Why Does Public News Augment Information Asymmetries?

Julio A. Crego[*]

Latest version: HERE
This version: January 30, 2018

Abstract

The arrival of a public signal worsens the adverse selection problem if informed investors are risk-averse. To illustrate the mechanism, I propose a dynamic model with a public signal and endogenous participation. In this set-up, the public signal reduces uncertainty which boosts informed investors' participation leading to a more toxic order flow. I confirm the model's empirical predictions by estimating the effect of the publication of the weekly change in oil inventories on liquidity via a difference-in-difference strategy. Precisely, the mean bid-ask spread doubles immediately after the release and volume increases by 32 percent regardless of the report's content.

Keywords: Public Information, News Release, Asymmetric Information, Liquidity.

JEL: G12, G14.

[*] *Tilburg University.* I am indebted to my Ph.D. advisor Dante Amengual for his encouragment and useful comments. For helpful suggestions, I also thank Manuel Arellano, Fabio Braggion, Sabrina Buti, Guillermo Caruana, Julio Gálvez, Thomas Gehrig, María Teresa González Pérez, Carolina Manzano, Monica Martinez-Bravo, Roberto Pascual, José Penalva, Rafael Repullo, Enrique Sentana, and Mark van Achter. Financial support from the Santander Research Chair at CEMFI is gratefully acknowledged. Errors and omissions are exclusively my own responsibility. *Email: jacrego@uvt.nl*

Public information is the source of large market fluctuations, however, contrary to common intuition, it augments information asymmetries.[1] Moreover, volume spikes immediately after a public announcement, independently of its content, which cannot be explained by existing theories based on rational agents (Bollerslev et al., forthcoming). This paper proposes a novel mechanism that explains these phenomena in a set-up where agents are rational.

Consider a market in which there are two types of traders. The first one comprises investors with liquidity or hedging needs, thus they sell and buy for exogenous reasons. Meanwhile, the second group consists of risk-averse speculators who own some private information about the asset and trade to monetize their informational advantage. Finally, there is a third agent, the market maker, who provides liquidity to both sides of the market but it is not able to identify the type of each trader. Glosten and Milgrom (1985) shows that this last agent provides a spread between the price to buy and the price to sell such that the loses from informed investors compensate the gains from the remaining traders.

The mechanism I propose hinges on how a public signal endogenously changes the composition of traders in the market due to risk aversion. Before the arrival of the signal, some informed agents refrain from trading because the gains from their informational advantage do not compensate the risk due to the remaining uncertainty about the asset value. When they have observed the signal, part of the uncertainty is resolved and they enter the market; at the same time, the market maker realizes that the adverse selection problem has worsened and widens the spread. Afterward, since a more informative order flow enhances learning from trades, information rents shrink inducing a reduction in spreads and informed trading. Therefore, the effect on volume and spreads diminishes over time and it finally dissipates.

I extend the baseline model in different dimensions to ensure that results do not strongly depend on the assumptions. First, I explicitly model investor with liquidity or hedging needs instead of assuming that they trade randomly. In this case, the empirical predictions remain unchanged but the effect on volume is alleviated. Intuitively, the

[1]Prior literature has shown this empirical fact after earning announcements (Lee et al., 1993; Krinsky and Lee, 1996), macroeconomic releases (Green, 2004) and Reuters releases (Riordan et al., 2013) among other publications.

spread is a measure of liquidity cost and these agents might react to the spike in spreads by leaving the market. Second, I show that the assumption of market maker's risk neutrality is a direct implication of perfect competition. Finally, I consider a public signal that is positively correlated with the private signal, instead of independent, and the main results prevail as long as the correlation is low.

The model provides clear empirical predictions. First, prices, volume, and bid-ask spreads should not be affected by the public signal before its release. At that time, the model predicts that prices react to the content of the announcement, and bid-ask spreads and volume increase. Afterward, prices should remain constant forever in expectation while volume and spreads should decrease over time until these variables return to the levels in the case of the absence of public news. Furthermore, according to the model, these movements in spreads and volume are independent of the public signal's realization since they are driven by the reduction in uncertainty. This independence provides an empirical prediction that differentiate my proposed channel from the ones previously discussed in the literature.[2] Hence, it is key to analyze the dynamics of volume and bid-ask spreads around a public announcement conditional on its content, which is the empirical contribution of this paper.

I test the implications of the theoretical model within the context of the Weekly Petroleum Status (WPS) report, which provides an appropriate setting for a credible empirical exercise. This document, which is released by the Energy Information Administration (EIA) every Wednesday at 10:30 a.m., contains the official oil inventory levels in the US. Moreover, it has four features necessary to test the empirical hypotheses of the model: First, it is crucial for pricing. I show that returns on Wednesdays at 10:30 a.m. are 28 times more volatile than the median minute. Second, unlike earnings announcements or press articles, the information content of the oil inventory report has a clear implication for oil prices. Specifically, inventory build-up is a signal of a negative shock in demand or a positive shock in supply which entails a decrease in prices. Third, the information is released automatically at a specific time, 10:30 a.m. This feature ensures that the release time does not depend on the signal realization, and allows me to identify the intraday

[2] Section 4 provides a brief discussion about the alternative explanations.

dynamics of the effect. Fourth, the report release affects firms highly dependent on oil, while the remaining sectors do not react to the report.

I use this last feature to design a difference-in-difference strategy. Precisely, I consider firms whose main input is oil as treated, and the remaining firms make up the control group. Similarly, the second difference compares Wednesdays with the remaining weekdays. I conduct the estimation across days and firms using data from a given minute. As a result, I identify the effect of the public release at every minute.

In line with the theoretical model, there is no significant effect before 10:30 a.m. At that moment, prices rise (decline) by 5.5 bps if inventories decrease (increase) by one standard deviation, but they do not change if the released data equals the expected change in inventories. The rest of the trading day prices remain on average constant. Meanwhile, bid-ask spreads widen by 2.79 bps and volume boosts by 32%, independently of the change in inventories, and this effect is persistent. Nonetheless, at the end of the day, effects become insignificant in agreement with the limiting results of the model.

These results remain valid if the signal content is substituted by its absolute value, if negative and positive values are allowed to have a different effect or if we divide the signal domain using binary variables, which supports the idea of independence between spreads and volume and the change in inventory. Further, to test the presence of parallel trends, I show that results are not a consequence of just one weekday, neither a specific firm. Moreover, results are more pronounced if we consider prices of oil ETFs instead of oil-firm equities as the information channel predicts.[3]

This paper is not the first one to study the theoretical implications of a public announcement. Kim and Verrecchia (1991), using earnings announcements as an example, considers the case in which only informed agents understand the public signal. In this situation, the public announcement worsens adverse selection because informed agents' rents increase. On the other hand, Pasquariello and Vega (2007) and Tetlock (2010) predict that this problem improves because public announcements disclose previously held private information. Kandel and Pearson (1995) obtains a similar prediction under the assumption that agents disagree about the content of the signal. These models focus on

[3]Some of these results are presented in the online appendix.

how public news modifies information sets, instead of participation; therefore they are not able to explain the increase in volume and spread regardless of the actual content.

I also contribute to a broader theoretical literature by constructing a model à la Glosten and Milgrom (1985) with a public signal and risk-averse informed traders. To my knowledge, this is the first paper to include a public signal in an information-based dynamic model.[4] Regarding risk-aversion, Holden and Subrahmanyam (1994) introduces risk-averse informed traders in Kyle (1985)'s model, but they do not consider public news. In a different vein, inventory models such as Ho and Stoll (1981), Madhavan and Smidt (1991), Hendershott and Menkveld (2014) and references therein, consider a risk-averse market maker. These papers, however, abstract from information asymmetries.

Even if I focus on the interaction between public news and risk aversion, the latter has important implications for the market maker's learning process. First, the number of past buys and sales is not enough to characterize the market maker's beliefs since the order matters. This feature arises from the endogenous composition of traders, if prices are high, which implies higher information rents for a low-value asset, sales are likely to be from informed agents. Therefore, a buy followed by a sale leads to a lower price than a buy led by a sell. Second, periods without trading become a signal about the value of the asset. Precisely, no-trading periods when prices are high signal a positive value, whereas the opposite occurs with no-trading periods during low prices.

Regarding the empirical contribution, the purpose of my paper is to analyze the effect of a public announcement on liquidity conditional on the actual data announced. This feature is missing in the existing literature on the effect of public news on asset markets.[5] However, Brown et al. (2009) and Riordan et al. (2013) already take a step in this direction by conditioning on the sign of the data. The former shows that positive earning surprises decrease information asymmetry with respect to zero-surprise announcements, whereas bad surprises increase it. Likewise, Riordan et al. (2013) identifies that the adverse

[4] O'Hara (1995) reviews the seminal information-based models.

[5] While previous research does not study the WPS, it analyzes macroeconomic announcements (Ederington and Lee, 1993; Fleming and Remolona, 1999; Green, 2004; Pasquariello and Vega, 2007; Hess et al., 2008; Elder et al., 2012), earnings announcements (Beaver, 1968; Ball and Brown, 1968; Foster et al., 1984; Bernard and Thomas, 1989, 1990; Lee et al., 1993; Krinsky and Lee, 1996; Vega, 2006; Savor and Wilson, 2016), and news feeds (Liu et al., 1990; Barber and Loeffler, 1993; Berry and Howe, 1994; Chan, 2003; Antweiler and Frank, 2004; Chae, 2005; Ranaldo, 2005; Tetlock, 2010).

selection problem worsens after Reuters releases some information, especially if it conveys negative news.

Besides considering the continuous realization of the signal, there are several differences between my study and the previous ones due to the nature of the public announcement and the empirical strategy. For instance, bad news and negative earnings surprises are more prevalent in times when uncertainty is higher (crises periods), therefore, results might be driven by changes in uncertainty instead of being a consequence of the news' content. Using a difference-in-difference methodology and month-year fixed effects, I tackle this concern. Another issue is the endogeneity that arises because journalists and CFOs select what to publish. While a CEO diverting money to her account might cover the front page of financial newspapers, we do not read the opposite information every day. The institutional set-up of the WPS report solves this problem since it is released by a governmental agency in a mechanical fashion.

In the following section, I outline the model and derive the empirical predictions. Next, in Section 2, I describe the institutional setting, and the data that I use to conduct the empirical analysis presented in Section 3. Section 4 compares the empirical results with the predictions of alternative mechanisms, and Section 5 concludes. Proofs of the theoretical results are gathered in Appendix A. Finally, I include some additional empirical results in the Online Appendix to which I refer adding the prefix *S*.

1 Model

Consider a market in which a unique asset can be traded at T different periods in time. The liquidation value of the asset equals the sum of three independent components: $v = \omega + \mu + \varepsilon$. The first one, ω, is subject to an asymmetric information problem, because some traders know it from the first period ($t = 0$) while the remaining agents can only learn it by observing the transaction history. Nonetheless, every agent in the model understands that ω can take two values: σ_ω with probability $\frac{1}{2}$ and $-\sigma_\omega$ otherwise. On the other hand, no trader knows μ before the public signal is released ($t = t_R$), and everyone does afterward. Finally, the third component, ε, represents the part of the asset value that is only known at liquidation. Without loss of generality, I assume that μ and ε have an expectation equal to zero and a variance equal to σ_μ and σ_ε respectively.

There are two types of traders in the market: a mass $1-\delta$ of noise traders who buy and sell randomly with probability $\frac{1}{2}$, and a mass δ of informed traders who know the realization of ω and maximize their expected utility. At each time t, one random trader buys, sells or does not trade, and she dies.[6] While for some not very liquid stocks, strategic behavior might be important; in the case of highly liquid stocks, describing traders as price takers is not a far-fetched assumption. Additionally, hereafter, I assume that both types of agents exist in the market ($0<\delta<1$).

Contrary to previous literature, informed agents are risk-averse. Specifically, informed agents' utility is given by $U(d_i) = \left(\mathbb{E}\left(v|\omega\right) - p\left(d_i\right)\right) d_i - \gamma_i Var\left(v|\omega\right) |d_i|$ where $d_i = 1$ ($d_i = -1$) if they buy (sell) and $d_i = 0$ if they do not take any action; and $p(1)$ is the ask price while $p(-1)$ denotes the bid price. Note that, as it is common in the literature, informed investors trade, at most, one unit.[7] Additionally, informed investors are heterogeneous in their risk aversion since γ is randomly distributed with c.d.f $F(\cdot)$. To ease the exposition, I assume that $F(\cdot)$ is differentiable and strictly increasing with support equal to $\mathbb{R}^+$. Direct evidence on informed investors' preferences is unachievable, nonetheless, using a structural model, Koijen (2014) estimates a risk aversion coefficient equal to 5.75 with a standard deviation of 10.53 which supports the heterogeneity in risk aversion. Likewise, insiders, or other proprietary traders, invest using their own money; thus, they are likely to act as risk-averse agents. Besides, differences in capital, leverage, or non-financial characteristics can create heterogeneity in their attitude towards risk.

There is a third type of agent, a competitive market maker, who sets the bid and ask price (A and B respectively) to make zero profits in expectation. This characterization of the market maker as a risk-neutral agent can be justified by competition under mild assumptions (see Section 1.4.2).

Following Glosten and Milgrom (1985), agents in the market are born at the beginning of each of the T periods knowing the whole sequence of past transactions and prices, and ω if they are informed, and they die at the end. The timing within a period is as follows:

[6]This assumption is common to previous literature (see Glosten and Milgrom, 1985; Easley and O'Hara, 1987, and references thereafter)

[7]An exception is the model proposed by Easley and O'Hara (1987) which allows for big and small trades.

first, and only in period t_R, the public information is revealed. Second, the market maker posts an ask and a bid price. Finally, a trader arrives at the market and buys, sells or leaves. Afterward, the next period starts.

Aside from the existence of public news, the set-up of the model closely relates to Glosten and Milgrom (1985) and Easley et al. (1997). Specifically, if $\gamma_i = 0\ \forall i$ the model is the same as Glosten and Milgrom (1985)'s model. On the other hand, if $\gamma_i = 0\ \forall i$ with probability α and $\gamma_i \to \infty\ \forall i$ with probability $1-\alpha$; the model becomes equivalent to Easley et al. (1997)'s model. To illustrate how this small change affects the model conclusions, I first present the agents' best responses without the release of public news; and then, I consider the case in which μ is revealed.

