$ defined by the Radon-Nikodym density

$$\frac{d\mathbb{P}_{\widetilde{W}(\cdot,s)}}{d\mu_\gamma} = e^{\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)}{\sigma^2(x)} d\omega_x - \frac{1}{2}\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x,s)^2}{\sigma^2(x)} dx},$$

the canonical process on $C^\gamma([\underline{x}, \bar{x}], \mathbb{R})$ has the same law as $\widetilde{W}(x,s)dx + \sigma(x)dB_x$ where (B_x) is a standard

Brownian motion. Therefore the probability measure on $C^\gamma([\underline{x},\bar{x}],\mathbb{R})\times S$ defined by

$$\left(e^{\int_{\underline{x}}^{\bar{x}}\frac{\widetilde{W}(x,s)}{\sigma^2(x)}d\omega_x-\frac{1}{2}\int_{\underline{x}}^{\bar{x}}\frac{\widetilde{W}(x,s)^2}{\sigma^2(x)}dx}\cdot\mu^\gamma(d\omega)\right)\otimes\pi_0(ds)=e^{\int_{\underline{x}}^{\bar{x}}\frac{\widetilde{W}(x,s)}{\sigma^2(x)}d\omega_x-\frac{1}{2}\int_{\underline{x}}^{\bar{x}}\frac{\widetilde{W}(x,s)^2}{\sigma^2(x)}dx}\cdot\mu^\gamma(d\omega)\otimes\pi_0(ds)$$

correctly specifies the joint probability law of (ω,s) according to the market maker's conjecture $\widetilde{W}$.

Under Assumption A.1(ii), we have by Lemma A.1(ii) that $e^{\int_{\underline{x}}^{\bar{x}}\frac{\widetilde{W}(x,s)}{\sigma^2(x)}d\omega_x-\frac{1}{2}\int_{\underline{x}}^{\bar{x}}\frac{\widetilde{W}(x,s)^2}{\sigma^2(x)}dx}$ is jointly measurable in (ω,s). Therefore, an application of Fubini-Tonelli Theorem shows that $\pi_1(ds,\omega;\widetilde{W})$ specifies the probability measure on S conditional on ω, i.e. $\{\pi_1(ds,\omega;\widetilde{W})\}_{\omega\in C^\gamma([\underline{x},\bar{x}],\mathbb{R})}$ is the ω-disintegration of the family $\{\mathbb{P}_{\widetilde{W}(\,\cdot\,,s)}\}_{s\in S}$, where $\mathbb{P}_{\widetilde{W}(\,\cdot\,,s)}$ is the probability law of aggregate order flow conditional on s, according to market maker belief $\widetilde{W}(\,\cdot\,,\,\cdot\,)$ regarding the insider's trading strategy. This proves the theorem.

A.3 Proof of Theorem 4.2

To make the dependence of expected AD price $\overline{P}(x,W;\widetilde{W})$ (18) on W more explicit, we write

$$\begin{aligned}\overline{P}(x,W;\widetilde{W})&=\mathbb{E}^{\mathbb{P}_W}[P(x,\omega\,;\widetilde{W})]\\&=\mathbb{E}^{\mathbb{P}_W}[\int_S\eta(x,s)\pi_1(ds,\omega;\widetilde{W})]\\&=\mathbb{E}^{\mathbb{P}_W}[\int_S\eta(x,s)^{e^{\mathcal{I}(\omega,s;\widetilde{W})}\pi_0(ds)}/_{C(\omega;\widetilde{W})}]\\&=\mathbb{E}^{\mu_\gamma}[\int_S\eta(x,s)\ {}^{e^{\mathcal{I}(\omega,s;\widetilde{W})}e^{\int_{\underline{x}}^{\bar{x}}\frac{\widetilde{W}(x',s)W(x')}{\sigma^2(x')}dx'}\cdot\pi_0(ds)}/_{C'(\omega;\widetilde{W})}],\end{aligned}\tag{A.2}$$

