Eliciting the private signal distribution from option prices

Julio A. Crego[*]

June 1, 2024

Abstract

I provide a theoretical framework that characterizes which option strike an informed agent buys or sells after a given signal. The informed agent faces a trade-off between higher exposure to the asset or a more favorable price for the option. In equilibrium, he implements a mixed strategy across strikes to camouflage himself as a noise trader. However, he only considers strikes within a segment of the strike line. This segment depends on the realization of the private signal. As a result, there is a one-to-one mapping between the asset distribution conditional on each possible signal realization and the price slope. Additionally, the model suggests that market makers can make the information asymmetry losses of noise traders independent of the private signal realization.

Keywords: Information asymmetries, options, return density.

JEL: G14, G23.

[*] *Tilburg University. E-mail: jacrego@uvt.nl, postal address: Warandelaan 2 5037AB, Tilburg.*

Easley et al. (1998) provide a model that rationalizes why informed agents should trade on the options market due to the intrinsic leverage of options. Their model characterizes the the informed agent decision of trading the asset or an available option when the market maker sets ask and bid prices in both markets endogenously. Nonetheless, their paper remains silent about which strike an informed investor trades if multiple strikes exist. In this paper, I consider the problem of an informed agent who has decided to trade in the options market and needs to optimally select the strike. At the same time, market makers in the option market set prices endogenously to cover the potential losses from adverse selection.

The model equilibrium presents three important features. First, informed agents use mixed strategies to camouflage as noise traders if the number of available strikes is high. Second, they trade completely different strikes after each signal realization. Consequently, the slope of option prices in different points of the strike line elicits the posterior distribution of informed agents after each signal realization. Third, multiple price levels sustain an equilibrium. Market makers might decide to set prices such that noise traders will lose more after very strong signal realizations or lose the same regardless of the private signal realization. This result suggests that the level of option bid-ask spreads does not contain much information about information asymmetries even if the slope does.

The model considers a call option market with many strikes à la Glosten and Milgrom (1985). Each option has an ask and a bid price set by a competitive market maker. On the other side of the transaction, nature selects whether the agent is informed or a noise trader. The latter trades a strike randomly. The former receives an imperfect private signal about the final value of the asset and decides whether to buy or sell and which strike. After observing the signal about the asset value, the informed trader faces a trade-off. He would like to get a higher exposure to the asset as in Easley et al. (1998) to obtain a higher benefit from his information. However, in equilibrium, strikes that offer a high exposure also have less favorable prices because the market maker internalizes that the order flow contains more information.

The optimal strategy for an informed agent consists of dividing the strike line into different segments. He buys an option in the segment that provides the higher exposure after the most positive signal. Nonetheless, within the segment, he plays mixed strategies and trades each strike with a positive probability. The market maker anticipates this behavior and sets a higher spread than for other strikes.[1] The informed agent buys a strike in the second segment with the highest exposure after the second most positive signal and so on. There are no profitable deviations because trading a strike with higher exposure is more expensive, while trading one with a lower spread leads to lower exposure; hence, lower information gains.

The market maker is competitive; therefore, prices reflect the equilibrium strategy of informed agents. I focus on the ask side of the book since both sides are symmetric. Since the informed agent separates the strike line into segments and chooses the segment based on the realization of the signal, prices also follow a segment structure. The informed agent trades the strikes with the highest exposure only after the most positive signal; therefore, these strikes have the highest spread to compensate for the adverse selection. Within the strike segment, not every strike is the same. Low strikes provide more exposure to the asset. Consequently, they are traded more by informed agents and the market maker sets wider spreads.

Contrary to most information models, market makers can have an active role. They can set the price segments wider with a lower spread or narrower with a higher spread. In both cases, they get zero profits. The difference is how much risk they eliminate from informed investors. These investors face risk before the signal realization as this signal might be good, bad, very good, very bad, etc. If the market maker selects a specific price function, any realization of the private signal leads to the same profits for the informed investors; hence, they face no risk, and the adverse selection cost is the same for each strike.

Regardless of the decision by the market maker, in equilibrium, the slope of the ask price with respect to the strike within a given segment must make informed agents indifferent to sustain a mix-strategy equilibrium. Therefore, the slope equals the loss due to a lower exposure, which boils down to the derivative of a conditional expectation. Consequently, analogous to Breeden and Litzenberger (1978), the slope of prices corresponds to a point of the cumulative density function of the asset liquidation value conditional on the private signal realization after which informed agents trade that segment. Since we have several strikes per segment, we can recover the conditional density after each signal for a wide range of distributions. This result arises from the risk neutrality of the informed investor and a high number of strikes.

This paper is the first to consider the optimal choice of the strike by informed investors and

[1] I define spread as the absolute difference between quoted price and the unconditional value.

endogenous prices at the same time. Augustin et al. (2023) considers the selection of strikes but disregards the change in prices due to the presence of informed investors. In contrast, my model focuses on characterizing the price function to elicit the possible information held by informed agents.

This paper also contributes to the literature that links option volume to future stock returns. The more popular measure to predict stock returns is the put-call ratio; that is, the put total volume over the call total volume. Pan and Poteshman (2006) show this measure outperforms any other measure based on prices. Bondarenko and Murayev links this predictability to insider trading by showing that the predictability dissipated after the SEC tightened the insider regulation following the arrest of Raj Rajaratnam. My paper adds the strike dimension to the problem. My model predicts that in-the-money volume leads to more extreme price changes than out-of-the-money volume. Nonetheless, it is silent about average predictability because it depends on the likelihood of each signal.

My model provides the theoretical result to elicit the distribution of private information. This distribution constitutes one of the main ingredients in learning models. For instance, Easley et al. (1996) considers a model in which the informed agent knows the asset with certainty. Then, informed agents trade according to their signal, and the market maker learns quickly. Cipriani and Guarino (2008) modify the model by considering an imperfect private signal with a particular distribution. This small modification leads to herding times during which informed agents disregard their own signal and follow publicly observed trades. My model provides a framework to differentiate empirically both models. If private information is almost certain, in-the-money options will present a steep slope. Instead, if private information is very noisy as Cipriani and Guarino (2008)'s calibration suggests, the slope should be flatter and almost constant across strikes.

Finally, my model adds to the vast literature that measures information asymmetries in different markets (See Ahern, 2020, for an evaluation of these metrics). Most of this literature focuses on measuring the average adverse selection cost in equities or bonds. My results provide guidance on how to extend these measures beyond the average adverse selection cost. For instance, investors might be willing to lose 0.1% of their investment with 50% probability and 0 otherwise; however, they might not accept losing 1% with 5% probability. The main insight of the model is that steep in-the-money option prices indicate the second case is more likely than the first one. Likewise, eliciting all the possible posteriors of informed agents allows market makers to set up spreads in related markets (e.g. CDS).

1 Stylized Model

In this section, I present a stylized model that captures the main insights and allows me to illustrate the equilibrium. The following sections discuss why the main results hold in more general set-ups.

Assets. There is an asset whose value at time T, θ, might be θ_U with probability α_0 or θ_L with probability $1-\alpha_0$. This asset is not directly traded but there are N different European call options written on the asset.[2] All of them expire at T. Options differ on their strikes, which range from K_L to K_U, such that $\theta_L \leq K_L < K_U \leq \theta_U$. For simplicity, and in line with usual markets, I consider equally-spaced strikes: $K_n = K_L + \frac{n-1}{N-1}(K_U - K_L)\ \ n \leq N\ n \in \mathbb{N}$.

Players. There is market maker who sets prices competitively. Analogous to Glosten and Milgrom (1985) model, there is a mass δ of risk-neutral informed agents who observe a signal s about the value at maturity and can buy or sell a single call option. The novelty is that they decide the strike. Despite the presence of informed agents, trade occurs because there is a mass $1-\delta$ of noise traders who buy and sell call options with probability ε and $1-\varepsilon$, respectively; and they select each strike with equal probability. To make the exposition clearer, I refer to the market maker as *she* and to an informed or noise trader as *he*.

Timing and strategies. Since this paper focuses on characterizing the spread across strikes, I consider a static model and assume a zero discount rate. At the beginning of the period, market makers submit limit orders to form the ask, $A(k)$, and bid, $B(k)$, of each option. Then, a random trader enters the market and trades. If this agent is an informed agent, he decides to buy or to sell, and chooses the strike. I allow mixed strategies; therefore, his strategies are defined in terms of probabilities pinned down in equilibrium. In particular, he buys (sells) a call with strike k with probability $q(k,s)$ $(p(k,s))$.

Private Signal. After observing the signal $s \in \mathcal{S} = \{s_1, \ldots, s_M\}$ $(1 < M < N)$, the informed agent posterior probability of the highest state is α_s. These signals reflect different

[2] Easley et al. (1998) discuss the trade-off between trading the asset or an option.

information acquisition processes such as knowledge from being an executive in the firm to a competitive advantage in reading accounting information; as well as different outcomes of these processes. Consider the case of an accounting expert who finds accounting fraud with probability one; if she finds fraud, the information gain becomes extreme; however, if she does not the information gain is minimal. This two outcomes, *fraud/not fraud* would constitute two different signals. The objective is to recover the posterior after those signals, but we cannot recover the nature of those; hence, I follow a general notation for the signals. Precisely, I assume $\alpha_s > \alpha_t \iff s > t$, and $\alpha_s > \alpha_0 \iff s > 0$ without loss of generality. Finally, I denote as $g(s)$ the probability mass function of the signals.