1.1 Best Responses without Public News

Since every informed investor receives the same signal, in equilibrium, their best response does not depend on past quotes and transactions. Instead, they take current quotes as given and maximize their utility conditional on ω. As a result, their reaction function is given by:[8]

$$d_i^*(A,B) = \begin{cases} 1 & \text{if } \gamma_i < \dfrac{\omega - A}{\sigma_\mu + \sigma_\varepsilon} \\ -1 & \text{if } \gamma_i < \dfrac{B - \omega}{\sigma_\mu + \sigma_\varepsilon} \\ 0 & \text{otherwise} \end{cases}$$

Informed investors buy if their information rents, $\omega - A$, cover their disutility from risk, $\gamma_i (\sigma_\mu + \sigma_\varepsilon)$. Therefore, the participation of informed traders highly depends on the quotes. If bid and ask prices are high, informed investors are less likely to buy; thus, buys do not carry much information. On the other hand, no-trades, and sales provide much more information as they are more likely to be from an informed trader.

Meanwhile, the market maker sets a bid and ask price to break even in expectation, conditional on the whole sequence of transactions up to time t. Precisely, her profit function is given by

$$U_t^{MM}(A,B) = \mathbb{E}\left[\mathbb{1}\{d_t = 1\}(v - B) + \mathbb{1}\{d_t = -1\}(A - v)\,|\mathcal{H}_t\right]$$

[8] I do not consider the case in which $B > A$, as it does not occur in equilibrium.

where $\mathcal{H}_t$ denotes the sequence of transactions up to time t. The assumption of perfect competition implies that equilibrium prices satisfy:

$$A_t^* = -\sigma_\omega + 2\Pi_t^+ \sigma_\omega, \text{ and } B_t^* = -\sigma_\omega + 2\Pi_t^- \sigma_\omega$$

where Π_t^+ (Π_t^-) are the market maker beliefs that $\omega = \sigma_\omega$ conditional on all the trades up to t and a buy (sell) order $d_t = 1$ ($d_t = -1$).[9] Π_t^+ can easily be obtained recursively, even if it depends on the whole sequence of buys and sales.

In previous models, the probability of informed trading is constant and symmetric during all trading periods. In my model, however, this probability changes as time passes by, and it is different if we consider the buy or the sell side. To illustrate the dynamics of traders' composition, consider the probability of informed buying, $PIB = \delta F\left(\frac{\sigma_\omega - A_t^*}{\sigma_\mu + \sigma_\epsilon}\right)$. If the true state is $\omega = \sigma_\omega$, the market maker learns through trades and the information rents $(\omega - A_t^*)$ reduce, which on average, leads to a reduction in informed trading. Additionally, at each point in time A_t^* adjusts inducing variation in the PIB.

The absence of symmetry is a direct implication of the endogenous composition of traders and creates two important differences with respect to previous models. First, the lack of trading signals the value of the asset. When prices are close to σ_ω, the information rents from buying are small and most informed traders exit the market without transacting if the value of the asset is high. The market maker internalizes this behavior and raises prices even more. Likewise, the opposite situation happens with low prices. This finding relates to Easley et al. (1997) which shows that lack of trading can signal the absence of private information. My model takes one step further and proves that it can also signal the sign of this information.

Secondly, the order of trades is important, hence the number of buys and sales is not enough to characterize prices. To exemplify why the order is important, consider the change in prices after a buy followed by a sale and vice-versa which corresponds to the middle branches of the tree in Figure 1. In the former situation, the sale takes place when prices are high due to the prior buy, therefore the seller is likely to be an informed trader. In the opposite case, one sale after a buy, the initial decrease in prices is lower than the

[9] To avoid extra notation, I denote by d, the current order (d_t), and the reaction of informed investors (d_i). Along the paper, I use stars to represent equilibrium quantities and actions.

previous one as prices are lower. The posterior increase, however, is greater by cause of lower prices. As a result, prices after the two transactions are lower in the first case than in the later. Instead, most existing models predict that both situations should lead to the same prices; graphically, Figure 1 would become a recombining tree.

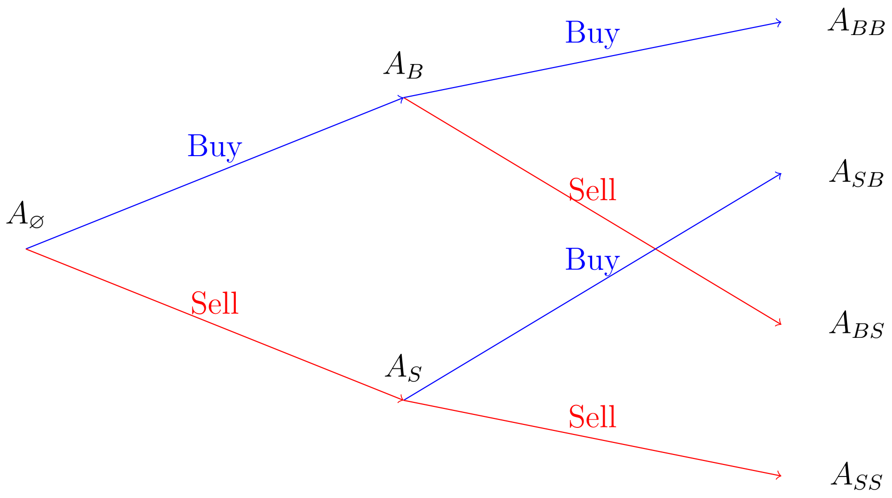

Figure 1: Example of any two periods in the model. A represent ask prices and they are order from the highest (top) to the lowest (bottom). $A_\varnothing$ is the ask price at any period, A_B (A_S) is the ask price one period after $A_\varnothing$ if the last transaction is a buy (sell). Finally, A_{hj} is the ask price two periods from the beginning if the first transaction was a buy ($h = B$) or a sell ($h = S$) and the last one was a buy ($j = B$) or a sell ($j = S$).

Finally, note that agents are not forward-looking. Accordingly, the model without public news is equivalent to releasing a public announcement at $T + 1$, after the market closes. Hence, I describe the equilibrium and its properties for the more general case in which there is a public announcement at t_R.

1.2 Equilibrium with Public News

An equilibrium in this model is given by the triples $\{d_i^*(A_t^*, B_t^*, \omega), A^*(\mathcal{H}_t), B^*(\mathcal{H}_t)\}_0^{t_{R-1}}$, and $\{d_i^*(A_t^*, B_t^*, \mu, \omega), A^*(\mathcal{H}_t, \mu), B^*(\mathcal{H}_t, \mu)\}_{t_R}^T$ where d_i^* is the informed agents' best response and the market maker obtains zero profits in expectation with quotes equal to A^*, and B^*. Importantly, quotes depend on the whole sequence of transactions ($\mathcal{H}_t$). Proposition 1 describes the equilibrium prices and strategies. Additionally, Corollary 1 shows that the spread is always positive, and informed traders do not trade against their information.

Proposition 1. *An equilibrium exists and it is unique. Moreover, it is given by:*

If $t < t_R$,

$$d_i^*(A^*, B^*, \omega) = \begin{cases} 1 & \text{if } \gamma_i < \dfrac{\omega - A^*}{\sigma_\mu + \sigma_\varepsilon} \\ -1 & \text{if } \gamma_i < \dfrac{B^* - \omega}{\sigma_\mu + \sigma_\varepsilon} \\ 0 & \text{otherwise} \end{cases}$$

$$A_t^* = -\sigma_\omega + 2\frac{\frac{1}{2}(1-\delta) + F\left(\frac{\sigma_\omega - A_t^*}{\sigma_\mu + \sigma_\varepsilon}\right)}{\frac{1}{2}(1-\delta) + \Pi_t F\left(\frac{\sigma_\omega - A_t^*}{\sigma_\mu + \sigma_\varepsilon}\right)}\Pi_t \, \sigma_\omega$$

$$B_t^* = -\sigma_\omega + \frac{(1-\delta)}{\frac{1}{2}(1-\delta) + (1-\Pi_t)F\left(\frac{B_t^* + \sigma_\omega}{\sigma_\mu + \sigma_\varepsilon}\right)}\Pi_t \, \sigma_\omega$$

If $t \geq t_R$,

$$d_i^*(A^*, B^*, \mu, \omega) = \begin{cases} 1 & \text{if } \gamma_i < \dfrac{\mu + \omega - A^*}{\sigma_\varepsilon} \\ -1 & \text{if } \gamma_i < \dfrac{B^* - \mu - \omega}{\sigma_\varepsilon} \\ 0 & \text{otherwise} \end{cases}$$

$$A_t^* = \mu - \sigma_\omega + 2\frac{\frac{1}{2}(1-\delta) + F\left(\frac{\mu + \sigma_\omega - A_t^*}{\sigma_\varepsilon}\right)}{\frac{1}{2}(1-\delta) + \Pi_t F\left(\frac{\mu + \sigma_\omega - A_t^*}{\sigma_\varepsilon}\right)}\Pi_t \, \sigma_\omega$$

$$B_t^* = \mu - \sigma_\omega + \frac{(1-\delta)}{\frac{1}{2}(1-\delta) + (1-\Pi_t)F\left(\frac{B_t^* - \mu + \sigma_\omega}{\sigma_\varepsilon}\right)}\Pi_t \, \sigma_\omega$$

where Π_t *are the market maker beliefs that* $\omega = \sigma_\omega$ *conditional on all the trades up to* t.

Corollary 1.

$$\sigma_\omega > A_t^* > B_t^* > -\sigma_\omega \text{ if } t < t_R \text{ and } \mu + \sigma_\omega > A_t^* > B_t^* > \mu - \sigma_\omega \text{ if } t \geq t_R.$$

Endogenous participation generates two effects when the market maker increases the spreads. First, she earns higher profits because of the higher price. Second, her profits increase because some informed agents leave the market. The latter effect is the contribution of this model and it plays an important role when the public news is published. At that point in time, the uncertainty about the asset value reduces, and more informed traders participate. Therefore, the order flow becomes more informative which forces the market maker to charge a wider spread.

In terms of comparative statics, for a given sequence $\mathcal{H}_t$, a higher private information volatility, σ_ω, generates higher spreads, and higher volume, as information rents are higher. Similarly, an increase in the proportion of informed agents δ leads to a rise in the spreads but decreases volume since noise traders always transact. Instead, an increment of σ_ε decreases spreads and volume since informed agents are less likely to participate.

1.3 Model Predictions

Since the empirical part relies on a difference-in-difference methodology, to create proper hypotheses, I compare the model predictions if there is no public news (labeled with a subscript 0) against the predictions with public news (labeled with subscript 1), keeping the realized variables constant.

Figure 2 sketches the main mechanism of the model, when $\omega = \sigma_\omega$ and $d_{t_R-1} = -1$. Before the time of release, informed agents whose risk-aversion coefficient is above $\underline{\gamma}$ leave the market without trading, regardless if μ will be revealed or not. At t_R the two models differ. If a public signal does not exist, then the market maker decreases the ask price because, after observing a sale ($d_{tR} = -1$), she updates negatively her beliefs about ω. This price movement increases the informational rents, thus fosters informed participation and every trader with risk-aversion below $\hat{\gamma}$ buys. If traders observe μ, we have an additional effect; the risk that informed traders face decreases which leads to even a more informative order flow. As a consequence, adverse selection costs rise which leads the market maker to set a wider spread, thus diminishing the profits of informed agents. Nonetheless, in equilibrium every investor below $\overline{\gamma} > \hat{\gamma}$ participates in the market.[10]

[10] The order of the different thresholds is a direct result of Proposition 3

The boost in participation creates a complementarity between public and private news; specifically, introducing a public signal increases the speed at which private information is incorporated into prices.

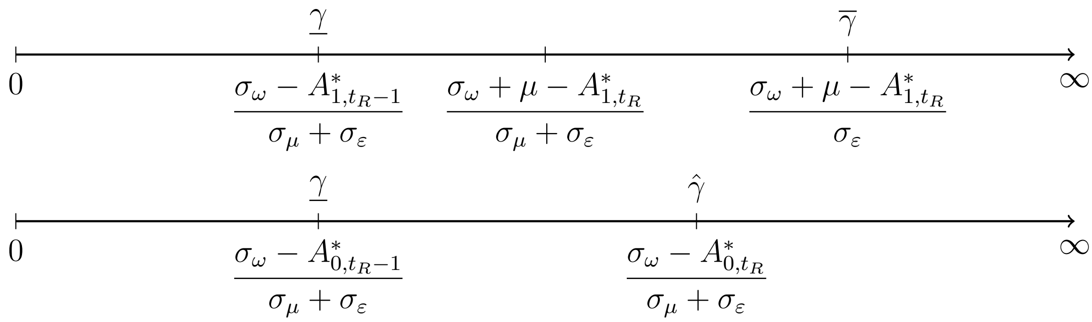

Figure 2: Mechanism. This figure sketches the main mechanism in the model when $\omega = \sigma_\omega$ and $d_{tR-1} = 1$. Both lines represent the support of the risk-aversion parameter. The upper line depicts the relevant thresholds under the existence of a public signal whereas the bottom line consider the case without one.

Along the paper, I consider three different variables that I observe in the data: midpoint returns, bid-ask spread, and volume. The first one is a measure about price dynamics and is defined as $m_t - m_{t-1}$ where $m_t = \frac{A_t + B_t}{2}$. Proposition 2 characterizes the dynamics of m_t in both scenarios, with and without public news. Consistent with the semi-strong market hypothesis, the midpoint rises by μ as soon as the public information is disclosed and remains constant afterward. In contrast, risk aversion and asymmetric information do not have any effect on the midpoint.

Proposition 2. *(Midpoint) In equilibrium, the expected midpoint is given by:*

$$\mathbb{E}_{\mathcal{H}_t}(m^*_{1t}) = \mathbb{E}_{\mathcal{H}_t}(m^*_{0t}) = 0 \textit{ if } t < t_R$$

$$\mathbb{E}_{\mathcal{H}_t}(m^*_{1t}) = \mu \textit{ and } \mathbb{E}_{\mathcal{H}_t}(m^*_{0t}) = 0 \textit{ if } t \geq t_R$$

The second variable I consider is the bid-ask spread, which is defined as $A_t - B_t$. Intuitively, it is a buffer the market maker needs to have in order to maintain zero profits under the presence of informed investors. Proposition 3 states that as soon as μ becomes common knowledge, the bid-ask spread increases, moreover, this rise is independent of μ as it is a result of a composition effect. If informed investors do not know μ they are less willing to act on their information since it entails a high risk. Once they observe μ these investors start trading. To maintain zero profits in expectation, the market maker reacts widening the spread. After the release of the signal, the market maker learns quicker about ω from trades since more informed traders participate. Hence, the bid-ask spread decreases and converges to the one in the model without news.