where

$$\mathcal{I}(\omega,s;\widetilde{W})=\int_{\underline{x}}^{\bar{x}}\frac{\widetilde{W}(x,s)}{\sigma^2(x)}d\omega_x-\frac{1}{2}\int_{\underline{x}}^{\bar{x}}\frac{\widetilde{W}(x,s)^2}{\sigma^2(x)}dx$$

and $C'(\omega,W;\widetilde{W})=\int_S e^{\mathcal{I}(\omega,s';\widetilde{W})}e^{\int_{\underline{x}}^{\bar{x}}\frac{\widetilde{W}(x',s')W(x')}{\sigma^2(x')}dx'}\pi_0(ds')$. The equality (A.2) follows from the fact that the law of the canonical process $x\mapsto\omega_x$ under $\mathbb{P}_W$ is the same as the law of $x\mapsto\int_0^x W(x')dx'+\omega_x$ under μ_γ.

The Gâteaux derivative of the payoff functional $W\mapsto\int_{\underline{x}}^{\bar{x}}W(x)\eta(x,s)ds$ is trivially identified with $\eta(\,\cdot\,,s)$, and it suffices to consider the cost functional

$$J_c(W)=\int_{\underline{x}}^{\bar{x}}W(x)\cdot\overline{P}(x,W;\widetilde{W})dx.$$

Let $v\in C^\delta([\underline{x},\bar{x}],\mathbb{R})$ and define

$$f(\varepsilon)=J_c(W+\varepsilon v).$$

The Gâteaux derivative $dJ_c(W)$ of $J_c(W)$ evaluated at v can be computed by invoking the Dominated Convergence Theorem and differentiating under the integral signs:

$$\begin{aligned} dJ_c(W) &= f'(0) \\ &= \int_{\underline{x}}^{\bar{x}} v(x) \cdot \overline{P}(x, W; \widetilde{W}) dx + \int_{\underline{x}}^{\bar{x}} W(x) g(x) dx \end{aligned} \tag{A.3}$$

for some function $g(x)$. The first integral in (A.3) verifies the AD term $N_{AD}(W)(v)$ of Equation (20). It remains to show the price impact term $N_K(W)(v)$ of Equation (21) is the second integral in (A.3).

The function $g(x)$ is of the form $\mathbb{E}^{\mu_\gamma}[\psi(\omega; x, \widetilde{W})]$. $\psi(\omega; x, \widetilde{W})$ is the random variable on Ω_γ given by

$$\int_S \eta(x, s) \frac{C'(\omega; \widetilde{W}) l(s, \omega; \widetilde{W}) \int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x', s)}{\sigma^2(x')} v(x') dx' - l(s, \omega; \widetilde{W}) \int_S l(s', \omega; \widetilde{W}) \int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x', s')}{\sigma^2(x')} v(x') dx' \pi_0(ds')}{C'(\omega; \widetilde{W})^2} \pi_0(ds)$$

where

$$C'(\omega; \widetilde{W}) = \int_S e^{\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x', s') W(x')}{\sigma^2(x')} dx' + \cdots} \pi_0(ds') = \int_S l(s', \omega; \widetilde{W}) \pi_0(ds')$$

is the posterior normalization constant of the market maker's posterior under the probability measure μ_γ, and

$$l(s, \omega; \widetilde{W}) = e^{\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x', s) W(x')}{\sigma^2(x')} dx' + \cdots} \tag{A.4}$$

is the likelihood of ω under μ_γ. For clarity of notation, in (A.4) we have put "$\cdots$" for terms not relevant for this calculation. We note that

$$\frac{l(s, \omega; \widetilde{W}) \pi_0(ds)}{C'(\omega; \widetilde{W})}$$

is the market maker's posterior condition on ω, under μ_γ. It is clear that $\psi(\omega; x, \widetilde{W})$ is the difference between the posterior expectation of the product of $\eta(x, \cdot)$ and $\int_{\underline{x}}^{\bar{x}} \frac{\widetilde{W}(x', \cdot)}{\sigma^2(x')} v(x') dx'$ and the product of their posterior expectations. In other words, $\psi(\omega; x, \widetilde{W})$ is the posterior covariance between these two quantities. This verifies the price impact term $N_K(W)(v)$ of Equation (21) and proves the theorem.