In order to avoid discussing corner cases, I assume that informed traders are for sure informed ($s \neq 0 \ \forall s$). Although Easley and O'hara (1987) show that relaxing this assumption leads to interesting learning patterns and Easley et al. (1996) proves its relevance to describe the empirical probability of informed trading, the aim of this model is to provide a characterization of the cross-section of strikes; hence, the time dimension, thus learning, plays a minor role. Actually, in this model, adding uninformative signals is isomorphic to increase the proportion of noise traders.[3]

Extra notation. I denote as Inf the fact that an informed is selected from the pool of traders, and as $d_k = 1$ ($d_k = -1$) the fact that someone buys (sells) an option with strike k from the market maker. Despite entailing extra notation, I use ΔK to refer to the difference between two consecutive strikes ($\Delta K = \frac{K_U - K_L}{N}$) and the set of all strikes $\mathcal{K}$. I denote the equilibrium functions with an asterisk. To save space I refer to the call payout $max\{\theta - K, 0\}$ as $(\theta - K)^+$.

Equilibrium. Similar to Glosten and Milgrom (1985) and Easley and O'hara (1987), I define an equilibrium in this model as a set of functions $\{q^*(k,s), p^*(k,s), A^*(k), B^*(k)\}$ such that:

1. The market maker makes zero profits from each strike: $\pi^A(A^*(k),k) = 0$, $\pi^B(B^*(k),k) = 0$, where π^A and π^B denote the expected profits of the market maker from strike k on the ask and bid side respectively:

$$\pi^A(A(k),k) = A(k) - \mathbb{E}((\theta-k)^+|d_k = 1) \quad \pi^B(B(k),k) = \mathbb{E}((\theta-k)^+|d_k = -1) - B(k) \ \forall k \in \mathcal{K}$$

[3] This assertion follows from assuming that informed agents when they do not receive a signal trade randomly mimicking noise traders as in Easley and O'hara (1987), Easley et al. (1996), and Cipriani and Guarino (2014).

2. The informed agent maximizes profits:

$$\{q^*(k,s), p^*(k,s)\} = \arg\max_{q,p} \Pi(q,p,s) \forall s \in \mathcal{S}, k \in \mathcal{K}.$$

$$\text{subject to} \sum_{k=K_L}^{K_N} p(k,s) + q(k,s) \leq 1 \ \forall s \text{ and } q(k,s) \geq 0, \ \ p(k,s) \geq 0$$

$$\Pi(q,p,s) = \sum_{k=K_L}^{K_N} q(k,s)\Big(\mathbb{E}\big((\theta-k)^+|s\big) - A^*(k)\Big) + \sum_{k=K_L}^{K_N} p(k,s)\Big(B^*(k) - \big(\mathbb{E}\big((\theta-k)^+|s\big)\Big)$$

3. Informed investors participate after each signal:

$$\Pi(q,p,s) > 0$$

Condition 1 represent the perfect competition of market makers in reduced form. Avoiding to explicitly model competition among market makers simplifies the model as the market maker becomes a passive player. Hence, it is enough to characterize the equilibrium strategy of informed investors at equilibrium prices, which reduces the dimensionality of the functions $q(\cdot)$ and $p(\cdot)$. Since there is a one-to-one mapping from strikes to (equilibrium) prices, I omit prices as arguments of the informed agent's strategies. Condition 2 borrows the informed agent's profit function from the seminal papers and aggregates those profits across strikes.

Condition 3 trivially holds, even out of equilibrium, in any paper with risk-neutral agents and binary state and signal space. Extending the signal space might result in some signals after which the informed investor does not find optimal to trade and leaves the market. However, since they do not participate, they do not affect prices. Hence, I focus the model on full participation and the posterior distributions (α_s) should be interpreted as those of participating investors.

1.1 Road to equilibrium

The complexity of this model resides in its flexibility. By allowing several strikes and signals, the equilibrium consists of several functions. Therefore, to obtain the equilibrium, we first need to reduce the possible functions. I do so through a sequence of lemmas that isolate each of the main moving pieces of the model.

Lemma 1. ***Market side separation.*** *Given a signal s, the informed agent either sells or buys call options. Mathematically,*

$$\sum_{k=K_L}^{K_U} q^*(k,s) = 1 \wedge p^*(k,s) = 0 \ \ \forall k \in \mathcal{K} \ \textit{or} \ \sum_{k=K_L}^{K_U} p^*(k,s) = 1 \wedge q^*(k,s) = 0 \ \ \forall k \in \mathcal{K}$$

Lemma 1 indicates that there are some signals after which the informed agent always buys while after others he sells or leaves the market without trading. Therefore, mixed strategies in which informed agents buy some strikes and sell others cannot constitute an equilibrium. This result eases significantly the problem as it allows us to solve for the equilibrium in the ask and bid side of the market separately. This separation occurs in most seminal models, such as Easley and O'hara (1987). Although in these cases, informed traders are either buyers or sellers at every point in time, instead of at every strike. Similar to those models, this result arises from the clear order of the signals implied by the binary distribution. If the posterior probability of the upper state is greater than α_0, the expected value of buying any call option is higher than the expected value of selling any of them.

Using Lemma 1 we can separate the informed agent decision in two steps. First, he decides to buy ($\alpha_s > \alpha_0$) or sell. Then he chooses the strike. When taking this last decision he faces a trade-off. On the one hand, lower strikes provide a higher exposure to the asset; hence, a higher expected payoff. On the other hand, in equilibrium, the market maker understands that lower prices entail higher adverse selection costs and increases the spread for those call options.

Lemma 2. ***No gaps.*** *In equilibrium, if the informed agent trades strike k' with positive probability, he trades any strike $k < k'$ with positive probability. Mathematically,*

$$\exists s \in \mathcal{S} \ s.t. \ q^*(k', s) > 0 \Rightarrow \exists s' \in \mathcal{S} \ s.t. \ q^*(k, s') > 0 \ \ \forall k < k' \ k, k' \in \mathcal{K}$$

$$\exists s \in \mathcal{S} \ s.t. \ p^*(k', s) > 0 \Rightarrow \exists s' \in \mathcal{S} \ s.t. \ p^*(k, s') > 0 \ \ \forall k < k' \ k, k' \in \mathcal{K}$$

Lemma 2 formalizes a clear implication of the previous trade-off: low strikes are always traded by informed traders. Consider a potential equilibrium in which informed traders do not buy strike k' then, the ask price equals the unconditional value due to competition. This price constitutes a great opportunity for informed agents, who might deviate and earn higher rents. If, in this potential equilibrium, they trade a higher strike, they would deviate for sure; hence it is not an equilibrium. However, if they trade a lower strike, they might decide not to deviate because the benefit of a higher exposure to the asset offsets the benefit of a lower price. Therefore, this model might provide options with no spread in equilibrium, but these options will be those with highest strikes.

Corollary 1. ***Underlying.*** *In equilibrium, informed agents always trade the lowest strike. Mathematically,*

$$\exists s \in \mathcal{S} \ s.t. \ q^*(K_L, s) > 0 \ and \ \exists s \in \mathcal{S} \ s.t. \ p^*(K_L, s) > 0$$

Corollary 1 to Lemma 2 indicates that the lowest strike always carries adverse selection costs. In the extreme case in which the asset is traded ($K_L = 0$), the corollary implies the informed agent trades the asset with positive probability. This result hinges on the signal being about the level of the liquidation value and therefore arises in many previous models as Easley et al. (1998).

Lemma 3. ***Signal order.*** *In equilibrium, if informed agents trade strike k' with positive probability after signal s', they do not trade any higher strike after a stronger signal. Mathematically,*

$$\exists s' \ s.t. \ q^*(k', s') > 0 \Rightarrow q^*(k, s) = 0 \ \ \forall k > k', s > s', \ k, k' \in \mathcal{K}, \ s, s' \in \mathcal{S}$$

$$\exists s' \ s.t. \ p^*(k', s') > 0 \Rightarrow p^*(k, s) = 0 \ \ \forall k > k', s < s', \ k, k' \in \mathcal{K}, \ s, s' \in \mathcal{S}$$

The benefit of a higher exposure to the asset depends on the information owned by the informed trader. After a signal with a very high posterior expectation, the marginal benefit of increasing the exposure is higher than after a signal which generates a posterior close to the unconditional value. This mechanism constitutes the basis of Lemma 3 which sets an order for the signals. It specifies that as the strike increases, the posterior expectation of the informed agent who trades decreases. This result reduces significantly the set of potential equilibria and makes the problem feasible.