Proposition 3. *(Bid-ask spread) At the release time, $t = t_R$, the bid-ask spread satisfies:*

$$A^*_{1,t_R} - B^*_{1,t_R} > A^*_{0,t_R} - B^*_{0,t_R}$$

Moreover, the difference $A^*_{1,t_R} - B^*_{1,t_R} - (A^*_{0,t_R} - B^*_{0,t_R})$, does not depend on μ. Additionally, in the limit, the spread in both scenarios converges. Precisely,

$$lim_{T\to\infty}\, A^*_{1,T} - B^*_{1,T} = lim_{T\to\infty}\, A^*_{0,T} - B^*_{0,T} = 0 \tag{1}$$

The last variable I analyze is volume, $\mathbb{E}(|d_t|)$, which is a relevant measure of liquidity. Proposition 4 specifies that volume increases at t_R and this effect decreases as the gains from trade decrease. Similar to spreads, the signal content does not affect the increment at the time of the release, nor the dynamics afterward.

Proposition 4. *(Volume) When the public information is released, volume the days without news is higher than the days with news:*

$$\mathbb{E}\left(|d^*_{1,t_R}|\right) > \mathbb{E}\left(|d^*_{0,t_R}|\right)$$

Further, the difference does not depend on μ. Moreover, it disappears in the limit:

$$lim_{T\to\infty}\mathbb{E}\left(|d^*_{1,T}|\right) = lim_{T\to\infty}\mathbb{E}\left(|d^*_{0,T}|\right) = (1-\delta) \tag{2}$$

1.4 Extensions

In this section, I extend the model along several dimensions to assess the importance of the main assumptions. Although it is a standard assumption in the literature, the insensitivity of noise traders to prices might have an effect on the empirical predictions. The first extension illustrates that even if noise traders might decide not to trade at unfavorable prices, as long as their private value is high enough, the results of the model do not change. Another concern is the asymmetry between market makers and informed traders.

While the former are risk-neutral by assumption, the later are heterogeneous and risk-averse. The second extension shows that this assumption is actually the result of perfect competition between heterogeneous and risk-averse market makers. The third extension addresses the implications of the independence assumption between the public and the private signal by considering that the public signal may reveal the private information. If the probability of losing the private information is low the baseline results hold; in contrast, if this probability is very high volume and spreads might decrease at the time of release. Ultimately, it is an empirical question.

1.4.1 Sensitive Noise Traders

The baseline model highlights the behavior of informed agents who maximize utility and trade with a market maker that makes zero profits in expectation due to perfect competition. In this extension, I model explicitly the third type of agents, the noise traders.

Noise traders represent investors that obtain a private value for the asset besides the common value. Passive funds that need to rebalance their portfolio, asset managers with hedging necessities, or investors looking for liquidity are examples of noise traders. When taking the decision between trading or not, these investors compare the private benefit they obtain with the price of the asset. Hence, I characterize noise traders utility as $U(d_i) = (\mathbb{E}(v) - p(d_i) + \theta\tau_i)\, d_i$ where τ_i is the private value for investor i. To simplify the exposition, I assume τ has a symmetric distribution around 0. Precisely, with probability $\frac{1}{2}$ the private value is positive with c.d.f and p.d.f given by $\Phi(\tau)$ and $\phi(\tau)$; otherwise τ is negative with c.d.f equal to $1 - \Phi(-\tau)$ and p.d.f $\phi(-\tau)$. Note that I implicitly assumed $\Phi(x) = 0$ if $x < 0$; moreover similar to the baseline case I assume that, at least, some noise traders are present in the market regardless of the quotes: $\Phi(0) > 0$.

Given equilibrium prices A_t^* and B_t^*, the best response of a noise trader is characterized by:[11]

$$d_i = 1 \text{ (Buy) if } \tau > \frac{A_t^* - \mathbb{E}(v|\mathcal{H}_t)}{\theta},$$
$$d_i = -1 \text{ (Sell) if } -\tau > \frac{\mathbb{E}(v|\mathcal{H}_t) - B_t^*}{\theta},$$
$$\text{and } d_i = 0 \text{ (Leave) otherwise}$$

[11] For simplicity, I do not consider $A_t^* < B_t^*$ as it does never hold in equilibrium.

A noise investor buys if the liquidity cost $A_t^* - \mathbb{E}(v|\mathcal{H}_t)$ is smaller than her private benefit $\theta\tau$. As a result, the probability that a noise trader buys is $\frac{1}{2}\left(1 - \Phi\left(\frac{A_t^* - \mathbb{E}(v|\mathcal{H}_t)}{\theta}\right)\right)$, and the probability that she sells is $\frac{1}{2}\left(1 - \Phi\left(\frac{\mathbb{E}(v|\mathcal{H}_t) - B_t^*}{\theta}\right)\right)$. The specific case of $\theta \to \infty$ corresponds to the baseline model.

The best response of the informed traders is exactly the same as in the baseline model. The pricing rule of the market maker, however, has to take into account the price sensitivity of noise traders. In the baseline case, if the market maker widens the spread, she increase her profits because she earns more from the noise traders and, at the same time, some informed traders leave the market. In contrast, if noise traders are sensitive to prices they might leave the market as well, which erodes the market maker's profit. To maintain the existence of an equilibrium, this latter channel cannot play a major role; that is, the price sensitivity of noise traders cannot be much higher than the one of informed agents. Proposition 5 formalizes the condition that guarantees the existence of equilibrium.

Proposition 5. *An equilibrium exists and it is unique at every t if*

$$\frac{f(x)}{F(x)(\sigma_\mu + \sigma_\epsilon)} - \frac{\phi(y)}{(1 - \Phi(y))\,\theta} > 0$$

for all $x \in \left(0, \frac{2\sigma_\omega}{\sigma_\epsilon + \sigma_\mu}\right)$ and $y \in \left(0, \frac{2\sigma_\omega}{\theta}\right)$ where $f(x)$ is the p.d.f of γ.[12]

The condition in Proposition 5 is convoluted since it links the risk-aversion distribution with the distribution of private values. Yet, it is the a sufficient condition for the market maker profits to increase as the bid-ask spread increases. If this is not satisfied we have two cases. Either the bid-ask spread decreases the market maker profits, or it increases or decreases the profits depending on the quote level. In the former case, we do not have an equilibrium as the market maker never makes zero profits in expectation because there are not enough noise traders. The latter case can lead to non-existence or multiple equilibria depending on the parameters. Consider $A^\dagger - B^\dagger$ is a spread that makes the market maker

[12] In the simplifying case that $\gamma \sim Exp(\lambda_\gamma)$ and $\tau \sim Exp(\lambda_\tau)$, a sufficient condition is:

$$\lambda_\gamma < \frac{\sigma_\epsilon + \sigma_\mu}{2\sigma_\omega} log\left(\frac{\theta}{\lambda_\tau(\sigma_\epsilon + \sigma_\mu)} + 1\right)$$

break even. As the bid-ask spread increases, noise traders leave the market which reduces profits but also informed traders leave the market. If the condition in Proposition 5 is not satisfied there can be another pair of quotes $A^{\ddagger}$ and $B^{\ddagger}$ with the same proportion of informed to uninformed, hence the same profits but a lower volume.

The main model predictions hold in the new set-up. Regarding the midpoint, the market maker incorporates immediately the public information, hence she changes the midpoint by μ at the time of the release. Meanwhile, the bid-ask spread increases because more informed agents enter the market. In terms of volume, however, the prediction depends on the model parameters. Whereas a lower uncertainty encourages informed investors to trade, the widening of the bid-ask spread discourages some noise traders to stay in the market which affects the total volume. The net effect of these two forces depends on the sensitivity of noise traders but unfortunately, the explicit condition cannot be obtained in closed form.

To illustrate the result I solve the model under a precise parameter configuration. Precisely, I assume that the risk aversion parameter (γ) and the private value (τ) have an exponential distribution with expectation equal to 5. I fix the variance of the private information, the public information, and the residual noise ($\sigma_\omega^2, \sigma_\mu, \sigma_\varepsilon$) to $0.5, 1$ and 0.25 respectively, and the release time (t_R) to 31. The realization of the public signal is set to 1. Figure 3 presents the results for different degrees of sensitivity. The blue solid line corresponds to the baseline case ($\theta \to \infty$) whereas the red dashed line and the black circles are the solution to the extended model with $\theta = 0.5$ and $\theta = 0.25$, respectively. In terms of magnitude, the former implies that the private value is on average twice the standard deviation of the value of the asset, while the second parameter reduces it to one standard deviation.

Figure 3 shows that the price sensitivity of noise traders does not affect the midpoint returns but it modifies the effect that public news has on the bid-ask spread and volume. The spread increases at the time of the release; furthermore, this effect is amplified when private values are low because the market maker needs to charge more to the noise traders that remain in the market in order to compensate for those who leave. In contrast, the level of price sensitivity of noise traders lessens the surge of trading and volume might

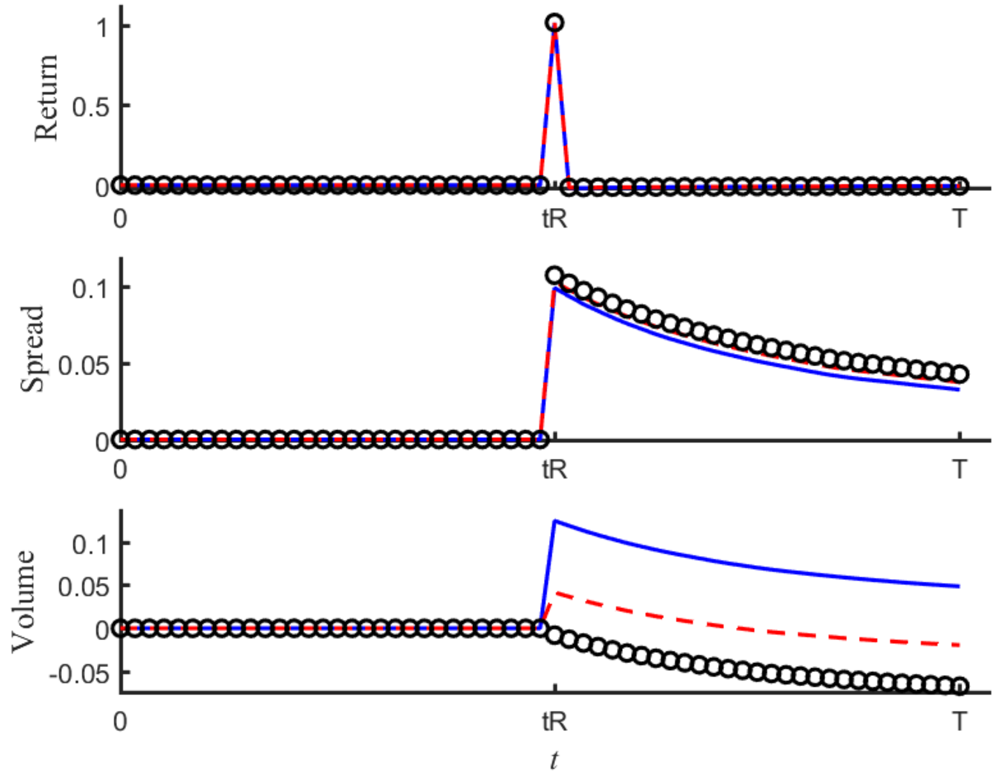

Figure 3: Sensitive Noise Traders. The average path of return, spread and volume over 10,000 simulations. Each path consists of 60 trading rounds, and the public news is released at tR=31 and equals 1 ($\mu = 1$). The risk-aversion and private values are assume to be distributed according to an $Exp\,(5)$. I fix the variance of the private information, the public information, and the residual noise $(\sigma^2_\omega, \sigma_\mu, \sigma_\varepsilon)$ to $0.25, 1$ and 0.25 respectively. Following Cipriani and Guarino (2014), I set $\delta = 40\%$. The blue line considers $\theta \to \infty$ whereas the red dashed line and the black circles assume $\theta = 0.5$ and $\theta = 0.25$ respectively.

even decrease after the signal realization.

1.4.2 Risk-averse market makers

The market maker in the baseline model makes zero profits in expectation. In this section I extend the baseline model by considering that there is a pool of market makers with the following utility function:

$$U_t^{MM}(A, B) = \mathbb{E}\left[\mathbb{1}\{d_t = 1\}\,(v - B) + \mathbb{1}\{d_t = -1\}\,(A - v) - \alpha_i\, Var_t\,(v|\mathcal{H}_t, d_t)\,\mathbb{1}\{d \neq 0\}|\mathcal{H}_t\right]$$

where $\mathbb{1}\{\cdot\}$ is a dummy variable that takes value 1 if the expression within the brackets is true and 0 otherwise. The utility is a mean-variance utility function which is consistent with the objective function of the informed agents. Similar to informed agents, market makers are heterogeneous in their risk aversion whose distribution is characterized by the parameter α with c.d.f $G(\cdot)$. Similar to $F(\cdot)$, I consider $G(c) > 0$ for all $c > 0$.

Lemma 1 establishes that only the market makers with the lowest risk aversion participate and characterizes the pricing decision of these agents.

Lemma 1. *If market makers compete à la Bertrand, transactions will take place at the following prices:*

$$A_t^* = \mathbb{E}\left(v|\mathcal{H}_t, d=1\right), \quad B_t^* = \mathbb{E}\left(v|\mathcal{H}_t, d=-1\right)$$

and only market makers with $\alpha_i = 0$ participate in the market.

Risk-averse market makers are not competitive as they require higher prices to be compensated for the risk. As a result, the relevant market maker is characterized by risk-neutrality, and this extension is equivalent to the baseline model.