A.4 Proof of Proposition 5.3

For aggregate order flow $\omega \in \Omega_\gamma$, define $B(\omega)$ to be the vector $(\int_{\underline{x}}^{\bar{x}} \eta(x, s_i) d\omega_x)_{i=1,\cdots,I} \in \mathbb{R}^I$. In terms of d and $\widetilde{D}$ (defined in (27) and (29), respectively), the general market maker posterior over signals

$\pi_1(ds, \omega; \widetilde{W})$ obtained in Theorem 3.1(i) can be written explicitly as the probability mass function

$$\pi_1(i, \omega, d; \widetilde{D}) = \frac{e^{(\tilde{d}^{(i)})^T \frac{\mathbf{M}(\eta)}{\sigma^2} d + (\tilde{d}^{(i)})^T \frac{B(\omega)}{\sigma^2} - \frac{1}{2}(\tilde{d}^{(i)})^T \frac{\mathbf{M}(\eta)}{\sigma^2} \tilde{d}^{(i)}}}{\sum_{j=1}^{I} e^{(\tilde{d}^{(j)})^T \frac{\mathbf{M}(\eta)}{\sigma^2} d + (\tilde{d}^{(j)})^T \frac{B(\omega)}{\sigma^2} - \frac{1}{2}(\tilde{d}^{(j)})^T \frac{\mathbf{M}(\eta)}{\sigma^2} \tilde{d}^{(j)}}}, \quad i = 1, \cdots, I \tag{A.5}$$

on S. For the expectation of $\big(\pi_1(i, \omega, d; \widetilde{D})\big)_{i=1,\cdots,I} \in \mathbb{R}^I$, we write

$$\bar{\pi}_1(d; \widetilde{D}) = \Big(\mathbb{E}^{\mu_\gamma}[\pi_1(s_i, \omega, d; \widetilde{D})]\Big)_{i=1,\cdots,I} \in \mathbb{R}^I, \tag{A.6}$$

where $\mathbb{E}^{\mu_\gamma}[\,\cdot\,]$ is taken with respect to distribution over possible order flows ω if the insider chooses d and the market maker holds belief $\widetilde{D}$. The insider's expected utility maximization problem (15), conditional on observing s_i, now takes the simple finite-dimensional form

$$\max_{d \in \mathbb{R}^I} e_i^T \mathbf{M}(\eta) d - \bar{\pi}_1(d; \widetilde{D})^T \mathbf{M}(\eta) d \equiv \max_{d \in \mathbb{R}^I} J(d; \widetilde{D}, i). \tag{A.7}$$

The Bayesian trading game between the insider and the market maker therefore reduces to one where the market maker's posterior belief is specified by (A.5), and the insider's problem is (A.7).

The change of basis transformation (28)

$$\hat{d} = \frac{1}{\sigma} \mathbf{M}^{\frac{1}{2}}(\eta) d$$

replaces $\mathbf{M}$, σ, and d in the market maker's posterior $\pi_1(i, \omega, d; \widetilde{D})$ of (A.5) by $\mathbf{I}$, 1, and $\hat{d}$, respectively. Under this change of basis, the random vector $\widehat{X} = \mathbf{M}^{-\frac{1}{2}} \frac{B(\omega)}{\sigma}$ can be re-written as $\int \xi_i(x) dZ_x$ where $\int \xi_i(x) \xi_j(x) dx = \delta_{ij}$ and (Z_x) is a standard Brownian motion. Therefore, by Itô isometry,

$$\widehat{X}_i \overset{i.i.d.}{\sim} \mathcal{N}(0, 1), \quad i = 1, \cdots, I. \tag{A.8}$$

This gives the market maker's posterior (30) in the canonical game.