Lemma 4. ***Signal separation.*** *In equilibrium, if informed agents play mixed strategies across some strikes after signal s, they do not play mixed strategies across those strikes after any other signal. Mathematically,*

$$\exists s \in \mathcal{S} \ s.t. \ q^*(k, s) \in (0,1) \Rightarrow \nexists s' \in \mathcal{S} \ s.t. \ q^*(k, s') \in (0,1)$$

$$\exists s \in \mathcal{S} \ s.t. \ p^*(k, s) \in (0,1) \Rightarrow \nexists s' \in \mathcal{S} \ s.t. \ p^*(k, s') \in (0,1)$$

The different marginal benefit of exposure according to the signal has important implications for mixed strategy equilibria. In this type of equilibria, after a signal realization, the informed agent trades different strikes randomly, which allows the informed agent to camouflage as a noise trader. An equilibrium of this type requires that after the signal, the informed agent obtains the same profit across all the strikes at which he trades with a positive probability. Otherwise, he would deviate and trade the strike that provides the highest profit with probability one. This indifference condition generates a separation of strikes according to the signal. Therefore, in a mixed strategy equilibria some strikes are traded by informed agents after one signal but not after any other signal.

Lemma 5. ***Price slope.*** *In equilibrium, if informed agents play mixed strategies across some strikes, the slope of the price function of those strikes reveals the posterior probability. Mathematically,*

$$\exists s \ s.t. \ q^*(k,s) \in (0,1) \ \ and \ \sum_{i=K_L}^{k} q^*(k,s) < 1 \Rightarrow \alpha_s = -\frac{A(k+\Delta K) - A(k)}{\Delta K}$$

$$\exists s \ s.t. \ p^*(k,s) \in (0,1) \ \ and \ \sum_{i=K_L}^{k} p^*(k,s) < 1 \Rightarrow \alpha_s = -\frac{B(k+\Delta K) - B(k)}{\Delta K}$$

Since the informed trader only trades a particular strike after a specific signal, the market maker knows for sure the posterior distribution of the informed agent if he trades. Therefore, the separation of signals transfers to the pricing function. Lemma 5 states that the slope of the pricing function provides us with enough information to recover the complete set of posterior distributions with which informed traders might end up after the signal realization. Very steep pricing functions indicate that informed traders own very precise information. Similarly, if the bid pricing function is steeper than the ask pricing function, it implies that informed agents are better informed about downward movements than about upward ones.

1.2 Equilibrium with a continuum of strikes

The previous lemmas hold for any number of strikes, but solving for the equilibrium in this general case implies solving dozens of cases unless we restrict to very few strikes. However, in the data, we observe a large number of strikes and, actually, most econometric methods exploit the high number of strikes to make inference (see Figlewski, 2018, for a review). Therefore, instead of restricting to few strikes, I solve the model in the limit when the number of strikes tends to infinity.

Lemma 6. ***Only mixed strategies.*** *If $\Delta K \to 0$; then, only mixed strategies can constitute an equilibrium. Mathematically, in the limit,*

$$q^*(k,s) = \lim_{\Delta K \to 0} \beta(k,s)\Delta K$$

$$p^*(k,s) = \lim_{\Delta K \to 0} \sigma(k,s)\Delta K$$

and $0 < \lim_{\Delta K \to 0} \beta(k,s) < \infty$ *or* $\beta(k,s) = 0$*, and* $0 < \lim_{\Delta K \to 0} \sigma(k,s) < \infty$ *or* $\sigma(k,s) = 0$.

This last lemma ensures that, in the limit, informed agents mix across strikes regardless of the signal. Otherwise, informed trading would result obvious to the market maker since the probability of noise traders reduces to a differential. Note that, even in the nonlimiting case, mixing helps the informed trader to disguise; however, this is at the expense of a lower exposure. Hence, if some of the signals provides extreme information rents and the space between strikes is high, the informed investor might decide to buy one strike with probability one. This is the reason why Lemma 6 is unique to the limiting case although approximates the discrete case.

Lemma 6 also specifies that the informed agent trades with probabilities proportional to the difference between strikes. If he trades any strike with a higher probability, the market maker would set the spread to eliminate the adverse selection loss since the benefit obtained by noise traders would be negligible. Instead, if the informed agent decides to mix across more strikes and trade with intensity lower than proportional to ΔK, then some strikes would not be traded in equilibrium and the informed agent would benefit from trading those strikes.

Since the equilibrium must be in mixed strategies, we can use the lemmas in the previous section to depict the possible equilibrium. Lemma 4 implies that informed investors buy or sell a given strike after one specific signal, which I denote as s_k^*[4]. Meanwhile, Lemma 2 and Lemma 3 ensure that all the strikes traded after a signal must be consecutive. Hence, equilibrium prices consists of M different segments and the market maker faces stronger signals in the lowest-strike segments. Lemma 5 characterizes the behavior of prices and the strategy of informed traders in a possible equilibrium through a differential equation. Solving the differential equation, we arrive to the following equilibrium response at the ask side:

$$\beta(k, s_k^*) = \frac{1}{\delta g(s_k^*)} \left(\lambda(s_k^*)(\theta_U - k) - PN\right) \text{ if } s_k^* > 0 \text{ and } \beta(k,s) = 0 \; \forall s \neq s_k^* \tag{1}$$

[4] There are two s_k^* per strike, a negative and a positive one after which informed traders sell or buy the strike

where $\lambda(s_k^*)$ is a positive constant, and where PN is a measure of noise trading:

$$PN = \frac{(1-\delta)\varepsilon}{(K_U - K_L)} \text{ if } s_k^* > 0; PN = \frac{(1-\delta)(1-\varepsilon)}{(K_U - K_L)} \text{ if } s_k^* < 0$$

The symmetric case holds for the bid side of the market:

$$\sigma(k, s_k^*) = \frac{1}{\delta g(s_k^*)} \left(\lambda(s_k^*)(\theta_U - k) - PN\right) \text{ if } s_k^* < 0 \text{ and } \beta(k, s_k^*) = 0 \; \forall s \neq s_k^* \tag{2}$$

Using these expressions, we can find the prices through the zero profit condition:

$$A^*(k) = \alpha_{s_k^*}(\theta_U - k) - \frac{PN}{\lambda(s_k^*)}(\alpha_{s_k^*} - \alpha_0) \text{ if } \exists s_k^* > 0 \text{ s.t } \beta(k, s_k^*) > 0 \tag{3}$$

$$\text{and } B^*(k) = \alpha_{s_k^*}(\theta_U - k) + \frac{PN}{\lambda(s_k^*)}(\alpha_0 - \alpha_{s_k^*}) \text{ if } \exists s_k^* < 0 \text{ s.t } \sigma(k, s_k^*) > 0 \tag{4}$$

Consistent with the intuition that an informed trader wants exposure to the asset, he trades more intensively call options with lower strikes. As a result, they totally offset the extra information rents, and the market maker sets the same slope for all options within a segment. While the slope does not depend on the mass of noise traders, the level of prices does, as we observe in the first term of the second summand. In line with previous nonstrategic models, the mass of noise traders interacts with the size of information rents. Therefore, for a given s_k^*, the higher the likelihood of informed trading, the lower the deviation of prices from the posterior mean value, and this effect is amplified for extreme signals.

At this point, we know that any equilibrium includes prices with the same slope within a segment and we know the intensity of informed trading except for its slope. Next, we need to characterize the relative width of every segment. To do that, I introduce more notation and denote as $\underline{K}(s_k^*)$ and $\overline{K}(s_k^*)$ the limits of the segment. Lemma 1 states that the law of total probability constraint always binds in equilibrium. Hence, the width should be such that this restriction binds:

$$\int_{\underline{K}(s_k^*)}^{\overline{K}(s_k^*)} \beta(k, s_k^*) dk = 1, \; s_k^* > 0 \text{ and } \int_{\underline{K}(s_k^*)}^{\overline{K}(s_k^*)} \sigma(k, s_k^*) dk = 1, \; s_k^* < 0$$

which implies that the upper strike is equal to:

$$\overline{K}(s_i^*) = \left(\theta_U - \frac{PN}{\lambda(s_i^*)}\right) - \sqrt{\left(\left(\theta_U - \frac{PN}{\lambda(s_i^*)}\right) - \underline{K}(s_i^*)\right)^2 - 2\frac{\delta g(s_i^*)}{\lambda(s_i^*)}} \tag{5}$$

According to Lemma 2, the upper strike of one segment and the lower strike of the subsequent segment coincide; moreover, Lemma 1 implies that the first segment starts at K_L. As a consequence, we have characterized the width of every segment as a function of $\lambda(s_k^*)$.

Any $\lambda(s_k^*) > 0$ sustains an equilibrium. At the moment of setting prices, the market maker has a trade-off for each signal. Let's consider the highest possible signal; hence, the lowest-strike segment. If the market maker chooses a high λ, the price level is higher but the intensity of trade by informed agents is also higher. We might think informed agents would be worse off and have a profitable deviation; however, this deviation does not happen because the segment width reduces providing a higher profit if the second highest signal realizes. Given this trade-off, there are many options for the market maker. To completely characterize the equilibrium, I introduce a regularization.