1.4.3 Correlated Signals

In the baseline model, public information is assumed to be independent of the private information. Sometimes this might not be the case, and the public signal discloses the private information held by some traders. In this extension, I consider the same baseline model but at the time of the release, besides μ, the market maker observes ω with probability ρ.

This modification of the information structure does not change the optimal decision of the informed agents before the release of the public information as they are not forward-looking. Intuitively, if an informed agent has an opportunity to obtain positive utility, she would do so regardless if her information will be revealed afterward. After the public signal realizes, the behavior of traders depends on the disclosure of information. With probability ρ, informed agents lose their advantage and leave the market. As a consequence, volume decreases to $1-\delta$ and the spread decreases to 0. Meanwhile, with probability $1-\rho$ the informational advantage remains exactly the same as in the baseline case.[13]

[13] The assumption of perfect competition among market makers forces them to reveal (through quotes) that they know ω. Hence, assuming that informed agents realize that they have lost their advantage or not leads to the same results.

I focus on the results at t_R. To ease the exposition, I define $\Delta Volume^B$, $\Delta BidAsk^B$ and $\Delta midpoint^B$ as the difference at $t = t_R$ between the baseline model with and without public news. Following the arguments above, proposition 6 characterizes the same differences under this new set-up.

Proposition 6. *When we consider the possibility that private information might be disclosed by the public signal, the effects of the presence of public news at the time of the release on the midpoint coincides with those of the baseline model*

$$\Delta Midpoint^{CS} = \Delta Midpoint^B = \mu.$$

However, the effect on bid-ask spreads is smaller than in the baseline case, and it might even have the opposite sign,

$$\Delta BidAsk^{CS} = (1-\rho)\Delta BidAsk^B - \rho\left(A^*_{0,t_R} - B^*_{0,t_R}\right).$$

Volume follows a similar pattern than spreads,

$$\Delta Volume^{CS} = (1-\rho)\Delta Volume^B - \rho\frac{\delta}{2}\left(F\left(\frac{\sigma_\omega - A^*_{0,t_R}}{\sigma_\epsilon + \sigma_\mu}\right) + F\left(\frac{B^*_{0,t_R} + \sigma_\omega}{\sigma_\epsilon + \sigma_\mu}\right)\right).$$

The predictions about the expected midpoint are exactly the same as in the baseline case. The effect on spreads and volume is the net effect of two forces. On the one hand, the public signal reduces uncertainty which boosts participation of informed traders. On the other hand, it might also eliminate their private information forcing them to leave the market. The overall effect depends on the relative strength of the latter channel which is measured by ρ. If ρ is high enough, volume and spreads decrease when a public signal is released; otherwise, the predictions of the baseline model remain valid. Therefore, the dominant channel can only be found empirically by observing if spreads increase or decrease after an announcement.

2 Institutional Framework and Data

Every Wednesday at 10:30 a.m., the Energy Information Administration posts the Weekly Petroleum Status Report which contains information about the stock of oil stored in the US by geographical area and product (crude, gasoline, etc.). The report contains a press note highlighting the main figures: oil refined, imports and oil inventories, a table that summarizes the data, and several spreadsheets which contain the disaggregated information. All these documents are available online and they are easily machine readable as the format is constant across weeks.[14]

The information in the report is the official and the only public data available. However, there are private companies that provide some data on inventories before Wednesdays. For instance, a data service company relies on drones to measure the amount of stock inside the tanks and provide the information to its clients on Mondays at 10:00 a.m. While this information might be valuable, the official report remains the main driver of oil price movements during trading hours. To see this, I plot in Figure 4 the percentage of the weekly intraday variance of the nearest-to-maturity oil future for each minute and weekday. In line with the previous statement, the maximum intraday variance is concentrated on the exact time of the report release. In fact, the pattern is mainly flat across days and minutes except on Wednesday at 10:30 a.m, and the closing of the open outcry at Chicago Mercantile Exchange at 2:30 p.m.

Aside from the current oil stock level, the Energy Information Administration also makes the historical ones available which constitute the first source of data in the paper. From the reports, I use the weekly differences in the total stock of oil, excluding the strategic petroleum reserve, to quantify the information that the public news convey. Note that the choice of oil product and location is irrelevant for the results since all of them are close substitutes; thus, their inventories are highly correlated. At the same time, an ideal measure of news should not be predictable. Thus, I extract the unpredictable component of the increase in inventories by regressing the change in oil stocks on its lag and week-of-the-year dummy variables. The regression residuals are my measure of news which does not contain the predictable component due to autocorrelation or seasonality. Furthermore, to ease the interpretation, I normalize it to have a standard deviation equal to one. I label the resulting variable as ΔInv.[15] While there might be other data sources that help to forecast changes in inventory, for estimation purposes I only assume that my measure of unpredictable information correlates positively with the actual unpredictable

[14]The report can be found at http://ir.eia.gov/wpsr.

[15]Results in the paper are maintained if I consider raw inventory variations.

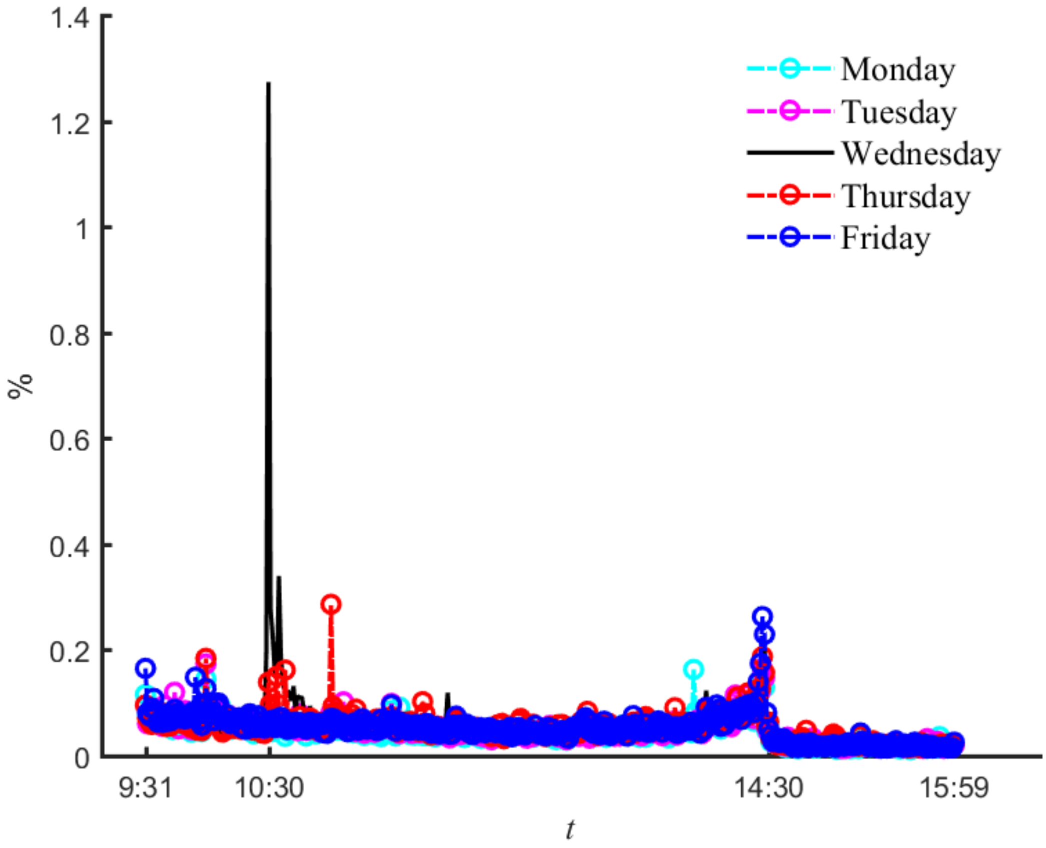

Figure 4: Intraday Variance Decomposition. This figure present the variance of the nearest-maturity future by minute normalized to sum to 100% every week. Overnight returns are not considered.

component.

Supplementary to this data, the effect of the report should be very different on industries that highly depend on oil versus those who do not; therefore, it is important to differentiate these two groups. To classify firms as oil or non-oil, I obtain how many dollars of this commodity a particular industry needs in order to produce one dollar of total output using the Bureau of Economic Analysis Input-Output tables from 1999. In general, industries require between three and ten cents of oil; however, there are two industries in my sample that stand out: oil extraction firms (0.76$) and refineries (0.38$). Consequently, those make up my definition of the oil sector. The decision to rely on 1999 data is driven by the fact that it is the last release that it is possible to link to historical SIC codes available in Center for Research in Security Prices (CRSP) database. The index remains very similar with modern tables even if it is subject to reclassification issues.

Lastly, I construct market variables from the Trade and Quote dataset which includes each transaction price and size, as well as the best bid and ask quotes. My available sample includes data on 50 randomly chosen firms, stratified in two volume buckets, from January 2007 to June 2013.[16] Regarding data processing, I apply the filters proposed by Holden and Jacobsen (2014), I aggregate the variables inside a minute, and I winsorize every variable at the 1% in each minute. As a result, each observation in my dataset is a minute inside a given day for a specific firm. Additionally, I drop from the sample minutes without transactions.

From the final sample, I construct three main variables: midpoint returns, proportional effective bid-ask spread, and volume. The midpoint returns measure changes in the mean beliefs about the value of the asset and is defined as: $r_t = log(\bar{m}_t) - log(\bar{m}_{t-1})$, where $\bar{m}_t = \frac{1}{K_t}\sum_{\tau}\frac{A_\tau + B_\tau}{2}$, A_τ and B_τ are the ask and bid prices posted at the time of transaction τ, and K_t is the total number of transactions in minute t. The second variable I analyze, the spread, is an indicator of transaction costs and has been associated with the degree of information asymmetry (see Glosten and Milgrom, 1985). I construct it as $sp_t = \frac{1}{K_t}\sum_{\tau}|\frac{p_\tau - m_\tau}{m_\tau}|$ where p_τ is the actual transaction price. The final variable I consider is volume, which reflects market activity, as it consists of the number of shares traded in a minute.

These variables match the variables considered in the theoretical framework in Propositions 2, 3, and 4. Additionally, through the lenses of the model, ΔInv is similar to μ, t_R corresponds to 10:30 a.m. and ω represents private information that some traders have related to oil prices. In my set-up private information can be thought as insiders who receive some private signal; but it can also be understood as investment companies or departments who process information, such as the worldwide weather, better than most investors. Finally, following Easley et al. (1997), I interpret that the model represents a day, and at the beginning of each day, the game starts again. Therefore, the difference between Wednesdays and any other day of the week equals the difference between the model with and without public news, if everything else is constant.

[16]Firms are divided into two volume buckets according to the data on CRSP at 2005.

3 Empirical Analysis

I test the model hypotheses using a difference-in-difference approach. I consider oil firms as treated and non-oil firms as the control group. At the same time, I compare Wednesdays with the other weekdays. To be precise, I estimate the following equation,

$$y_{i,t} = \mu + \delta_m + \theta_0 Oil_i + \theta_1 \Delta Inv_t + \theta_2 Wed_t + \theta_3 Wed_t \cdot \Delta Inv_t + \\ (\gamma_0 \Delta Inv_t + \gamma_1 Wed_t + \gamma_2 Wed_t \cdot \Delta Inv_t) \cdot Oil_i + \varepsilon_{i,t} \quad (3)$$

where $y_{i,t}$ is the value of the variable of interest for firm i on day t, δ_m are month-year fixed effects, and Oil_i and Wed_t are dummy variables that indicate if the firm belongs to the oil sector and if the day of the week is Wednesday. In order to characterize the complete dynamics, I estimate Equation (3) for every minute independently. Hence, observations are at the daily-firm level. Since ΔInv varies on a weekly basis, and to account for some correlation across weeks, I cluster standard errors at the monthly level.

To identify the effect of the presence of a public announcement, I make three main assumptions: First, non-oil firms are not affected by the oil report. Second, the presence of the report publication only affects Wednesdays, and third, there is no other factor that affects differently oil and non-oil firms at 10:30 a.m. on Wednesdays. The online appendix provides evidence in favor of these assumptions. Table S.3 indicates that the price, bid-ask spread and volume of non-oil firms before and after the time of release are similar; therefore, these firms are a reliable control group. As regards to the control period, Section S.2 compares Wednesdays with each of the other weekdays separately, and results are unaltered regardless of the weekday used as a control. Finally, Section S.5 addresses the issue of an omitted factor by conducting a difference-in-difference-in-difference where the third difference corresponds to the change before and after the EIA started publishing the report on Wednesdays at 10:30 a.m. and the same results prevail.

The key parameters of interest are γ_1 and γ_2. While the first one captures the effect irrespective of the news, the second one measures the reaction to the report's content. To focus the attention on the release, I restrict the sample to the period from 10:00 a.m. to 10:59 a.m. I refer to parameters from different minutes with a superscript. For instance, $\gamma_1^{10:00}$ is the estimate of γ_1 using only data at 10:00 a.m. and variation across days and firms.

The theoretical results state that prices should react instantaneously in the direction of public information. Since a decrease in inventories implies a high μ in terms of the model, the empirical hypothesis is that $\gamma_2^{10:30} < 0$. At any other time different from the release time, the model predicts that price dynamics are equal with and without news; equivalently, $\gamma_2^m = 0$ for every m except $m =$ 10:30. In the upper plot of Figure 5 I depict the estimates of γ_2 and their 95% confidence intervals. They confirm the previous hypotheses. Precisely, prices rise immediately by 5.5 bps if inventories drop by one standard deviation, but they remain constant afterward. Meanwhile, the lower plot depicts the estimates of γ_1. Consistent with the model prediction, the disclosure of a public signal is irrelevant besides its content; equivalently, if the report does not provide new data, $\mu = 0$, prices do not react.

Apart from the empirical support to my model, these results support the assumption that informed investors lack prior information about oil stocks. Otherwise, these investors would buy before an upcoming news of a decrease which would push prices up as the market maker learns from the fundamental. This mechanism creates a negative correlation between the change in inventories and returns before the release that is not consistent with the data.