The change of basis (28) replaces $\mathbf{M}$ and d in the objective function $J(d; \widetilde{D}, i)$ of (A.7) by $\mathbf{I}$ and $\hat{d}$, respectively. This gives the insider's problem (31) in the canonical game and proves the proposition.

A.5 Proof of Theorem 5.5

Lemma A.2. *Let $a' > 0$ be a solution to the equilibrium equation $\Phi(a) = 0$. In the canonical game, suppose the market maker holds belief $a'\mathbf{Q}$. Then the insider's objective function* (31) *conditional on each s_i is concave, which implies that first-order conditions are sufficient for optimality. Therefore, since the insider's strategy $a'\mathbf{Q}$ satisfies the joint first-order conditions $\Phi(a')\mathbf{Q} = 0$, $a'\mathbf{Q}$ is an equilibrium of* *the canonical game, as defined in* (33).

Proof. Given the market maker's belief $\widehat{D} = a'\mathbf{Q}$ in the canonical game, the Hessian matrix of the insider's objective function (31) conditional on s_i can be computed by differentiating directly the first-order condition (35), which gives

$$-a'\mathbf{Q}\,\underbrace{\mathbb{E}\left[\begin{bmatrix} p_1 & -p_1p_2 & \cdots \\ -p_1p_2 & p_2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}\right]}_{S} = -a'\mathbf{Q}S.$$

By the Cauchy-Schwarz inequality,

$$\mathbb{E}[p_ip_j]^2 \leq \mathbb{E}[p_i^2]\mathbb{E}[p_j^2] \leq \mathbb{E}[p_i]\mathbb{E}[p_j], \quad \forall 1 \leq i, j \leq I.$$

Therefore S is positive semidefinite. Since $\mathbf{Q}S = S\mathbf{Q}$, $\mathbf{Q}S$ is also positive semidefinite. Therefore the Hessian $-a'\mathbf{Q}S$ is negative semidefinite. It follows that the insider's objective function is concave. This proves the lemma. □

Lemma A.3. *Let*

$$\mathbf{p} = \begin{bmatrix} p_1 \\ \vdots \\ p_I \end{bmatrix}$$

be the market maker's posterior in the canonical game, conditional on the insider observing signal s_i. $\mathbf{p}$ *is a random probability measure on the signal space* $S = \{s_1, \cdots, s_I\}$. *Let* $X \overset{d}{\sim} \mathcal{N}(0, a^2\mathbf{Q})$ *be a random vector that is multivariate normal with mean* 0 *and covariance matrix* $a^2\mathbf{Q}$.

Under the equilibrium ansatz $a\mathbf{Q}$, *the probability law of* $\mathbf{p}$ *is given by*[75]

$$\mathbf{p} \overset{d}{\sim} \left(\frac{e^{X_1}}{\sum_j \cdots}, \cdots, \frac{e^{a^2+X_i}}{\sum_j \cdots}, \cdots \frac{e^{X_I}}{\sum_j \cdots} \right)^T.$$

Proof. Substituting the equilibrium ansatz $a\mathbf{Q}$ into the expression (A.5) for the market maker's posterior gives immediately that, conditional on signal s_i and order flow ω,

$$\pi_1(s_i, \omega, \beta^{(i)}; a\mathbf{Q}) = \frac{e^{a^2(1-\frac{1}{I}) + ae_i^T\mathbf{Q}\mathbf{M}^{-\frac{1}{2}}(\eta)\frac{B(\omega)}{\sigma}}}{\sum_j \cdots}, \tag{A.9}$$

[75] "$\sum_j \cdots$" is a random normalization constant so that $\sum_i p_i = 1$.

and, for $k \neq i$,

$$\pi_1(s_k, \omega, \beta^{(i)}; a\mathbf{Q}) = \frac{e^{a^2(-\frac{1}{I}) + a e_k^T \mathbf{Q} \mathbf{M}^{-\frac{1}{2}}(\eta)\frac{B(\omega)}{\sigma}}}{\sum_j \cdots}, \tag{A.10}$$

where $\sum_j \cdots$ is a random normalization constant.