Regularization 1. ***No risk.*** *Market makers set prices to make losses by noise traders independent of the private signal. Mathematically,*

$$\mathbb{E}((\theta_U - k)^+ | s_k^*) - A(k) = \mathbb{E}((\theta_U - K_L)^+ | s_{K_L}^*) - A(K_L)$$

The regularization implies

$$\lambda(s_k^*) = \frac{\alpha_{s_k^*} - \alpha_0}{\alpha_{s_{K_L}^*} - \alpha_0} \lambda(s_{K_L}^*), \quad s_k^* > 0;$$

therefore, it characterizes the equilibrium except for $\lambda(s_{K_L}^*)$. Moreover, it provides a nice insight. Market makers can set prices to eliminate the uncertainty about the private signal. In this equilibrium, the informed agent does not care about receiving a very high or a high signal because the profits are the same. If he receives a very high signal, he mixes across few strikes but low ones. Instead, if the signal is lower, he mixes across many strikes camouflaging better.

At this point, we have characterize the intensity of trading such that informed agents optimize, and the corresponding prices for which the expected profit is zero, for those strikes with positive informed trading. From Lemma 2, we know that they trade the lowest strikes; hence if there are strikes with no informed trading must be those after the last segment ($K > \overline{K}(s_\alpha) \;\; s_\alpha > 0 \; s_{\alpha-1} < 0$). Those strikes exist if the profit from lower strikes exceeds the profit of trading the strikes with prices equal to the unconditional value:

$$\mathbb{E}((\theta_U - k)^+ | s_k^*) - A(k) \geq (\alpha_{s_k^*} - \alpha_0)(\theta_U - k') \;\forall k' \geq \overline{K}(s_\alpha) \iff \overline{K}(s_\alpha) \geq \theta_U - \frac{PN}{\lambda(s_\alpha)}.$$

This expression just provides a region where the last traded strike might lie and a similar expression holds for the bid side. Nonetheless, if we combine them with the nonnegativity constraint of the trading intensity:

$$\beta(k, s_k^*) > 0 \;\; \forall k < \overline{K}(s_\alpha) \Rightarrow \overline{K}(s_\alpha) \leq \theta_U - \frac{PN}{\lambda(s_\alpha)},$$

we obtain a unique last strike traded by informed agents. Hence, if there exists $\lambda(s_{K_L}^*) > 0$ that make Equation (5) hold and $\overline{K}(s_\alpha) = \theta_U - \frac{PN}{\lambda(s_{K_L}^*)} \leq K_U$, we reached an equilibrium. Generally, multiple $\lambda(s_{K_L}^*) > 0$ will satisfy these conditions and the market maker can decide to face information asymmetries in all strikes or leave some strikes without information asymmetries.

1.3 Graphical illustration

To provide some intuition about the model, I present in Figure 1 the ask price under different set-ups. The baseline set-up considers $\frac{2}{3}$ of the investors are informed, and the remaining $\frac{1}{3}$ buy and sell with equal probability. The asset might be valued 100 or 0 with the same probability $\left(\frac{1}{2}\right)$ and the market maker provides call options with strikes ranging from 30 to 70. I consider a very simple information structure in which informed investors receive a positive signal with probability $\frac{1}{2}$. Nonetheless, there are two possible positive signals: a very informative one that leads to a certain posterior $\alpha_2 = 1$ with probability 0.4 and a noise one with posterior $\alpha_1 = 0.55$ and probability 0.6. Finally, I consider $\lambda(K_L) = 0.0006$ [5]

Panel (a) plots the equilibrium ask price and the informed intensity ($\beta(k, s_k^*)$) for all strikes below 50. Due to the limited amount of signals, we can easily observe the distinction between the different segments. There are three segments. The first segment is the one preferred after the strong signal; hence, the market maker sets a very wide spread. Note that the payoff of the call of strike 30 in the upper state is 70 but the unconditional payoff is 35. Therefore, the market maker almost eliminates the information gain. The slope of this segment is very steep because the information is very precise. The second segment corresponds to the noisy information. In this case, the segment is much wider and, consequently, the intensity of informed trading is lower and less steep because the informed agent can mix across many strikes. As a result, the adverse selection faced by the market maker in one strike is low, leading to a lower spread. Finally, the

[5] The parametrization provides an example easy to visualize. More realistic parametrizations would provide the same insights.

informed agent will never trade an option with strike higher than 44.4 because it does not provide enough cost saving compare to more in-the-money options that provide a higher exposure.

I change two of the main parameters in Panel (b). The dashed line consider the case in which only half of the investors are informed. The decrease in informed investors reduces the incentive to mix across strikes because the trades are less revealing. Consequently, informed investors trade more aggressively reducing the width of the ask segments. The dotted line considers a 2% decrease in the highest payoff. Therefore, the information rents decrease and the benefit of a higher exposure to the asset declines. Since informed investors trade off disguising as noise traders and exposure to the asset, we observe more mixing; hence, wider segments. Additionally, the price declines because the call option payoff is lower.

Panel (c) illustrates Lemma 5 because the slope of the price function does not depend on any of the parameters. I plot the ask price under the baseline parametrization and two alternative ones. The first alternative parametrization (crosses) considers two positive signals with posterior probabilities 0.75 and 0.55 and the same unconditional mean as the baseline distribution. Instead, the second parametrization (circles) considers two positive signals with the same posterior distribution of the baseline case but in which the very informative signal realizes with probability 0.2. Panel (c) shows that the baseline and the first alternative parametrization can be differentiated in the data because the slope of the ask price is different. On the other hand, the baseline and the second alternative parametrization might be observationally equivalent unless we know the remaining parameters and the true model.

Finally, panel (d) considers the role of $\lambda(K_L)$. I set this parameter arbitrarily and a range of possible values sustain an equilibrium. In panel (c), the parameter is common to all the signal specifications. Instead, in panel (d), I set the parameter such that the strikes with adverse selection are the same regardless of the information structure and equal to the maximum range across specifications. The first alternative signal structure (crosses) presents the biggest change because it had the most limited strike range. If $\lambda(K_L)$ lowers, the market maker accommodates a wider range of strikes allowing the informed agent to mix across many strikes. The result is a lower adverse selection per strike; hence a lower spread in those strikes that are traded by the informed agents in the high $\lambda(K_L)$ case. However, a low $\lambda(K_L)$ implies that more strikes have a positive spread. Noticeably, panel (d) illustrate the limit to identification. The baseline (solid line) and the second alternative signal structure (circles) lead to almost observationally equivalent ask prices despite considering a different expected value conditional on the signal being positive.

2 Extensions

The model disregards some important aspects of the market. For instance, assets do not have a binary distribution. These simplifications allow me to characterize the equilibrium and illustrate the mechanism at play. Nonetheless, the insights of the model are more general. In the following sections, I discuss different extensions of the model using the lemmas in Section 1.1.

2.1 Beyond the binary distribution

The equilibrium strategies can be described succinctly and in close form only in the case of a continuum of strikes and a binary asset value distribution. However, while maintaining the case of the continuum of strikes, we can generalize the asset distribution and still obtain most of the previous results. To do that, I introduce an assumption about the ordering of the signals.

Assumption 1. ***General asset distribution.*** *Signals can be ordered according to first-order stochastic dominance. Mathematically,*

$$s > s' \Rightarrow F(\theta|s) < F(\theta|s') \; \forall s, s' \in \mathcal{S} \cup \emptyset$$

where $F(\cdot|s)$ is the cumulative distribution function of the asset value conditional on the realization s of the private signal.

The binary distribution is the simplest case that satisfies the above assumptions but there are many other common cases. For instance, if the informed investor holds a Gaussian prior before observing the signal realization, and the signal and asset value form a joint Gaussian distribution; then, the posterior distribution of the asset satisfies the assumptions above.

Assumption 1 ensure that Lemmas 1 to 6 hold. Nonetheless, I presented Lemma 5 using the binary specification so I reformulate it using any distribution.

Lemma 7. ***Price slope.*** *In equilibrium, if informed agents play mixed strategies across some strikes, the slope of the price function of those strikes reveals the posterior probability. Mathe-*

matically,

$$\exists s \ s.t. \ q^*(k,s) \in (0,1) \Rightarrow 1 - F(k|s) = -\frac{dA(k)}{dk}$$
$$\exists s \ s.t. \ p^*(k,s) \in (0,1) \Rightarrow 1 - F(k|s) = -\frac{dB(k)}{dk}$$

In the binary case, the left-hand side remains constant across strikes. Instead, in the general case it varies across strikes because for each strike within a segment, we recover a different part of the conditional distribution function. This result allows us to elicit conditional distributions with several parameters; e.g. Gaussian distributions.