Regarding the spread, the model predicts an increase after the release, independently on the information content ($\gamma_1^{10:30} > 0$). Moreover, this effect lasts for some periods as the market maker learns about the private signal although, it finally vanishes. Therefore, I hypothesize that $\gamma_1^m > 0$ if $m \geq$ 10:30 but $\gamma_1^m = 0$ if $m <$ 10:30, or $m \gg$ 10:30. The duration of the effect depends on the parameters of the model. The upper plot of Figure 6 shows that the spread increases at 10:30 a.m. by 2.79 bps and it slowly decreases, and it becomes insignificant after thirty minutes. At the same time, in line with the model, the lower plot confirms that the content of the report is irrelevant for the spread.

With respect to volume, the mechanism behind the model implies a strong comovement between the effect on this variable and the spread. Figure 7 present the estimates of Equation (3) with volume as a dependent variable. We observe a similar pattern to the one in the analog figure for the spread (Figure 6). Precisely, volume rises by 32% at

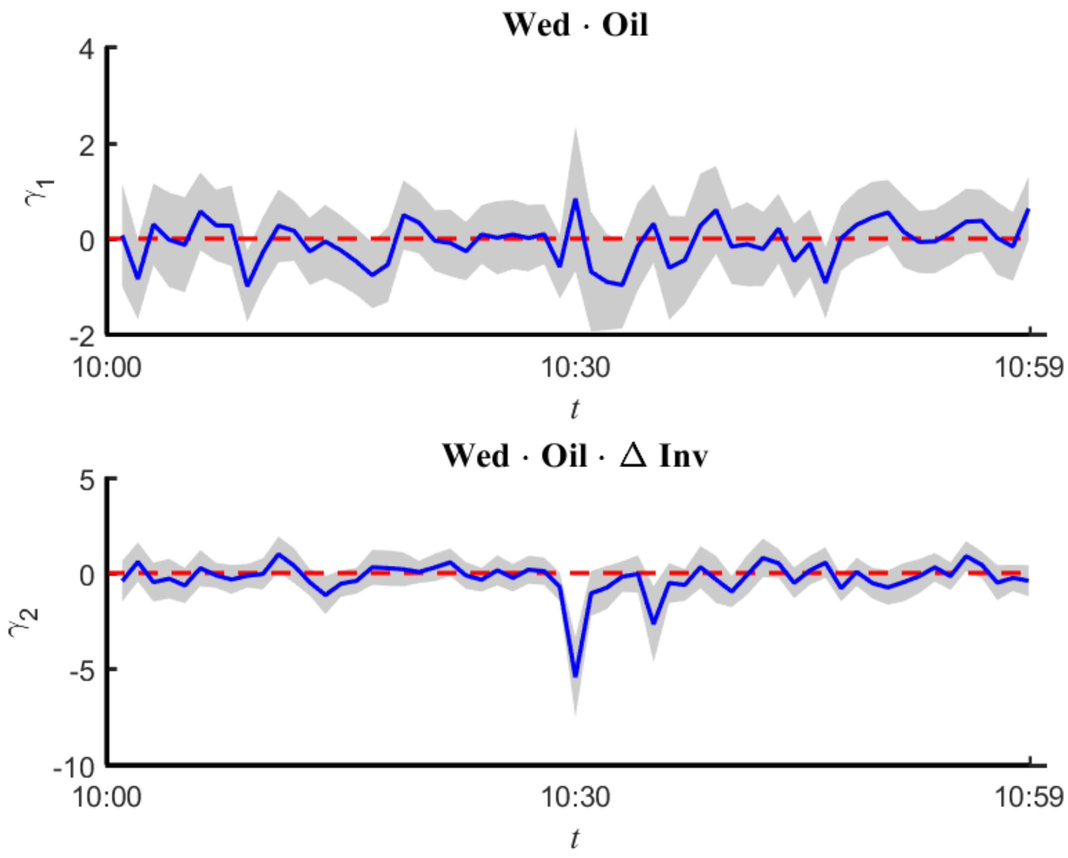

Figure 5: Estimates Return. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (3) using midpoint returns as a dependent variable. I estimate the coefficients independently for every minute exploiting variation across days. The upper plot corresponds to γ_1, while the lower one refers to γ_2. I describe the sample in Section 2. The red dashed line indicates the zero.

the moment of the announcement and it decreases afterward. In addition, the presence of news does not affect market activity before its release. This lack of reaction suggests that traders neither know the information before nor do they strategically defer their trading.[17] Regarding the report's content, it does not affect the increment in volume, or the posterior dynamics, as the model suggests.

Lastly, the model predicts that the case with a news release, eventually, converges to the one without news. I show that this convergence is immediate as regards returns, but it takes some minutes in terms of volume and spreads. To test the limiting results, in

[17] There is a significant decrease in volume just before the release. While it might be because of trading deferral, it becomes insignificant in most of the robustness checks.

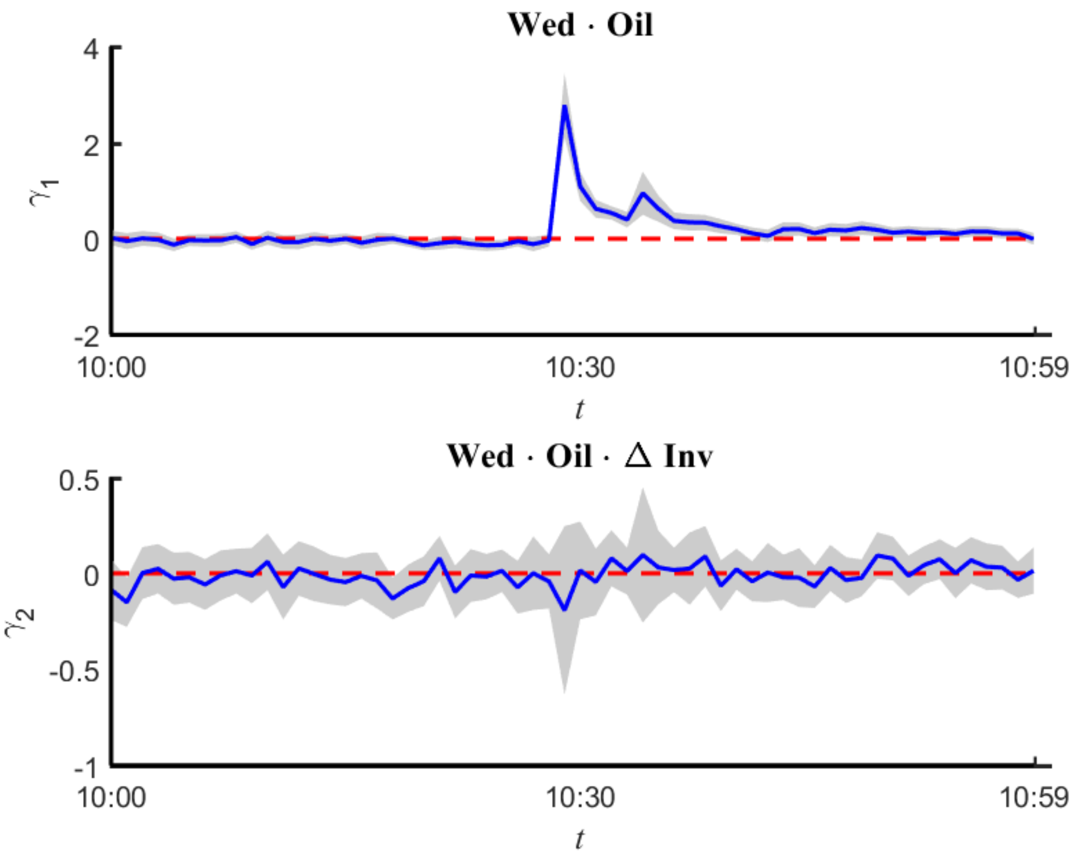

Figure 6: Estimates Spread. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (3) using proportional effective bid-ask spreads as a dependent variable. I estimate the coefficients independently for every minute exploiting variation across days. The upper plot corresponds to γ_1, while the lower one refers to γ_2. I describe the sample in Section 2. The red dashed line indicates the zero.

Figure 8 I plot the estimate for the last hour of the trading day. The graphs confirm the dissipation of the effects since estimates are not significant in all cases. Likewise, these results constitute a placebo test that reinforces the validity of previous results.

3.1 The effect of magnitude

The theoretical model concludes that the effect on spreads and volume should be independent to the content of the report. However, the previous results just control for the change in inventory; instead, the absolute value of the public signal might be the relevant variable as Kim and Verrecchia (1991)'s model predicts. To explore this alternative

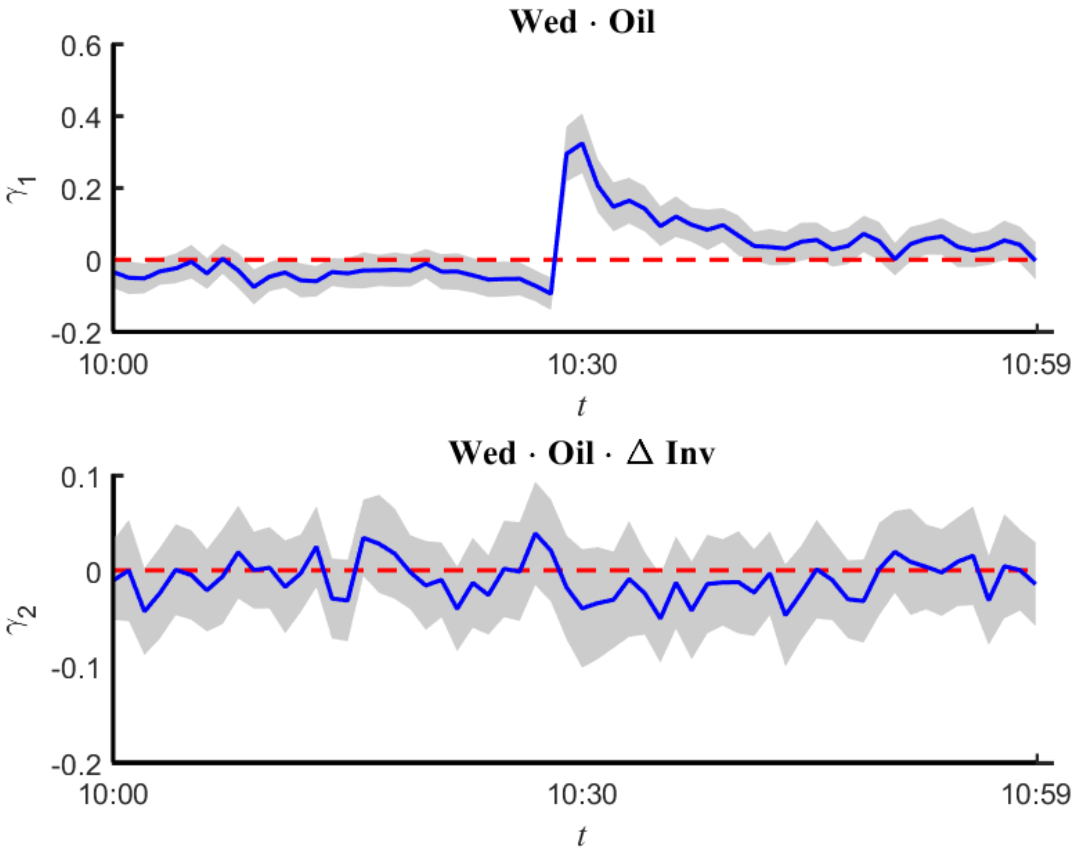

Figure 7: Estimates Volume. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (3) using number of transactions as a dependent variable. I estimate the coefficients independently for every minute exploiting variation across days. The upper plot corresponds to γ_1, while the lower one refers to γ_2. I describe the sample in Section 2. The red dashed line indicates the zero.

channel, I estimate the following model:

$$y_{i,t} = \mu + \delta_t + \theta_0 Oil_i + \theta_1 \left|\Delta Inv_t\right| + \theta_2 Wed_t + \theta_3 Wed_t \cdot \left|\Delta Inv_t\right| + \\ \left(\gamma_0 \left|\Delta Inv_t\right| + \gamma_1 Wed_t + \gamma_2 Wed_t \cdot \left|\Delta Inv_t\right|\right) \cdot Oil_i + \varepsilon_{i,t}. \quad (4)$$

Fixed effects are important in this specification as we have two opposite effects. On the one hand, periods with volatile changes in inventory (high σ_μ) should present bigger effects according to the model. On the other hand, if the model parameters are fixed, the absolute change in inventories should be irrelevant. By the use of month-year dummies, I control for low-frequency variability. Hence, we can interpret the other coefficients as if the parameters do not vary. Therefore, according to the model, $\gamma_2^m = 0$ for all m, and

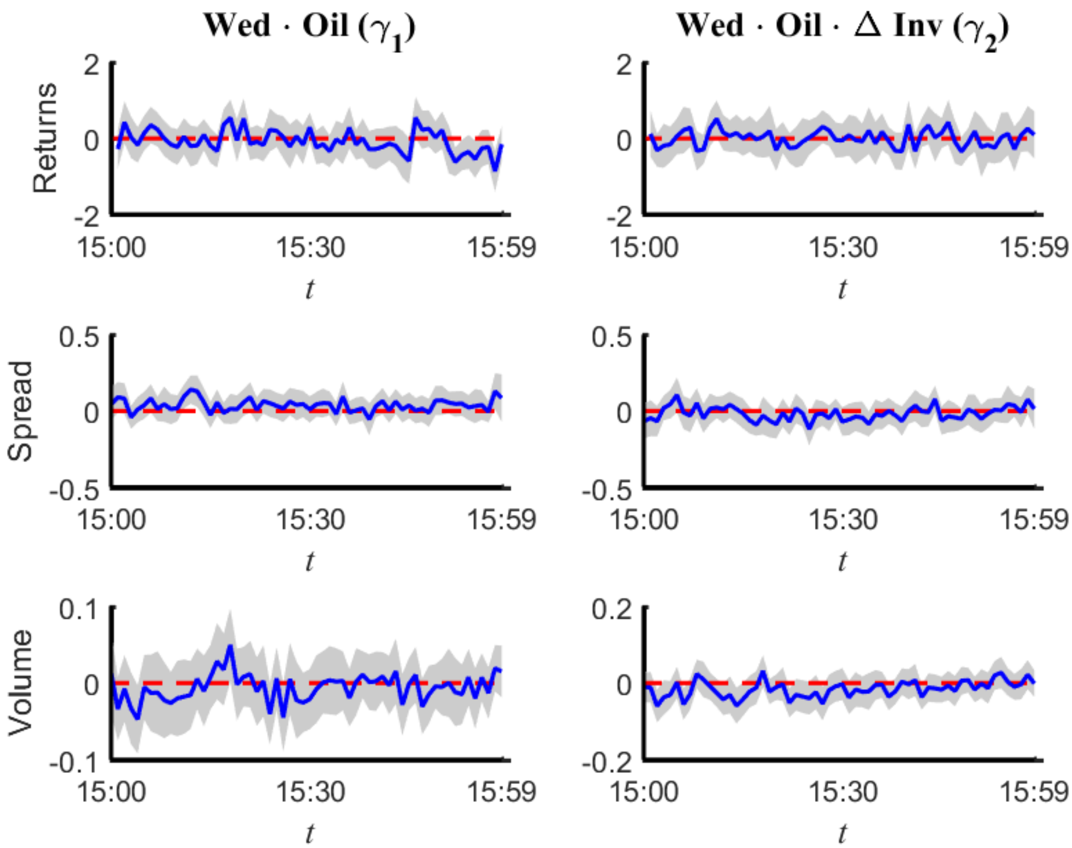

Figure 8: Afternoon Sample. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (3). Each row consider a different dependent variable, from top to bottom: midpoint return, proportional effective bid-ask spread, and number of transactions. I estimate the coefficients independently for every minute exploiting variation across days. The left-hand side plots correspond to γ_1, while the right-hand side ones refers to γ_2. I describe the sample in Section 2. The red dashed line indicates the zero.

the estimate of γ_1 should not change with respect to the same coefficients in equation (3). Figure 9 shows that this coefficient does not change for any of the three market variables. Besides, the absolute value of inventories does not have an effect on the market.