The common factor $e^{a^2(-\frac{1}{I})}$ in (A.9) and (A.10) cancels after normalization. It remains to consider the random vector $\mathbf{M}^{-\frac{1}{2}}(\eta)\frac{B(\omega)}{\sigma}$. By the same Itô isometry argument as that for (A.8), $\mathbf{M}^{-\frac{1}{2}}(\eta)\frac{B(\omega)}{\sigma} \in \mathbb{R}^I$ is distributed multivariate standard normal. Therefore $a\mathbf{Q}\frac{B(\omega)}{\sigma} \overset{d}{\sim} \mathcal{N}(0, a^2\mathbf{Q})$. This proves the lemma. □

Proof of Theorem

<u>*Derivation of Equilibrium Equation* (39).</u>

Substituting the equilibrium ansatz $\widehat{D} = a\mathbf{Q}$ of (37) into the insider's first-order condition (35) for the canonical game conditional on s_1 gives

$$\mathbb{E}\left[e_1 - \begin{bmatrix} p_1 \\ \vdots \\ p_I \end{bmatrix} - a^2\mathbf{Q}\begin{bmatrix} p_1 & & \\ & \ddots & \\ & & p_I \end{bmatrix}\mathbf{Q}e_1 + a^2\mathbf{Q}\begin{bmatrix} p_1 \\ \vdots \\ p_I \end{bmatrix}\begin{bmatrix} p_1 & \cdots & p_I \end{bmatrix}\mathbf{Q}e_1\right] = 0, \tag{A.11}$$

where $\mathbb{E}[\,\cdot\,]$ is taken with respect to the probability law of the random probability measure $\mathbf{p} = \begin{bmatrix} p_1 \\ \vdots \\ p_I \end{bmatrix}$ under the conjectured equilibrium.

The first-order condition (A.11) conditional on s_1 simplifies to

$$\mathbb{E}\left[e_1 - \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_I \end{bmatrix} - a^2\left(\begin{bmatrix} p_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - p_1\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_I \end{bmatrix}\right)\right] = \mathbb{E}\left[(1 - a^2 p_1)\begin{bmatrix} 1 - p_1 \\ -p_2 \\ \vdots \\ -p_I \end{bmatrix}\right] = 0. \tag{A.12}$$

The probability law of $\mathbf{p}$ is as characterized in Lemma A.3 above, with $i = 1$. Under this probability law, the moments in Equation (A.12) are the moments of a logistic normal distribution. We now show that, by substituting for the appropriate relationships between corresponding moments, (A.12) can be written as

$$\Phi(a)\mathbf{Q}e_1 = 0, \tag{A.13}$$

for some $\Phi\colon [0, \infty) \to \mathbb{R}$.

To show (A.13), it suffices to show

$$\mathbb{E}\left[(1-a^2p_1)(1-p_1)\right] = (I-1)\mathbb{E}\left[(1-a^2p_1)p_2\right]. \tag{A.14}$$

Under the probability law of $\mathbf{p}$ obtained in Lemma A.3, (A.14) will imply

$$\mathbb{E}\left[(1-a^2p_1)(1-p_1)\right] = (I-1)\mathbb{E}\left[(1-a^2p_1)p_i\right], \quad i = 3, \cdots, I. \tag{A.15}$$

In turn, (A.14) holds if

$$\mathbb{E}\left[p_1(1-p_1)\right] = (I-1)\mathbb{E}\left[p_1p_2\right]. \tag{A.16}$$

(A.16) holds since

$$\mathbb{E}\left[p_1(1-p_1)\right] = \mathbb{E}\left[p_1(p_2 + \cdots + p_I)\right] = (I-1)\mathbb{E}\left[p_1p_2\right],$$

where $\mathbb{E}[p_1p_i] = \mathbb{E}[p_1p_j]$ for all $i, j \neq 1$ under the logistic-normal probability law of $\mathbf{p}$ given by Lemma A.3, conditional on s_1.