2.2 The role of non-uniform noise trading

Bryzgalova et al. (2023) and Bogousslavsky and Muravyev (2024) show that retail investors do not trade uniformly across strikes. They tend to trade at-the-money and slightly out-of-the-money options. This evidence is consistent with Admati and Pfleiderer (1988). They endogenize the participation of liquidity traders in the equity market and show they optimally choose to concentrate their trading in time. Consequently, the assumption of uniform trading by noise traders is likely violated. The model can be easily extended to capture this pattern.

Consider noise traders buy with probability ε and sell with the complement probability as they do in the baseline model. However, conditional on buying an option, the probability of buying a strike k is $h^+(k)$. Similarly, those who sell choose the strike according to the probability $h^-(k)$. If noise traders do not trade a strike, no one will. Therefore, I impose that there is always a possibility of a trade by a noise trader ($h^+(k) > 0, \ h^-(k) > 0 \ \forall \ k$). I also impose that noise traders must trade one strike $\left(\int_{K_L}^{K_U} h^+(k)dk = \int_{K_L}^{K_U} h^-(k)dk = 1\right)$.

This new set-up provide the same results as the baseline. Lemma 1 holds because it depends on whether noise traders are buyers or sellers, which remains unaltered. Lemma 2 and 3 hinge on the decreasing information advantage with respect to the strike, which also holds in the extended model. Lemmas 4 and 5 hold because the marginal gain of increasing the strike is different across signals and the same within a signal. The distribution of noise traders across all strikes generates incentives to mix; hence Lemma 6 holds.

The difference between the extended and the baseline model lies on the informed intensity to trade. Similar to the effect of PN in the baseline specification, the informed investor will trade more aggressively strikes with higher noise trading probability up to the point that the marginal benefit across all strikes traded after the same signal is the same. This result is consistent with Admati and Pfleiderer (1988) and suggests that the informed investor aligns with noise traders; hence, the market maker inventory is more uniform than the noise traders demand.

2.3 The role of inventory management

In the stylized model, the market maker is competitive and makes zero profit. This assumption is common to most information models and complex to relax[6]. Yet, inventory matters, specially in option prices. Muravyev (2016) shows that two thirds of the price impact is due to inventory and information asymmetries drive the remaining. In this section, I relax the perfect competitive assumption of market makers in terms of expected gain and allow compensation for holding positions and discuss the implications through the lenses of the model.

Consider the market maker instead of requiring zero expected profits to trade strike k in the direction d_k, she requires a compensation equal to $a_k d_k + \gamma(k)|d_k|$. I assume $\gamma(k) \geq 0\ \forall k$. We can map this compensation to Ho and Stoll (1981) seminal paper. a_k corresponds to the shifts in prices due to outstanding inventory. If market makers own a positive delta position, they might decide to decrease the bid and ask, even losing money with the trades on one side ($a_k < 0$) to avoid building more inventory. $\gamma(k)$ through the lenses of Ho and Stoll (1981) equals the risk aversion of market makers times the variance of owning a position in strike k. This modeling choice has the advantage to also map to transaction costs, outside options different per strike, etc. Nonetheless, it approximate an inventory model but it does not directly corresponds to one. The missing mechanism is the reduction in the cost of holding inventory when the market maker learns from the trade; hence reduces the conditional variance. This interaction between inventory and information models generate significant mathematical complexity and it is unlikely to be of first-order importance.

If $a_k = 0$ and $\gamma_k = \gamma$, Lemmas 1-6 hold.[7]. Therefore, if market makers are close to the optimal inventory and every option has a similar effect on their inventory, all insights hold. Instead, if γ_k varies across strikes, Lemmas 2 and 3 might not hold although the other lemmas hold. Lemma 1 follows from the original Ho and Stoll (1981): inventory concerns adds a positive component to the spread to compensate the market maker. Lemmas 4, 5 and 6 rely on the incentives

[6] Ying (2020) merges information and inventory models by assuming the latter consume at a lower frequency.

[7] Proofs are identical substituting the ask price $A(k)$ by the net ask price $A(k) - \gamma$

of informed agents to use mixed strategy equilibria to disguise within noise traders. These incentives remain unchanged. The result might seem counterintuitive because market makers may favor some strikes more than others modifying the price slopes. Instead, the informed agent internalizes this change and trades with a higher likelihood lower strikes, increasing adverse selection, and restoring the slope in prices.

When market makers are far away from the optimal inventory ($\exists\ k$ s.t. $a_k >> 0$), no lemma will hold unless we introduce significant structure in the market maker problem. In this case, an information model as the one I propose is not suited since informed agents might disregard their information and just trade the strike "on sale" due to inventory.

2.4 The role of leverage

Previous literature points to leverage as the driver of the choice between stocks and options (Easley et al., 1998, see, for example). Conditional on transacting on options, the consequences of leverage are smaller. For instance, consider a stock trades at 191 and some investors know it will trade at 200 at option maturity.[8]. If they hold 400 dollars to invest, they can buy two shares and earn 18 dollars in profits. Instead, they can buy one option contract with strike 187.5 traded at 400 dollars, and earn 1250 dollars. An alternative would be to trade an out-of-the money option and buy 1538 contracts with strike 197.5, resulting in a profit of 3845. These back-of-the-envelop calculations exemplify how options provide huge leverage in comparison to stocks (100x) but the difference between in- and at-the money options is small (2.5x). In this section, I discuss whether the main conclusions from the stylized model hold if we account for leverage.

To incorporate leverage to the model, I consider investors can trade $h(k)$ units of the option with strike k. In this case, if $h(k)$ is very high for out-of-the-money options, informed investors will trade those options after the strongest signals (Lemmas 2 and 3 might not hold). Nonetheless, they still have the incentive to mix across similar strikes; therefore, Lemma 6 remains valid. Likewise, the traded quantity cannot depend on the signal which mantains the validity of Lemma 4. Altogether, informed investors trade with certain probabilities across strikes but they only trade a given strike after one specific signal. The difference with the baseline case is that they might trade high-strike options after the strongest signals.

[8] Data for Apple on May 20,2024 and maturity May 31,2021

Consider informed investors follow a strategy similar to the baseline case in which they trade within a segment after one signal but not the others. Then, the condition for an equilibrium in mixed strategies is that they are indifferent between strikes within the same segment:

$$\frac{d}{dk}h(k)\left(\mathbb{E}\left[(\theta-k)^{+}|s\right]-A(k)\right)=0$$

Rearranging and following the same proof as in the baseline model:

$$-\frac{dA(k)}{dk}=(1-F(k|s))\left(1+\frac{dh}{k}\frac{\mathbb{E}\left[(\theta-k)^{+}|s\right]-A(k)}{h(k)}\right)$$

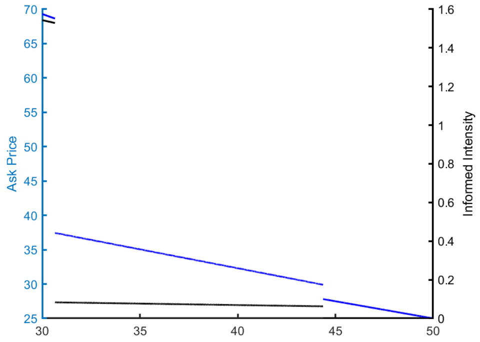
In this case, there is an extra term that captures the benefit for informed agents to trade a strike that allows a higher quantity of contracts. If we know the function $h(k)$, we can elicit the possible posterior distributions of the informed agent from the slope and level of ask prices.

3 Conclusion

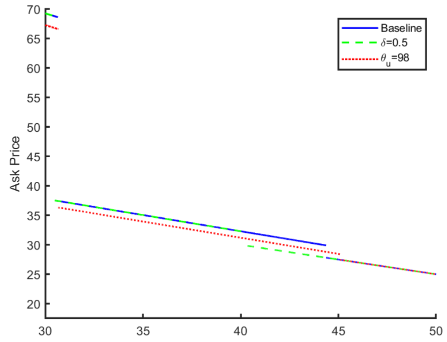
This paper provides a stylized model that characterizes informed trading in the option markets. If informed investors have a continuum of strikes available, they use mixed strategies to camouflage as noise traders. However, they do not mix across all strikes because they would lose exposure to the asset. Instead, they mix within a segment of the strike line. This segment is determined by the realization of the private signal. In particular, different realizations of the signal lead to non-overlapping segments within which they trade.

This behavior by informed agents results in a very particular price function. First, different segments of the strike line have different spread levels and different slopes with respect to the strike. Second, the slope of each segment corresponds to a section of the cumulative distribution function of the asset realization conditional on a private signal realization. Third, the signal after which informed agents trade a given segment is the conditioning set that defines the slope. Altogether, my model provides the theoretical foundation to estimate the possible posterior probabilities of informed agents trading in the options market.