3.2 Asymmetric reaction to news

Along the paper, I assume the effect of the report's information is symmetric. In other words, I consider that an inventory decrease of one standard deviation affects the market as a build-up of the same magnitude but with opposite sign. While this assumption is consistent with the theoretical model, previous literature suggests that it might not be

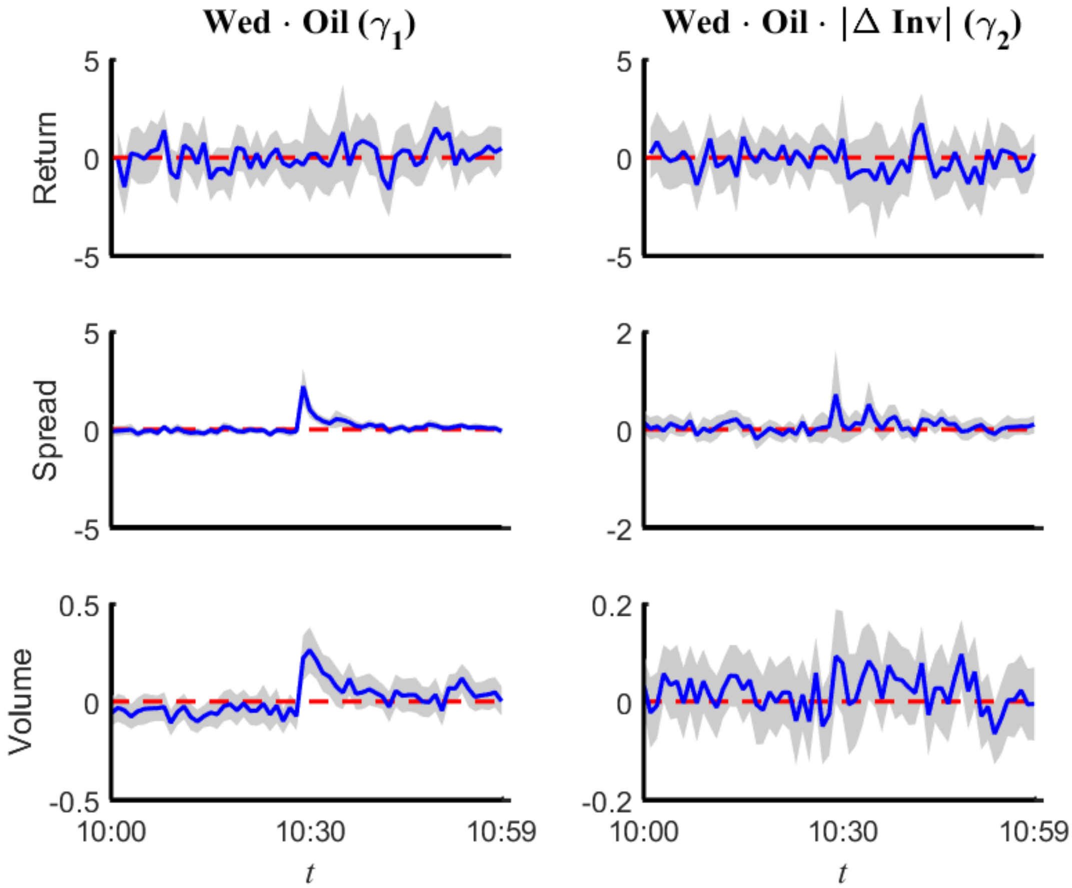

Figure 9: More Noise Traders. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (4). Each row consider a different dependent variable, from top to bottom: midpoint return, proportional effective bid-ask spread, and number of transactions. I estimate the coefficients independently for every minute exploiting variation across days. The left-hand side plots correspond to γ_1, while the right-hand side ones refers to γ_2. I describe the sample in Section 2. The red dashed line indicates the zero.

valid. For instance, traders may suffer from negative bias. In this line, Tetlock (2007) shows that high media pessimism generates downward pressure on market prices but optimism does not have an effect. Another possible channel is attention, as negative news are more salient on media outlets (see Soroka, 2006), it might affect market activity more strongly. In fact, Brown et al. (2009) presents evidence supporting this theory in the case of earnings surprises. More specific to this case, an increase in inventories might contain more, or less, information than a decrease.

To allow for heterogeneous effects depending on the sign, I estimate the following equation:

$$y_{i,t} = \mu + \delta_t + \theta_0 Oil_i + \theta_1^+ \Delta Inv_t^+ + \theta_1^- \Delta Inv_t^- + \left(\theta_2 + \theta_3^+ \Delta Inv_t^+ + \theta_3^- \Delta Inv_t^-\right) \cdot Wed_t +$$
$$\left(\gamma_0^+ \Delta Inv_t^+ + \gamma_0^- \Delta Inv_t^- + \gamma_1 Wed_t + \left(\gamma_2^+ \Delta Inv_t^+ + \gamma_2^- \Delta Inv_t^-\right) \cdot Wed_t\right) \cdot Oil_i + \varepsilon_{i,t} \quad (5)$$

where $\Delta Inv_t^- = min\{0, \Delta Inv_t\}$ captures the effect of positive news whereas $\Delta Inv_t^+ = max\{0, \Delta Inv_t\}$ corresponds to negative news.

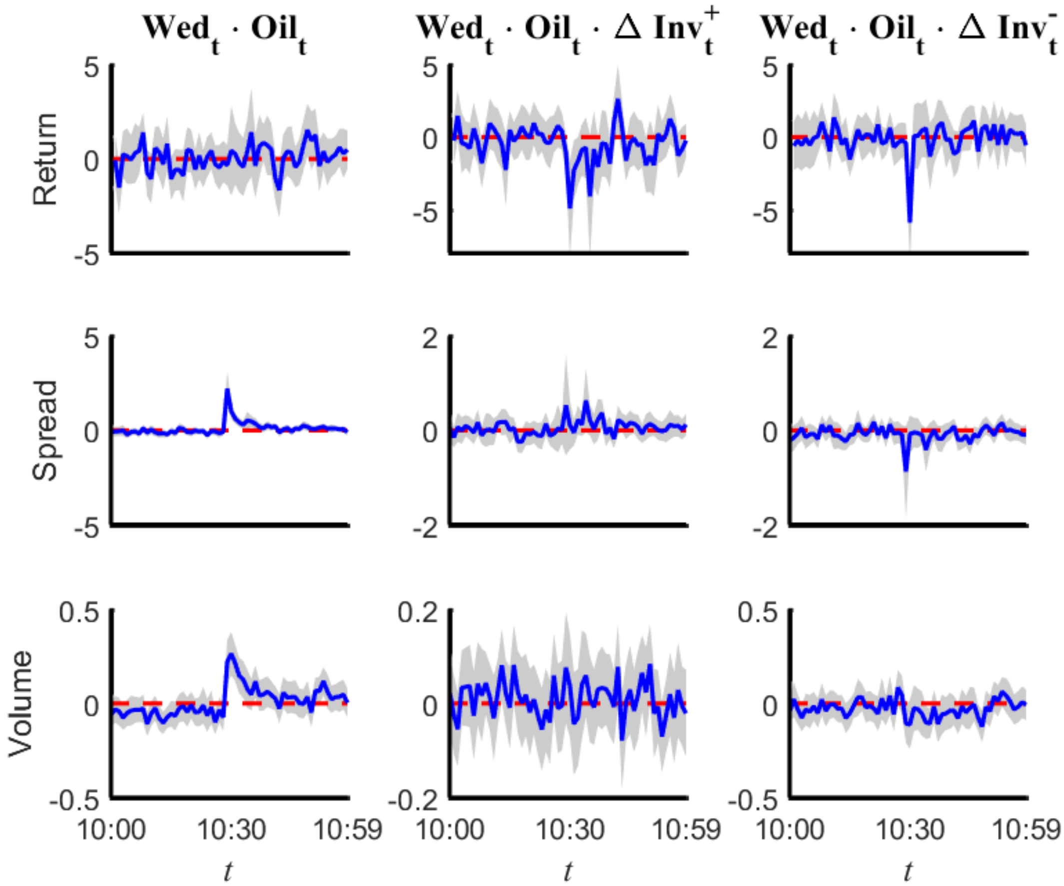

Figure 10: Asymmetric Reaction. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (5) Each row consider a different dependent variable, from top to bottom: midpoint return, proportional effective bid-ask spread, and number of transactions. I estimate the coefficients independently for every minute exploiting variation across days. The right-hand side plots correspond to γ_2^-, and the middle ones to γ_2^+; meanwhile, the left-hand side ones refers to γ_1. I describe the sample in Section 2. The red dashed line indicates the zero.

In Figure 10, I plot the estimates for every minute in the morning sample (from 10:00 a.m. to 10:59 a.m). The left and middle columns present the results for negative and positive information respectively. We observe that both cases are extremely similar.

Actually, the symmetry hypothesis ($\gamma_2^+ = \gamma_2^-$) cannot be rejected at the usual significance levels.[18] Likewise, the right column depicts the results for γ_1. Even if the violation of the symmetry assumption can affect the estimates of this parameter, the magnitude and sign of the coefficients are similar to the baseline case.

3.3 Non-linear reaction to news

An additional concern is the linearity assumption. Although smaller variations in inventories do not affect volume and spreads, extraordinary changes might. To test some possible non-linearity, I divide news content in three categories: *positive* if $\Delta Inv_t < -1.65$, *negative* if $\Delta Inv_t > 1.65$ or *zero* otherwise.[19]. Then, I estimate the following regression:

$$\begin{aligned} y_{i,t} = \mu + \delta_t + \theta_0 Oil_i - \theta_1^+ neg_t + \theta_1^- pos_t + \left(\theta_2 - \theta_3^+ neg_t + \theta_3^- pos_t\right) \cdot Wed_t \\ + \left(\gamma_0^- pos_t - \gamma_0^+ neg_t + \gamma_1 Wed_t + \left(\gamma_2^- pos_t - \gamma_2^+ neg_t\right) \cdot Wed_t\right) \cdot Oil_i + \varepsilon_{i,t} \quad (6) \end{aligned}$$

where neg_t (pos_t) takes value one if the observation belongs to the *negative* (*positive*) group, and *zero* is the excluded category. Note that I introduce a minus sign in front of the coefficients corresponding to *bad news*, and its interactions. Thus, the symmetry assumption remains $\gamma_2^+ = \gamma_2^-$.

Figure 11 shows that results remain unchanged under a binary identification. Further, I cannot reject that the effect is symmetric. This evidence reinforces the previous findings on the independence between the report content and the reaction of volume and spread.

4 Alternative Mechanisms

Although alternative mechanisms predict some of the previous empirical findings, they cannot explain all of them. For instance, Kim and Verrecchia (1991) proposes a model in which informed traders are the only ones able to interpret the public signal. Hence, when the signal is realized, the asymmetric information problem worsens and volume surges because informed traders trade more aggressively. Additionally, the strength of this effect

[18] I reject the hypothesis in three, five and four minutes (out of 60) at the 5% significance level in the case of returns, spreads, and volume, respectively. Furthermore, these minutes are not between 10:25 and 10:35 a.m.