Therefore, in (A.13) we can take $\Phi(a)$ to be (up to a scalar multiple) the left-hand side of (A.15), i.e.

$$\Phi(a) = \mathbb{E}\left[1 - p_1 - a^2p_1 + ap_1^2\right]$$

where

$$p_1 = \frac{e^{a^2+X_1}}{e^{a^2+X_1} + \sum\limits_{i\neq 1} e^{X_i}}, \quad X = (X_i)_{1\leq i\leq I} \overset{d}{\sim} \mathcal{N}(0, a^2\mathbf{Q}).$$

By symmetry, the first-order condition for $i \neq 1$ is identical to (A.13) after permuting the indices 1 and i. Stacking the I symmetric first-order conditions,

$$\Phi(a)\mathbf{Q}e_i = 0, \ 1 \leq i \leq I,$$

side-by-side gives the matrix equation $\Phi(a)\mathbf{Q} = 0$ of (38). By Lemma A.2, an equilibrium is given by a solution $a^* > 0$ to $\Phi(a) = 0$, which is the equilibrium equation of (39).

Existence of Equilibrium.

We have $\Phi(0) > 0$. Moreover, $1 - p_1 - a^2p_1 + ap_1^2 \to 0$ and $1 - p_1 - a^2p_1 + ap_1^2 < 0$ eventually almost surely as $a \to \infty$. By Fatou's Lemma, $\Phi(a) \to 0$ from below as $a \to \infty$. By the Intermediate Value Theorem, there exists $a^* > 0$ such that $\Phi(a^*) = 0$. This proves (i), and (ii) follows immediately from Proposition 5.4.

A.6 Proof of Proposition 6.9

This is an immediate corollary of Lemma A.3 in Proof of Theorem 5.5, with $a = a^*$.

References

Aït-Sahalia, Y. and R. Kimmel (2007). Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics 83*(2), 413–452.

Aït-Sahalia, Y., C. Li, and C. X. Li (2021). Implied stochastic volatility models. *Review of Financial Studies 34*(1), 394–450.

Augustin, P., M. Brenner, and M. G. Subrahmanyam (2019). Informed options trading prior to takeover announcements: Insider trading? *Management Science 65*(12), 5697–5720.

Back, K. (1993). Asymmetric information and options. *Review of Financial Studies 6*(3), 435–472.

Back, K., C. H. Cao, and G. A. Willard (2000). Imperfect competition among informed traders. *Journal of Finance 55*(5), 2117–2155.

Bali, T. G. and S. Murray (2013). Does risk-neutral skewness predict the cross section of equity option portfolio returns? *Journal of Financial and Quantitative Analysis 48*(4), 1145–1171.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. *Review of Financial Studies 9*(1), 69–107.

Bates, D. S. (2000). Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics 94*(1-2), 181–238.

Berestycki, H., J. Busca, and I. Florent (2002). Asymptotics and calibration of local volatility models. *Quantitative Finance 2*(1), 61.

Biais, B. and P. Hillion (1994). Insider and liquidity trading in stock and options markets. *Review of Financial Studies 7*(4), 743–780.

Breeden, D. T. and R. H. Litzenberger (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business 51*(4), 621–651.

Britten-Jones, M. and A. Neuberger (2000). Option prices, implied price processes, and stochastic volatility. *Journal of Finance 55*(2), 839–866.

Caballe, J. and M. Krishnan (1994). Imperfect competition in a multi-security market with risk neutrality. *Econometrica 62*(3), 695–704.

Cao, C., Z. Chen, and J. M. Griffin (2005). Informational content of option volume prior to takeovers. *Journal of Business 78*(3), 1073–1109.

Cao, J. and B. Han (2013). Cross section of option returns and idiosyncratic stock volatility. *Journal of Financial Economics 108*(1), 231–249.