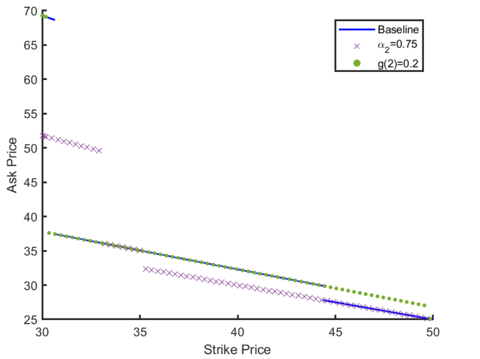
A Figures



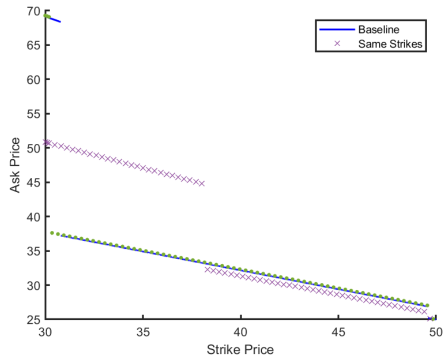
(a) Ask price vs informed intensity



(b) Effect of parameters



(c) Different signal structure



(d) Fixed strike range for informed

Figure 1: Model graphical illustration

B Auxiliary lemmas

I start by proving some lemmas that will be useful in the remaining proofs. These lemmas consider the stylized model with the generalized asset distribution.

The first lemma shows that first-order stochastic dominance is a sufficient condition for ordering signals in terms of expected payoffs regardless of the strike.

Lemma 8. ***Signals are ordered in terms of call payouts.***

$$\exists k' \ s.t. \ \mathbb{E}\big((\theta-k')^+|s\big)-\mathbb{E}\big((\theta-k')^+\big|\tilde{s}\big)>0 \Rightarrow \mathbb{E}\big((\theta-k)^+|s\big)-\mathbb{E}\big((\theta-k)^+\big|\tilde{s}\big)>0 \ \forall k\in\mathcal{K};\ \forall s\neq\tilde{s}\in\mathcal{S}\cup\emptyset$$

Proof. For any $s\in\mathcal{S}\cup\emptyset$,using integration by parts, we can write the conditional expectation as:

$$\mathbb{E}\big((\theta-k)^+|s\big)=lim_{\theta\to\infty}\theta F(\theta|s)-k-\int_k^{\bar{\theta}}F(\theta|s)d\theta$$

Therefore:

$$\mathbb{E}\big((\theta-k)^+|s\big)-\mathbb{E}\big((\theta-k)^+\big|\tilde{s}\big)=\int_k^{\bar{\theta}}\big(F(\theta|\tilde{s})-F(\theta|s)\big)d\theta$$

First order stochastic dominance (first implication below) ensures that if the difference is positive for one strike, it is positive for all other strikes.

$$\mathbb{E}\big((\theta-k')^+|s\big)-\mathbb{E}\big((\theta-k')^+\big|\tilde{s}\big)>0\Rightarrow F(\theta|\tilde{s})>F(\theta|s)\ \forall\theta\Rightarrow\mathbb{E}\big((\theta-k)^+|s\big)-\mathbb{E}\big((\theta-k)^+\big|\tilde{s}\big)>0\ \forall k$$

□

Actually, the lemma provides the difference between the expected payoff of the same option after two different signal. As a corollary, I prove that information advantage decreases with the strike.

Corollary 2. ***Information advantage decreases with strike***

$$\mathbb{E}\big((\theta-k')^+|s\big)-\mathbb{E}\big((\theta-k')^+\big|\tilde{s}\big)>\mathbb{E}\big((\theta-k)^+|s\big)-\mathbb{E}\big((\theta-k)^+\big|\tilde{s}\big)\ \forall k>k'\ |s|>|\tilde{s}|;\ \forall s\neq\tilde{s}\in\mathcal{S}\cup\emptyset$$

The next lemma proves the intuitive result that informed agents buy options whose expected payoff according to the signal is above the ask price.

Lemma 9. ***Informed traders buy undervalued assets and sell overvalued ones.***

$$q^*(k,s)>0\Rightarrow\mathbb{E}\big((\theta-k)^+|s\big)-A^*(k)\geq 0 \tag{6}$$

$$p^*(k,s)>0\Rightarrow B^*(k)-\mathbb{E}\big((\theta-k)^+|s\big)\geq 0 \tag{7}$$

Proof. Consider the profit function:

$$\Pi\equiv\sum_{k=K_L}^{K_U}\kappa(k)q(k,s)\Big(\mathbb{E}\big((\theta-k)^+|s\big)-A^*(k)\Big)+\sum_{k=K_L}^{K_U}\kappa(k)p(k,s)\Big(B^*(k)-\big(\mathbb{E}\big((\theta-k)^+|s\big)\Big) \tag{8}$$

subject to the law of total probability and nonnegativity constraints. Due to the linearity of the problem, one of the restrictions will always bind so $\sum_{k=K_L}^{K_U}q(k,s)=1\vee q^*(k,s)=0\ \forall k$ and the same holds for $p^*(k,s)$. From this linear problem, we clearly see the implication of the lemma as otherwise, the informed agent can set to 0 the p or q for which the right-hand implication is not true and increase his profits. □

A common result in models of private information in financial markets is the presence of a positive spread. Next lemma shows that this result holds in my model for every option.

Lemma 10. ***There is a spread around the unconditional value.***

$$A^*(k) \geq \mathbb{E}((\theta - k)^+) \geq B^*(k)$$

and the first (second) inequality is strict for strike k if there exists a signal s such that $q^*(k,s) > 0$ ($p^*(k,s) > 0$).

Proof. Let's start by proving $A^*(k) \geq \mathbb{E}((\theta - k)^+)$ in equilibrium. The per unit profits for the market maker are given by:

$$\begin{aligned}\pi^A(A^*(k),k) &= A(k) - \mathbb{E}((\theta-k)^+|d_k = 1)\\ &= A(k) - \sum_{s_1}^{s_M} \mathbb{E}((\theta-k)^+|s)q(k,s)P(Inf|d_k=1) - \mathbb{E}((\theta-k)^+)(1-P(Inf|d_k=1))\\ &= -\sum_{s_1}^{s_M}\Big[\mathbb{E}((\theta-k)^+|s) - A^*(k)\Big]P(Inf,s|d_k=1) - \Big[\mathbb{E}((\theta-k)^+) - A^*(k)\Big](1-P(Inf,s|d_k=1))\end{aligned}$$

Note that $P(Inf, s|d_k = 1)$ depends on $q(k,s)$. Using Bayes:

$$P(Inf,s|d_k=1) = \frac{\sum_{s_1}^{s_M} q^*(k,s)g(s)\delta}{P(d_k=1)}$$

Therefore we have two cases. First, if $q(k,s) = 0$; hence $P(Inf,s|d_k = 1) = 0$

$$\pi^A(A^*(k),k) = 0 \Rightarrow \mathbb{E}((\theta-k)^+) - A^*(k)) = 0 \Rightarrow A^*(k) = \mathbb{E}((\theta-k)^+)$$

Second, if $q(k,s) > 0$, we know from Lemma 9 that $\mathbb{E}((\theta-k)^+|s) - A^*(k) > 0$; therefore,

$$\pi^A(A^*(k),k) = 0 \Rightarrow \mathbb{E}((\theta-k)^+) - A^*(k) < 0 \Rightarrow A^*(k) > \mathbb{E}((\theta-k)^+)$$

□

C Proofs of the lemmas of the stylized model

I prove Lemmas 1 to 6 of the stylized model using the general asset distribution to reduce the number of proofs.

C.1 Lemma 1

Note that Lemmas 9 and 10 imply:

$$q^*(k,s) > 0 \underset{L9}{\Rightarrow} \mathbb{E}\left((\theta-k)^+|s\right) > A(k) \underset{L10}{>} B(k) \underset{L9}{\Rightarrow} p^*(k,s) = 0$$

In words, if the informed buys some option with positive probability, it does not sell the same option. Lemma 1 is more general because it requires that buying an option with positive probability implies that the informed agent does not sell any other option. Nonetheless, the general proof is similar but we require Lemma 8:

$$q^*(k,s) > 0 \underset{L9}{\Rightarrow} \mathbb{E}\left((\theta-k)^+|s\right) > A(k) \underset{L10}{>} \mathbb{E}\left((\theta-k)^+\right) \underset{L8}{\Rightarrow} \mathbb{E}\left((\theta-k')^+|s\right) > \mathbb{E}\left((\theta-k')^+\right) \forall k'$$

$$\mathbb{E}\left((\theta-k')^+|s\right) > \mathbb{E}\left((\theta-k')^+\right) \forall k' \underset{L9,10}{\Rightarrow} p^*(k',s) = 0 \;\forall k'$$

For the last implication, note that Lemma 10 implies $\mathbb{E}\left((\theta-k')^+\right) \geq B(k')$. The equality case implies $p^*(k',s) = 0$ directly. In the strict inequality case, Lemma 9 implies the result.