[19] I select the threshold following Bernile et al. (2016); nonetheless, results do not vary using 1.75 or 2 as cutoffs

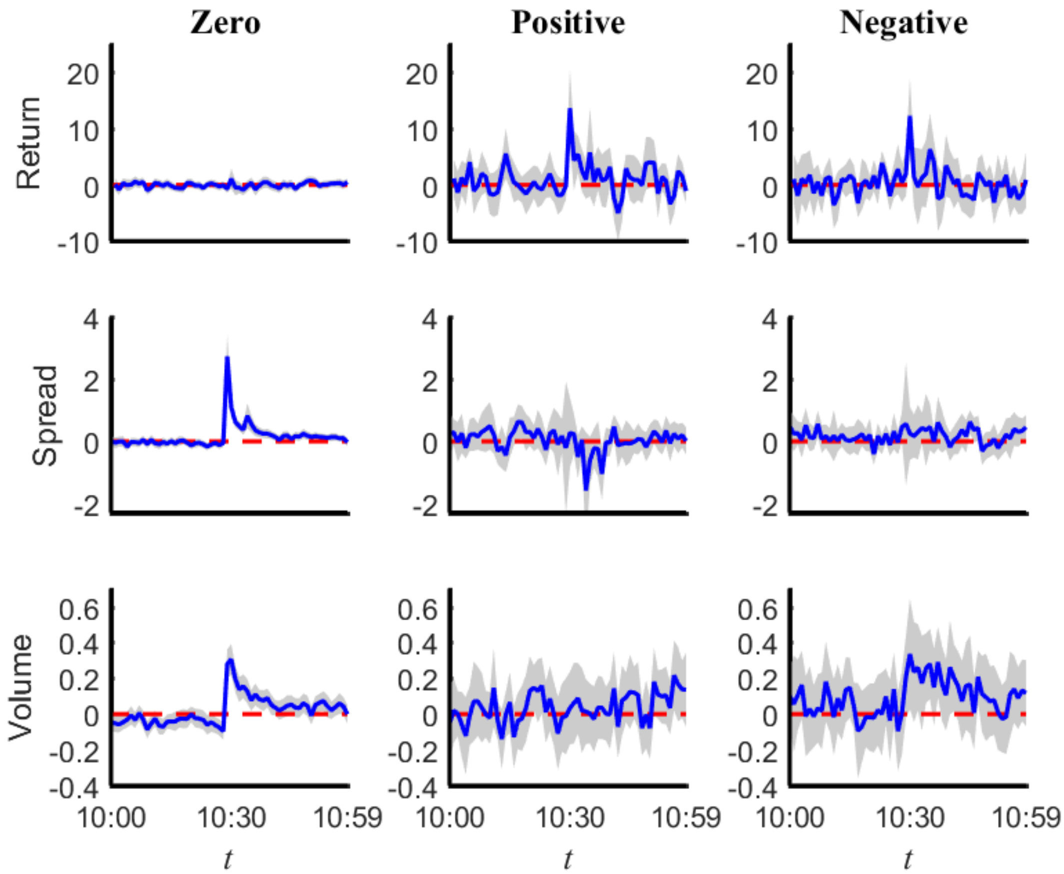

Figure 11: Binary Specification. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (6) Each row consider a different dependent variable, from top to bottom: midpoint return, proportional effective bid-ask spread, and number of transactions. I estimate the coefficients independently for every minute exploiting variation across days. The right-hand side plots correspond to γ_2^-, and the middle ones to γ_2^+; meanwhile, the left-hand side ones refers to γ_1. I describe the sample in Section 2. The red dashed line indicates the zero.

raises as the signal's realization moves away from its expectation since traders gain a larger informational advantage. Even if the dynamics of the variables is consistent with the data, the empirical results in Section 3.1 indicate that the change in inventories is irrelevant for the movement of volume and spreads.

While Kim and Verrecchia (1991) assumes that private and public signals are complements, they might be substitutes. In this regard, Pasquariello and Vega (2007) and Tetlock (2010) argue that liquidity improves and volume rises after the release of a private signal as the result of the reduction of the informed traders' advantage. Although the data support the increase in volume after the release of the WPS report, I present clear empirical evidence against the improvement of liquidity.

In a similar vein, Kandel and Pearson (1995) considers a market in which agents agree to disagree about the public signal. As a consequence, after the signal's realization, they are more willing to trade with each other which leads to higher volume and lower spreads. This negative relationship between volume and spreads is common to most models that include agree-to-disagree agents, but it is not present around public news' arrival. Nevertheless, (Bollerslev et al., forthcoming) advocates in favor of these models because they generate a volume-volatility elasticity below unity, in line with their empirical evidence but in contrast to most rational models. My model, while populated by rational agents, generates a volume-volatility elasticity below one with the precise number being a function of the shape of the risk aversion distribution and other parameters. Therefore, it can accommodate a low elasticity in addition to an increase in volume and spreads, reconciling the stylized facts and agents' rationality.

In sum, in the context of the announcement of the change in inventories, the data support the endogenous participation of risk-averse traders instead of alternative explanations. Nonetheless, it is very likely that every mechanism is at play and their importance differs according to the nature of the public announcement. For instance, understanding earnings announcements might require some knowledge about the firm, therefore Kim and Verrecchia (1991)'s channel becomes the most relevant. On the other hand, press releases disclose privately held information, hence we expect liquidity to improve after their publication consistent with the empirical findings in Tetlock (2010). Finally, there are important public announcements, as those related to macroeconomic or industry indicators, that are mostly unrelated to private information. In these situations, the endogenous participation of traders becomes first-order importance which explains why Pasquariello and Vega (2007) does not find evidence of a liquidity enhancement around macro-announcements, despite other predictions of their model are supported by the data.

5 Conclusions

This paper shows that the release of public information worsens the adverse selection problem if informed investors are risk-averse and participation is endogenous. Nonetheless, higher adverse selection costs come with more informative prices through two different channels. On the one hand, the market maker observes the public news and immediately adjust the quotes. On the other hand, the release of information resolves uncertainty and boosts the participation of informed agents which accelerates the market maker's learning about the private information.

The mechanism behind the model generates clear empirical implications that differ from the predictions of previous research; mainly, the signal's content is irrelevant to explain the movement of volume and bid-ask spreads around its release. I corroborate these hypotheses by estimating the causal effect of publishing the weekly change in oil inventories. Specifically, I find that spreads and volume increase regardless of the actual inventories' data whereas the midpoint reacts to this information. In line with the model, volume and spreads remain high for several minutes whereas the midpoint adjusts instantly.

The results of this paper have two important implication for policymakers whose objective is to enhance price informativeness. First, public information entails an amplification mechanism which must be taken into account when deciding the investment to compile information. Secondly, adverse selection costs rise, even if every investor obtains the same information from the public signal. Hence, the widening of bid-ask spreads at the time of a public announcement should not be taken as evidence to support a failure of standardization.

References

W. Antweiler and M. Z. Frank. Is all that talk just noise the information content of internet stock message boards. *The Journal of Finance*, 59(3):1259–1294, 2004.

R. Ball and P. Brown. An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, pages 159–178, 1968.

B. M. Barber and D. Loeffler. The dartboard column: Second-hand information and price pressure. *Journal of Financial and Quantitative Analysis*, 28(02):273–284, 1993.

W. H. Beaver. The information content of annual earnings announcements. *Journal of Accounting Research*, pages 67–92, 1968.

V. L. Bernard and J. K. Thomas. Post-earnings-announcement drift: delayed price response or risk premium? *Journal of Accounting Research*, pages 1–36, 1989.

V. L. Bernard and J. K. Thomas. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics*, 13(4):305–340, 1990.

G. Bernile, J. Hu, and Y. Tang. Can information be locked up? informed trading ahead of macro-news announcements. *Journal of Financial Economics*, 2016.

T. D. Berry and K. M. Howe. Public information arrival. *The Journal of Finance*, 49(4): 1331–1346, 1994.

T. Bollerslev, J. Li, and Y. Xue. Volume, volatility and public news announcements. Technical report, forthcoming.

S. Brown, S. A. Hillegeist, and K. Lo. The effect of earnings surprises on information asymmetry. *Journal of Accounting and Economics*, 47(3):208–225, 2009.

J. Chae. Trading volume, information asymmetry, and timing information. *The Journal of Finance*, 60(1):413–442, 2005.

W. S. Chan. Stock price reaction to news and no-news: drift and reversal after headlines. *Journal of Financial Economics*, 70(2):223–260, 2003.

M. Cipriani and A. Guarino. Estimating a structural model of herd behavior in financial markets. *The American Economic Review*, 104(1):224–251, 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, and M. O'Hara. One day in the life of a very common stock. *Review of Financial Studies*, 10(3):805–835, 1997.

L. H. Ederington and J. H. Lee. How markets process information: News releases and volatility. *The Journal of Finance*, 48(4):1161–1191, 1993.

J. Elder, H. Miao, and S. Ramchander. Impact of macroeconomic news on metal futures. *Journal of Banking & Finance*, 36(1):51–65, 2012.

M. J. Fleming and E. M. Remolona. Price formation and liquidity in the us treasury market: The response to public information. *The Journal of Finance*, 54(5):1901–1915, 1999.

G. Foster, C. Olsen, and T. Shevlin. Earnings releases, anomalies, and the behavior of security returns. *Accounting Review*, pages 574–603, 1984.

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. C. Green. Economic news and the impact of trading on bond prices. *The Journal of Finance*, 59(3):1201–1233, 2004.

T. Hendershott and A. J. Menkveld. Price pressures. *Journal of Financial Economics*, 114(3):405–423, 2014.

D. Hess, H. Huang, and A. Niessen. How do commodity futures respond to macroeconomic news? *Financial Markets and Portfolio Management*, 22(2):127–146, 2008.

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

C. W. Holden and S. Jacobsen. Liquidity measurement problems in fast, competitive markets: expensive and cheap solutions. *The Journal of Finance*, 69(4):1747–1785, 2014.

C. W. Holden and A. Subrahmanyam. Risk aversion, imperfect competition, and long-lived information. *Economics Letters*, 44(1-2):181–190, 1994.

E. Kandel and N. D. Pearson. Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy*, 103(4):831–872, 1995.

O. Kim and R. E. Verrecchia. Market reaction to anticipated announcements. *Journal of Financial Economics*, 30(2):273–309, 1991.

R. S. Koijen. The cross-section of managerial ability, incentives, and risk preferences. *The Journal of Finance*, 69(3):1051–1098, 2014.

I. Krinsky and J. Lee. Earnings announcements and the components of the bid-ask spread. *The Journal of Finance*, 51(4):1523–1535, 1996.

A. P. Kyle. Continuous Auctions and Insider Trading. *Econometrica*, 53(6):1315–1335, 1985.

C. M. Lee, B. Mucklow, and M. J. Ready. Spreads, depths, and the impact of earnings information: An intraday analysis. *Review of Financial Studies*, 6(2):345–374, 1993.

P. Liu, S. D. Smith, and A. A. Syed. Stock price reactions to the wall street journal's securities recommendations. *Journal of Financial and Quantitative Analysis*, 25(03): 399–410, 1990.

A. Madhavan and S. Smidt. A bayesian model of intraday specialist pricing. *Journal of Financial Economics*, 30(1):99–134, 1991.

M. O'Hara. *Market microstructure theory*, volume 108. Blackwell Cambridge, MA, 1995.

P. Pasquariello and C. Vega. Informed and strategic order flow in the bond markets. *Review of Financial Studies*, 20(6):1975–2019, 2007.

A. Ranaldo. Intraday market dynamics around public information arrivals. In *EFMA*

2002 London Meeting, Forthcoming, 2005.

R. Riordan, A. Storkenmaier, M. Wagener, and S. S. Zhang. Public information arrival: Price discovery and liquidity in electronic limit order markets. *Journal of Banking & Finance*, 37(4):1148–1159, 2013.

P. Savor and M. Wilson. Earnings announcements and systematic risk. *The Journal of Finance*, 71(1):83–138, 2016.

S. N. Soroka. Good news and bad news: Asymmetric responses to economic information. *Journal of Politics*, 68(2):372–385, 2006.

P. C. Tetlock. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168, 2007.

P. C. Tetlock. Does public financial news resolve asymmetric information? *Review of Financial Studies*, 23(9):3520–3557, 2010.

C. Vega. Stock price reaction to public and private information. *Journal of Financial Economics*, 82(1):103–133, 2006.

A Proofs

Proof. Proposition 1

Since the market maker problem is symmetric, I focus, first, on the ask side of the market. To shorten the proof I define σ as the residual variance for the informed investors, i.e. $\sigma = \sigma_\mu + \sigma_\varepsilon$ if $t < t_R$ and $\sigma = \sigma_\varepsilon$ otherwise; and, $\tilde{\mu}$ as the public information at time t, $\tilde{\mu} = 0$ before t_R and $\tilde{\mu} = \mu$ afterwards. In the case of no-news $\tilde{\mu} = 0$ and $\sigma = \sigma_\mu + \sigma_\varepsilon$ for all t.

To show that there is a unique equilibrium, I first show that the market maker's profits are increasing as the ask quote increases. Then I show that for very low quotes she loses money, whereas for very high quotes she earns profits. Hence, there is one, and only one, quote at which she breaks even.

The zero profit condition is given by:

$$A_t^* = \mathbb{E}\left[v|d=1,\mathcal{H}_t\right] = \tilde{\mu} + \mathbb{E}\left[\omega|d=1,\mathcal{H}_t\right]$$

where the last equality comes from independence between ω, μ and ε. Using the distribution of ω, we can write the market maker's profits as:

$$g_A(A) \equiv A + \sigma_\omega - P(\omega = \sigma_\omega|d=1,\mathcal{H}_t)2\sigma_\omega - \tilde{\mu} \tag{7}$$

where $P(\omega = \sigma_\omega|d=1,\mathcal{H}_t)$ is the probability of being in the high state of ω conditional of some agent buying and all the past information. Using Bayes' rule, this probability equals:

$$P(\omega = \sigma_\omega|d=1,\mathcal{H}_t) = P(d=1|\omega=\sigma_\omega,\mathcal{H}_t)\frac{P(\omega=\sigma_\omega|\mathcal{H}_t)}{P(d=1|\mathcal{H}_t)}$$

Moreover, note that $\mathcal{H}_t$ is irrelevant if we condition on ω, therefore the first term equals

$$P(d=1|\omega=\sigma_\omega,\mathcal{H}_t) = P(d=1|\omega=\sigma_\omega) = F\left(\frac{\sigma_\omega + \tilde{\mu} - A_t^*}{\sigma}\right)\delta + \frac{1}{2}\left(1-\delta\right)$$

which corresponds to the probability that an informed buys given that the asset value is high times the probability that the trader is actually informed plus the complementary probability times the probability that a noise trader buys, one half.

Likewise, the denominator equals the probability that an informed buys, weighted by the probability that the asset value is high, times the likelihood for him to be informed plus the probability that a noise trader buys:

$$P(d=1|\mathcal{H}_t) = P(\omega=\sigma_\omega|\mathcal{H}_t)F\left(\frac{\sigma_\omega+\tilde{\mu}-A_t^*}{\sigma}\right)\delta+\frac{1}{2}\left(1-\delta\right)$$

which leads to the following expression for $P(\omega=\sigma_\omega|d=1,\mathcal{H}_t)$:

$$P(\omega=\sigma_\omega|d=1,\mathcal{H}_t) = \frac{P(\omega=\sigma_\omega|\mathcal{H}_t)\left[F\left(\frac{\sigma_\omega+\tilde{\mu}-A_t^*}{\sigma}\right)\delta+\frac{1}{2}\left(1-\delta\right)\right]}{P(\omega=\sigma_\omega|\mathcal{H}_t)F\left(\frac{\sigma_\omega+\tilde{\mu}-A_t^*}{\sigma}\right)\delta+\frac{1}{2}\left(1-\delta\right)}.$$

Summing and subtracting $(1-P(\omega=\sigma_\omega|\mathcal{H}_t))\frac{1}{2}(1-\delta)$ from the numerator we obtain:

$$P(\omega=\sigma_\omega|d=1,\mathcal{H}_t) = 1-\frac{1-P(\omega=\sigma_\omega|\mathcal{H}_t)}{1+2P(\omega=\sigma_\omega|\mathcal{H}_t)\frac{\delta}{1-\delta}F\left(\frac{\sigma_\omega+\tilde{\mu}-A_t^*}{\sigma}\right)} \tag{8}$$

from which is obvious that $\frac{dP(\omega=\sigma_\omega|d=1,\mathcal{H}_t)}{dA_t^*}<0$. This result directly implies that $g_A'(A)>0$.