Carr, P. and L. Cousot (2012). Explicit constructions of martingales calibrated to given implied volatility smiles. *SIAM Journal on Financial Mathematics 3*(1), 182–214.

Chabakauri, G., K. Yuan, and K. E. Zachariadis (2022). Multi-asset noisy rational expectations equilibrium with contingent claims. *Review of Economic Studies 89*(5), 2445–2490.

Christoffersen, P., R. Goyenko, K. Jacobs, and M. Karoui (2018). Illiquidity premia in the equity options market. *Review of Financial Studies 31*(3), 811–851.

Christoffersen, P., K. Jacobs, and K. Mimouni (2010). Volatility dynamics for the S&P500: Evidence from realized volatility, daily returns, and option prices. *Review of Financial Studies 23*(8), 3141–3189.

Collin-Dufresne, P. and V. Fos (2016). Insider trading, stochastic liquidity, and equilibrium prices. *Econometrica 84*(4), 1441–1475.

Collin-Dufresne, P., V. Fos, and D. Muravyev (2021). Informed trading in the stock market and option-price discovery. *Journal of Financial and Quantitative Analysis 56*(6), 1945–1984.

Coval, J. D. and T. Shumway (2001). Expected option returns. *Journal of Finance 56*(3), 983–1009.

Driessen, J., P. J. Maenhout, and G. Vilkov (2009). The price of correlation risk: Evidence from equity options. *Journal of Finance 64*(3), 1377–1406.

Duffie, D., J. Pan, and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica 68*(6), 1343–1376.

Easley, D., M. O'Hara, and P. S. Srinivas (1998). Option volume and stock prices: Evidence on where informed traders trade. *Journal of Finance 53*(2), 431–465.

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

Friz, P. K. and M. Hairer (2020). *A Course on Rough Paths*. Springer.

Glosten, L. R. and L. E. Harris (1988). Estimating the components of the bid/ask spread. *Journal of Financial Economics 21*(1), 123–142.

Goyal, A. and A. Saretto (2009). Cross-section of option returns and volatility. *Journal of Financial Economics 94*(2), 310–326.

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics 92*(2), 153–181.

Hairer, M. (2014). Singular stochastic PDEs. *Proceedings of the International Congress of Mathematicians*.

Hendershott, T., C. M. Jones, and A. J. Menkveld (2011). Does algorithmic trading improve liquidity? *Journal of Finance 66*(1), 1–33.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6*(2), 327–343.

Huang, R. D. and H. R. Stoll (1997). The components of the bid-ask spread: A general approach. *Review of Financial Studies 10*(4), 995–1034.

Karatzas, I. and S. Shreve (2012). *Brownian Motion and Stochastic Calculus*. Springer.

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

Lin, J.-C., G. C. Sanger, and G. G. Booth (1995). Trade size and components of the bid-ask spread. *Review of Financial Studies 8*(4), 1153–1183.

Lyons, T. J., M. Caruana, and T. Lévy (2007). *Differential Equations Driven by Rough Paths*. Springer.

Makarov, I. and A. Schoar (2020). Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics 135*(2), 293–319.

Ni, S. X., J. Pan, and A. M. Poteshman (2008). Volatility information trading in the option market. *Journal of Finance 63*(3), 1059–1091.

Pan, J. and A. M. Poteshman (2006). The information in option volume for future stock prices. *Review of Financial Studies 19*(3), 871–908.

Revuz, D. and M. Yor (2013). *Continuous Martingales and Brownian Motion*. Springer.

Rochet, J.-C. and J.-L. Vila (1994). Insider trading without normality. *Review of Economic Studies 61*(1), 131–152.

Roll, R., E. Schwartz, and A. Subrahmanyam (2009). Options trading activity and firm valuation. *Journal of Financial Economics 94*(3), 345–360.

Zhan, X., B. Han, J. Cao, and Q. Tong (2022). Option return predictability. *Review of Financial Studies 35*(3), 1394–1442.