C.2 Lemma 2

Let's prove this lemma by contradiction assuming there is an equilibrium such that $q^*(k', s') > 0$ and $q^*(k, s) = 0 \forall s \;\; k < k'$. I focus on the ask side but the proof for the bid side is symmetric. By the first equilibrium condition and Lemma 10, we know:

$$A(k) = \mathbb{E}((\theta - k)^+ | d_k = 1) = \mathbb{E}((\theta - k)^+) \text{ and } A(k') > \mathbb{E}((\theta - k')^+)$$

since only noise traders trade at strike k but there is informed trading at k'. To arrive a contradiction, we just need to find a strategy of the informed agent that provides higher profits than the one assumed. Let consider an alternative $\widehat{q}(k, s)$ in which $\widehat{q}(k, s') = q^*(k', s')$, $\widehat{q}(k', s') = 0$, and $\widehat{q}(k, s) = q^*(k, s)$ for every other k, s. For completeness, consider $\widehat{p}(k, s) = p^*(k, s)$. Let denote the profits that the informed agent obtains from this new strategy as $\widehat{\Pi}(s')$ and from the assumed one as $\Pi(s')$. Then, we have:

$$\begin{aligned}
\widehat{\Pi}(s') - \Pi(s') &= q^*(k', s')\Big(\big(\mathbb{E}\big((\theta - k)^+ | s'\big) - A(k)\big) - \big(\mathbb{E}\big((\theta - k')^+ | s'\big) - A(k')\big) \Big) \\
&> q^*(k', s')\Big(\big(\mathbb{E}\big((\theta - k)^+ | s'\big) - \mathbb{E}((\theta - k)^+)\big) - \big(\mathbb{E}\big((\theta - k')^+ | s'\big) - \mathbb{E}((\theta - k')^+)\big) \Big) \\
&\underset{C2}{>} 0
\end{aligned}$$

which is a contradiction.

C.3 Lemma 3

I prove Lemma 3 by contradiction, similar to the previous lemma. Assume $q^*(k', s') > 0$ and $\exists k'', s''$ s.t. $q^*(k'', s'') > 0, k'' > k', s'' > s'$. I consider an alternative strategy such that $\hat{q}(k', s'') = q^*(k', s'') + q^*(k'', s'')$ and $\hat{q}(k'', s'') = 0$. The remaining part of the strategy remains unchanged. Importantly, I maintain the assumption that $q^*(k', s') > 0$. Let denote the profits that the informed agent obtains from this new strategy as $\widehat{\Pi}$ and from the assumed one as Π. Then, we have:

$$\widehat{\Pi} - \Pi = q^*(k'', s'')\Big(\big(\mathbb{E}\big((\theta - k')^+ | s''\big) - A(k')\big) - \big(\mathbb{E}\big((\theta - k'')^+ | s''\big) - A(k'')\big) \Big)$$

Note that $q^*(k', s') > 0$ implies:

$$\begin{aligned}
&\mathbb{E}\big((\theta - k')^+ | s'\big) - A(k') \geq \mathbb{E}\big((\theta - k'')^+ | s'\big) - A(k'') \\
&\Rightarrow A(k'') - A(k') > \mathbb{E}\big((\theta - k'')^+ | s'\big) - \mathbb{E}\big((\theta - k')^+ | s'\big) \qquad (9)
\end{aligned}$$

as otherwise we can decrease $q^*(k', s')$ and increase $q^*(k'', s')$ and earn a higher profit as we did in the proof of Lemma 2. Corollary **??** implies:

$$\begin{aligned}
&\mathbb{E}\big((\theta - k')^+ | s''\big) - \mathbb{E}\big((\theta - k')^+ | s'\big) > \mathbb{E}\big((\theta - k'')^+ | s''\big) - \mathbb{E}\big((\theta - k'')^+ | s'\big) \\
&\Rightarrow \mathbb{E}\big((\theta - k'')^+ | s'\big) - \mathbb{E}\big((\theta - k')^+ | s'\big) > \mathbb{E}\big((\theta - k'')^+ | s''\big) - \mathbb{E}\big((\theta - k')^+ | s''\big)
\end{aligned}$$

Hence, using (9)

$$\begin{aligned}
&A(k'') - A(k') > \mathbb{E}\big((\theta - k'')^+ | s''\big) - \mathbb{E}\big((\theta - k')^+ | s''\big) \\
&\Rightarrow \mathbb{E}\big((\theta - k')^+ | s''\big) - A(k') > \mathbb{E}\big((\theta - k'')^+ | s''\big) - A(k'')
\end{aligned}$$

Therefore, $\widehat{\Pi} - \Pi > 0$ and we arrived to a contradiction. The proof for the bid side is symmetric.

C.4 Lemma 4

Note that $q^*(k,s) > 0 \wedge q^*(k',s) > 0$ s.t. $k \neq k' \Rightarrow \mathbb{E}\big((\theta-k)^+|s\big) - A^*(k) = \mathbb{E}\big((\theta-k')^+|s\big) - A^*(k')$ as otherwise the strategy does not maximize the informed agent profits. Rearranging we get:

$$\mathbb{E}\big((\theta-k)^+|s\big) - \mathbb{E}\big((\theta-k')^+|s\big) = A^*(k) - A^*(k')$$

The right-hand side is set by the market maker; therefore, it cannot depend on the realized value of the signal. On the other hand, Corollary 2 shows that the left-hand change monotonically with the signal. Therefore, there exist at most one signal realization after which this expression holds. Note that $p^*(k,s) > 0 \wedge p^*(k',s) > 0$ has the same implication; thus, the proof applies to the bid side of the market.

C.5 Lemma 5

First, note that Lemma 2, Lemma 3, and Lemma 1 ensure that $q^*(k,s) \in (0,1)$ and $\sum_{i=K_1}^{k} q^*(k,s) < 1 \Rightarrow q^*(k+\Delta K, s) \in (0,1)$. Hence, using the last result from the indifference condition, we know::

$$\mathbb{E}\big((\theta-(k+\Delta K))^+|s\big) - \mathbb{E}\big((\theta-k)^+|s\big) = A(k+\Delta K) - A(k)$$

$$\int_{k}^{k+\Delta K} (\theta-k)dF(\theta|s) + \int_{k+\Delta K}^{\overline{\theta}} \Delta K dF(\theta|s) = -A(k+\Delta K) - A(k)$$

where $F(\theta|s)$ is the c.d.f. of the asset given signal s. Integrating by parts,

$$\Delta K - \int_{k}^{k+\Delta K} F(\theta|s)d\theta = -(A(k+\Delta K) - A(k))$$

$$\Rightarrow 1 - \frac{1}{\Delta K}\int_{k}^{k+\Delta K} F(\theta|s)d\theta = -\frac{A(k+\Delta K) - A(k)}{\Delta K}$$

Using the binary specification, $F(\theta|s) = 1 - \alpha_s$, we get the desired result.

Corollary 3. ***The limiting case of Lemma 5.***
It is convenient to consider the limiting case when $\Delta K \to 0$.

Proof. We can then use L'Hôpital's rule to get:

$$1 - F(k|s) = -\frac{dA(k)}{dk}$$

□

C.6 Lemma 6

We want to prove that

$$q^*(k,s) > 0 \Rightarrow lim_{N\to\infty} Nq^*(k,s) = \beta(k,s) < \infty$$

Proof. The conditional probability of being informed:

$$P(Inf|d_k = 1) = \frac{\delta \sum_{s_1}^{s_M} q^*(k,s)g(s)}{\delta \sum_{s_1}^{s_M} q^*(k,s)g(s) + (1-\delta)\varepsilon\frac{1}{N}}$$

Using Lemma 6 and $q^*(k,s) > 0$, we can simplify the expression to:

$$P(Inf|d_k = 1) = \frac{\delta q^*(k,s)g(s)}{\delta q^*(k,s)g(s) + (1-\delta)\varepsilon\frac{1}{N}}$$

Multiplying numerator and denominator by the number of strikes:

$$P(Inf|d_k = 1) = \frac{N\delta q^*(k,s)g(s)}{N\delta q^*(k,s)g(s) + (1-\delta)\varepsilon}$$

We have three cases:

1. $q^*(k,s) = o(N) \Rightarrow \lim_{N\to\infty} P(Inf|d_k = 1) = 0$
2. $q^*(k,s) > O(N) \Rightarrow \lim_{N\to\infty} P(Inf|d_k = 1) = 1$
3. $lim_{N\to\infty} Nq^*(k,s) = \beta(k,s) < \infty \Rightarrow P(Inf|d_k = 1) \in (0,1)$

I show that the first two cannot sustain an equilibrium. In the first case, the ask price of the call is $\lim_{N\to\infty} A(k) = \mathbb{E}((\theta - k)^+)$. Then, for any signal with $\mathbb{E}((\theta - k)^+|s) > \alpha$, the informed agent can obtain a higher profit by increasing $q^*(k,s)$. Therefore $q^*(k,s)$ cannot be an equilibrium unless the total probability restriction binds. The following proves it does not:

$$lim_{N\to\infty} \sum_{k=K_1}^{K_N} q^*(k,s) = \lim_{N\to\infty} \sum_{n=0}^{N-1} q^*\big(K_1 + \frac{n}{N}(K_N - K_1), s\big) \le \lim_{N\to\infty} N \max_k q^*(k,s) = 0 \tag{10}$$

In the second case, the ask price of the call is $\lim_{N\to\infty} A(k) = \mathbb{E}((\theta - k)^+|s)$ and the informed agent will make zero profits. Hence, he will increase the profits by trading any of the strikes with in which $Pr(Inf|d_k = 1) = 0$. We can prove that these strikes exists by contradiction.

Assume $\exists s\ s.t.\ lim_{N\to\infty} Nq^*(k,s) = \infty\, \forall k$, then the total probability of trading after any signal s is $\lim_{N\to\infty} \sum_{K_0}^{K_N} q^*(k,s) \ge \lim_{N\to\infty} \sum_{s=s_1}^{s_M} N \min_k q^*(k,s) = \infty$ which leads to a contradiction since $q^*(k,s)$ is a probability and must be lower than 1.

□

D Equilibrium derivations

Let's denote s_k^* the signal after which the informed investor buys strike k. Note that Lemma 4 ensures that s_k^* (if exists) is unique. we start by characterizing the intensity of informed trading per strike. Hence, we consider only strikes traded after some signal s.

The first equilibrium condition (perfect competition) leads to:

$$A(k) = \mathbb{E}((\theta - k)^+|s^*)P(Inf|d_k = 1) + \mathbb{E}((\theta - k)^+)P(\overline{Inf}|d_k = 1)$$

$$A(k) = \mathbb{E}((\theta - k)^+|s^*) + \Big(\mathbb{E}((\theta - k)^+) - \mathbb{E}((\theta - k)^*|s^*)\Big)P(\overline{Inf}|d_k = 1) \tag{11}$$

Taking derivatives:

$$\frac{dA(k)}{dk} = \frac{d\mathbb{E}((\theta - k)^+|s^*)}{dk} + \frac{d\Big(\mathbb{E}((\theta - k)^+) - \mathbb{E}((\theta - k)^+|s^*)\Big)P(\overline{Inf}|d_k = 1)}{dk}$$

Corollary 3 implies $\frac{dA(k)}{dk} = F(k|s^*) - 1$. Moreover, using Leibniz's rule:

$$\frac{d\mathbb{E}((\theta - k)^+|s^*)}{dk} = F(k|s^*) - 1$$

Therefore,

$$\frac{d\Big(\mathbb{E}((\theta - k)^+) - \mathbb{E}((\theta - k)^+|s^*)\Big)P(\overline{Inf}|d_k = 1)}{dk} = 0$$

$$\Rightarrow \frac{d\Big(\mathbb{E}((\theta-k)^+|s^*)-\mathbb{E}((\theta-k)^+)\Big)}{dk}P(\overline{Inf}|d_k=1)=\frac{dP(\overline{Inf}|d_k=1)}{dk}\Big(\mathbb{E}((\theta-k)^*)-\mathbb{E}((\theta-k)^*|s^*)\Big) \tag{12}$$

Intuitively, as the information advantage increases, the probability of informed trading must decrease. Note that:

$$P(\overline{Inf}|d_k=1)=\frac{P(d_k=1|\overline{Inf})(1-\delta)}{P(d_k=1)}$$

while the denominator depends on the intensity of informed trading the numerator does not; hence,

$$\begin{aligned}\frac{dP(\overline{Inf}|d_k=1)}{dk} &= -\frac{P(d_k=1|\overline{Inf})(1-\delta)}{P(d_k=1)^2}\frac{dP(d_k=1)}{dk}\\ &= -P(\overline{Inf}|d_k=1)\frac{dP(d_k=1)}{dk}\frac{1}{P(d_k=1)}\\ &= -\frac{dlog(P(d_k=1))}{dk}\end{aligned}$$

Therefore, Equation (12) simplifies to:

$$\frac{d\Big(\mathbb{E}((\theta-k)^+|s^*)-\mathbb{E}((\theta-k)^+)\Big)}{dk}=-\frac{dlog(P(d_k=1))}{dk}\Big(\mathbb{E}((\theta-k)^*)-\mathbb{E}((\theta-k)^*|s^*)\Big)$$

Using the binary assumption: $\mathbb{E}((\theta-k)^+|s^*)=\alpha_s(\theta_U-k)$, we get to:

$$\frac{1}{\theta_U-k}dk=dlog(P(d_k=1))$$

Integrating both sides and applying the exponential function:

$$P(d_k=1)=\lambda(s^*)(\theta_U-k) \tag{13}$$

$\lambda(s^*)>0$ arises as the integration constant. We can use the total probability theorem to compute $\beta(k,s)$:

$$P(d_k=1|\overline{Inf})(1-\delta)+\delta g(s^*)\beta(k,s^*)=\lambda(s^*)(\theta_U-k)\Rightarrow\beta(k,s^*)=\frac{1}{\delta g(s^*)}\left(\lambda(s^*)(\theta_U-k)+PN\right) \tag{14}$$

To compute prices we start by computing the probability of a noise trader:

$$P(\overline{Inf}|d_k=1)=\frac{P(d_k=1|\overline{Inf})(1-\delta)}{P(d_k=1)}$$

From Equations (13) and (14), we can substitute to obtain:

$$P(\overline{Inf}|d_k=1)=\frac{\lambda(s^*)(\theta_U-k)-\delta g(s^*)\beta(k,s^*)}{\lambda(s^*)(\theta_U-k)}=1-\frac{\delta g(s^*)\beta(k,s^*)}{\lambda(s^*)(\theta_U-k)}$$

Using this expression and the binary specification in equation (11), we get the price as a function of informed intensity:

$$A(k)=\alpha_s^*(\theta_U-k)-(\alpha_s^*-\alpha_0)\left(1-\frac{\delta g(s^*)\beta(k,s^*)}{\lambda(s^*)(\theta_U-k)}\right)(\theta_U-k)$$

Substituting $\beta(k,s^*)$, we obtain the result.

To obtain the limit of the segments, we ensure informed agents trade with probability equal to one:

$$\int_{\underline{K}}^{\overline{K}} \frac{1}{\delta g(s_k^*)} \left(\lambda(s_k^*)(\theta_U - k) - PN\right) = 1$$

The integral results in a second degree polynomial with two roots:

$$\overline{K}(s_i^*) = \left(\theta_U - \frac{PN}{\lambda(s_i^*)}\right) \pm \sqrt{\left(\left(\theta_U - \frac{PN}{\lambda(s_i^*)}\right) - \underline{K}(s_i^*)\right)^2 - 2\frac{\delta g(s_i^*)}{\lambda(s_i^*)}}$$

The highest root (plus sign) implies $\overline{K}(s_i^*) > \left(\theta_U - \frac{PN}{\lambda(s_i^*)}\right) \Rightarrow \exists k$ s.t. $\beta(k, s_k^*) < 0$; hence it cannot be an equilibrium. On the other hand, the lowest root ensures the intensity is positive for all strikes.

References

A. R. Admati and P. Pfleiderer. A theory of intraday patterns: Volume and price variability. *The review of financial studies*, 1(1):3–40, 1988.

K. R. Ahern. Do proxies for informed trading measure informed trading? evidence from illegal insider trades. *The Review of Asset Pricing Studies*, 10(3):397–440, 2020.

P. Augustin, M. Brenner, G. Grass, P. Orłowski, and M. G. Subrahmanyam. Informed options strategies before corporate events. *Journal of Financial Markets*, 63:100766, 2023.

V. Bogousslavsky and D. Muravyev. An anatomy of retail option trading. *Available at SSRN*, 2024.

O. Bondarenko and D. Murayev. What information do informed traders use? *SSRN*.

D. T. Breeden and R. H. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of business*, pages 621–651, 1978.

S. Bryzgalova, A. Pavlova, and T. Sikorskaya. Retail trading in options and the rise of the big three wholesalers. *The Journal of Finance*, 78(6):3465–3514, 2023.

M. Cipriani and A. Guarino. Herd behavior and contagion in financial markets. *The BE Journal of Theoretical Economics*, 8(1):0000102202193517041390, 2008.

M. Cipriani and A. Guarino. Estimating a structural model of herd behavior in financial markets. *American Economic Review*, 104(1):224–51, 2014.

D. Easley and M. O'hara. Price, trade size, and information in securities markets. *Journal of Financial economics*, 19(1):69–90, 1987.

D. Easley, N. M. Kiefer, M. O'hara, and J. B. Paperman. Liquidity, information, and infrequently traded stocks. *The Journal of Finance*, 51(4):1405–1436, 1996.

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

S. Figlewski. Risk-neutral densities: A review. *Annual Review of Financial Economics*, 10: 329–359, 2018.

L. R. Glosten and P. R. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics*, 14(1):71–100, 1985.

T. Ho and H. R. Stoll. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial economics*, 9(1):47–73, 1981.

D. Muravyev. Order flow and expected option returns. *The Journal of Finance*, 71(2):673–708, 2016.

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

C. Ying. The pre-fomc announcement drift and private information: Kyle meets macro-finance. *Available at SSRN 3644386*, 2020.