Using the definition of $g_A(\cdot)$ in (7), we can check that $g_A(\tilde{\mu}-\sigma_\omega)\leq 0$ and $g_A(\tilde{\mu}+\sigma_\omega)\geq 0$. Therefore, an equilibrium exists and it is unique. Moreover,

$$\tilde{\mu}-\sigma_\omega\leq A_t^*\leq\tilde{\mu}+\sigma_\omega \tag{9}$$

In the case of the bid price, the zero profit condition is:

$$B_t^* = \mathbb{E}\left[v|d=-1,\mathcal{H}_t\right] = \tilde{\mu}+\mathbb{E}\left[\omega|d=-1,\mathcal{H}_t\right]$$

and we can define the profits from buying as:

$$g_B(B)\equiv B+\sigma_\omega-P(\omega=\sigma_\omega|d=-1,\mathcal{H}_t)2\sigma_\omega-\tilde{\mu} \tag{10}$$

where by Bayes rule,

$$P(\omega=\sigma_\omega|d=-1,\mathcal{H}_t) = \frac{\frac{1}{2}(1-\delta)P(\omega=\sigma_\omega|\mathcal{H}_t)}{(1-P(\omega=\sigma_\omega|\mathcal{H}_t))F\left(\frac{B_t^*+\sigma_\omega-\tilde{\mu}}{\sigma}\right)\delta+\frac{1}{2}(1-\delta)} \tag{11}$$

From this expression is straightforward to check that $g_B'(B)>0\;\forall\;B$. Finally, substituting B in equation (10), we obtain that $g_B(\tilde{\mu}-\sigma_\omega)\leq 0$ and $g_B(\tilde{\mu}+\sigma_\omega)\geq 0$.

□

Lemma 2. *In equilibrium, under the assumption that* $0 < \delta < 1$ *and* $F(c) > 0\ \forall c > 0$,

$$0 < P(\omega = \sigma_\omega | d = 1, \mathcal{H}_t) < 1$$

Proof. According to equation (8), these inequalities hold if $0 < \delta < 1$, $F(c) > 0\, \forall c$ and $0 < P(\omega = \sigma_\omega | \mathcal{H}_t) < 1$. While the first two hold by assumption, the third one depends on an endogenous quantity that we can obtain in a recursive fashion. Precisely, there are three cases:

- $d_t = 1 \Rightarrow P(\omega = \sigma_\omega | \mathcal{H}_{t+1}) = P(\omega = \sigma_\omega | d = 1, \mathcal{H}_t)$
- $d_t = -1 \Rightarrow P(\omega = \sigma_\omega | \mathcal{H}_{t+1}) = P(\omega = \sigma_\omega | d = -1, \mathcal{H}_t)$
- $d_t = 0 \Rightarrow P(\omega = \sigma_\omega | \mathcal{H}_{t+1}) = P(\omega = \sigma_\omega | d = 0, \mathcal{H}_t)$

The first two cases are obtained in equation (8) and (11). The third one can be obtained as:

$$P(\omega = \sigma_\omega | d = 0, \mathcal{H}_t) = \cfrac{1}{1 + \cfrac{1 - P(\omega = \sigma_\omega | \mathcal{H}_t)}{P(\omega = \sigma_\omega | \mathcal{H}_t)} \cfrac{1 - F\left(\cfrac{\sigma_\omega + \tilde{\mu} - A_t^*}{\sigma}\right)}{1 - F\left(\cfrac{B_t^* + \sigma_\omega - \tilde{\mu}}{\sigma}\right)}}$$

In every case, if $0 < P(\omega = \sigma_\omega | d = 1, \mathcal{H}_t) < 1$ then $0 < P(\omega = \sigma_\omega | d = 1, \mathcal{H}_{t+1}) < 1$. Since $P(\omega = \sigma_\omega | \mathcal{H}_0) = \frac{1}{2}$, we proved that $0 < P(\omega = \sigma_\omega | \mathcal{H}_t) < 1\ \forall t$. □

Proof. Corollary 1

Using Lemma 2, is straightforward to show that inequalities in Equation (9) are strictly satisfied. We can follow a similar argument, and show that these inequalities hold for the bid price.

Besides, to show that $B_t^* < A_t^*$, we subtract $g_A(A_t^*)$ and $g_B(B_t^*)$ leading to the following expression:

$$A_t^* - B_t^* = 2\sigma_\omega \left[P(\omega = \sigma_\omega | d = 1, \mathcal{H}_t) - P(\omega = \sigma_\omega | d = -1, \mathcal{H}_t)\right]$$

Therefore the quoted spread is positive if and only if the right-hand side is positive for every A_t^* and B_t^*; equivalently, using Bayes rule, if and only if

$$\frac{P(d = 1 | \omega = \sigma_\omega, \mathcal{H}_t)}{P(d = 1 | \mathcal{H}_t)} > \frac{P(d = -1 | \omega = \sigma_\omega, \mathcal{H}_t)}{P(d = -1 | \mathcal{H}_t)}$$

Inverting both sides, and applying the law of total probability we get:

$$\frac{P(d=-1|\omega=-\sigma_\omega,\mathcal{H}_t)}{P(d=-1|\omega=\sigma_\omega,\mathcal{H}_t)} > \frac{P(d=1|\omega=-\sigma_\omega,\mathcal{H}_t)}{P(d=1|\omega=\sigma_\omega,\mathcal{H}_t)}$$

Finally, if we plug the corresponding values we obtain

$$B_t^* < A_t^* \iff \frac{\left[F\left(\frac{B_t^*+\sigma_\omega-\tilde{\mu}}{\sigma^2}\right)\delta+\frac{1}{2}(1-\delta)\right]}{\left(\frac{1}{2}(1-\delta)\right)} > \frac{\left(\frac{1}{2}(1-\delta)\right)}{\left[F\left(\frac{\sigma_\omega+\tilde{\mu}-A_t^*}{\sigma^2}\right)\delta+\frac{1}{2}(1-\delta)\right]}.$$

This inequality is satisfied due to the assumptions that $\delta > 0$ and $F(c) > 0$ if $c > 0$. □

Proof. Proposition 2

Given $\mathcal{H}_t = \{d_1, ..., d_{t-1}\}$, let me define $-\mathcal{H}_t = \{-d_1, ..., -d_{t-1}\}$. Next, after some algebra we can show that $(A^*(\mathcal{H}_t) - \tilde{\mu}) = -(B^*(-\mathcal{H}_t) - \tilde{\mu})$. This result implies that:

$$m_t(\mathcal{H}_t) = \frac{A^*(\mathcal{H}_t) + B^*(\mathcal{H}_t)}{2} = -\frac{A^*(-\mathcal{H}_t) + B^*(-\mathcal{H}_t)}{2} = -m_t(-\mathcal{H}_t)$$

Moreover, note that given the symmetry in the model, $P(\mathcal{H}_t|\omega=\sigma_\omega) = P(-\mathcal{H}_t|\omega=-\sigma_\omega)$. Therefore, $\mathbb{E}_{\mathcal{H}_t}(m_{1t}) = \tilde{\mu}$ if there are news; and $\mathbb{E}_{\mathcal{H}_t}(m_{0t}) = 0$, otherwise.

□

Lemma 3.

$$lim_{T\to\infty}\, A_T = lim_{T\to\infty}\, B_T = \omega + \tilde{\mu}$$

Proof. Let define the percentage of buys as N_B and the percentage of sales as N_S. Corollary 1 ensures that if $\omega = \sigma_\omega$ ($\omega = -\sigma_\omega$), informed investors do not buy (sell). Moreover, it also establish that there is always a probability that an informed investor trades, given the assumption that $F(c) > 0$ if $c > 0$. Then, using the law of large numbers (LLN), it is possible to show that

$$lim_{T\to\infty}\, P(N_B > N_S|\omega=\sigma_\omega) = 1 \text{ and } lim_{T\to\infty}\, P(N_B > N_S|\omega=-\sigma_\omega) = 0. \quad (12)$$

Note that the random arrival of traders and $0 < \delta < 1$ are sufficient conditions to apply the LLN.

Lastly, the results in Equation (12) imply that

$$lim_{T\to\infty}\, P(\omega=\sigma_\omega|\mathcal{H}_t) = 1 \text{ if } \omega=\sigma_\omega, \text{ and}$$
$$lim_{T\to\infty}\, P(\omega=-\sigma_\omega|\mathcal{H}_t) = 1 \text{ if } \omega=-\sigma_\omega.$$

□

Proof. Proposition 4

As noise traders trade equally by assumption, it is enough to show that informed traders trade more. Specifically, I show that:

$$F\left(\frac{\sigma_\omega + \mu - A^*_{1t}}{\sigma_\varepsilon}\right) > F\left(\frac{\sigma_\omega - A^*_{0t}}{\sigma_\varepsilon + \sigma_\mu}\right) \tag{13}$$

which is the probability that an investor trades given she is informed and $\omega = \sigma_\omega$. Note that, focusing in the ask side of the book, both quantities would be equal if:

$$A_{1t_R} = \mu + \sigma_\omega + (A^*_{0t_R} - \sigma_\omega)\frac{\sigma_\varepsilon}{\sigma_\varepsilon + \sigma_\mu} \equiv \bar{A}.$$

Since $g_A(A^*_{0,t}) = 0$ by definition; furthermore $P(\omega = \sigma_\omega | d = 1, \mathcal{H}_t)$ is the same in both scenarios as the probability of informed trading is the same,

$$g_A(\overline{A}) = \overline{A} - A^*_{0,t_R} \Rightarrow g(\overline{A}) = \left(1 - \frac{\sigma_\varepsilon}{\sigma_\varepsilon + \sigma_\mu}\right)\left(\sigma_\omega - A^*_{0t_R}\right) > 0.$$

As $g_A(\cdot)$ is increasing and $g_A(A^*_{1t_R}) = 0$ then, $\bar{A} > A^*_{1t_R}$. Therefore, (13) is satisfied.

The case of the bid side is very similar. The price which makes volume equal is:

$$\overline{B} = \mu - \sigma_\omega + (B^*_{0,t_R} + \sigma_\omega)\frac{\sigma_\varepsilon}{\sigma_\varepsilon + \sigma_\mu}.$$

Using the same strategy as before we get that

$$g_B(\overline{B}) = (B^*_{0,t_R} + \sigma_\omega)\left(\frac{\sigma_\varepsilon}{\sigma_\varepsilon + \sigma_\mu} - 1\right) < 0$$

Since $g_B(B^*_{1,t_R}) = 0$ and $g'_B(\cdot) > 0$ then $B^*_{1,t_R} > \overline{B}$. Therefore,

$$F\left(\frac{B^*_{1t} - \mu + \sigma_\omega}{\sigma_\varepsilon}\right) > F\left(\frac{B^*_{0t} + \sigma_\omega}{\sigma_\varepsilon + \sigma_\mu}\right)$$

To prove that the increase does not depend on μ is enough to show that volume is not affected by this variable. From Equation (7), we know that $A^*_{1t} - \tilde{\mu}$ does not depend on μ. As a consequence, volume, defined as the right hand side of inequality (13), does not depend on μ.

The limiting result is trivial using Lemma 3 □

Proof. Proposition 3

To show the effect on spreads, the following results are useful. $P_1(\omega = \sigma_\omega|\mathcal{H}_{t_R})$ equals $P_0(\omega = \sigma_\omega|\mathcal{H}_{t_R})$ by construction, as models are identical until t_R, and $P_0(d = -1|\omega = \sigma_\omega, \mathcal{H}_{t_R}) = \frac{1}{2}(1-\delta) = P_1(d = -1|\omega = \sigma_\omega, \mathcal{H}_{t_R})$, that is, regardless if there is public news or not, only uniformed sell when the value of the private component of the asset is high. From the zero-profit condition, we know that:

$$A^*_{1t_R} - A^*_{0t_R} = \mu + 2\sigma_\omega \left[P_1(\omega = \sigma_\omega|d_{t_R} = 1, \mathcal{H}_{t_R}) - P_0(\omega = \sigma_\omega|d_{t_R} = 1, \mathcal{H}_{t_R})\right]$$

which we can rewrite using Bayes' rule as:

$$A^*_{1t_R} - A^*_{0t_R} = \mu + \left[\frac{P_1\left(d_{t_R} = 1|\omega = \sigma_\omega \mathcal{H}_{t_R}\right)}{P_1\left(d_{t_R} = 1|\mathcal{H}_{t_R}\right)} + \frac{P_0\left(d_{t_R} = 1|\omega = \sigma_\omega \mathcal{H}_{t_R}\right)}{P_0\left(d_{t_R} = 1|\mathcal{H}_{t_R}\right)}\right] P(\omega = \sigma_\omega|\mathcal{H}_{t_R})2\sigma_\omega$$

Using the law of total probability and the results stated at the beginning, we get

$$A^*_{1t_R} - A^*_{0t_R} = \mu + \frac{P(\omega = \sigma_\omega|\mathcal{H}_{t_R})^2\,(1-\delta)\,\delta\left(F\left(\frac{\sigma_\omega + \mu - A^*_{1t}}{\sigma_\varepsilon}\right) - F\left(\frac{\sigma_\omega - A^*_{0t}}{\sigma_\varepsilon + \sigma_\mu}\right)\right)}{P_1\left(d_{t_R} = 1|\mathcal{H}_{t_R}\right)P_0\left(d_{t_R} = 1|\mathcal{H}_{t_R}\right)}\sigma_\omega$$

therefore, given (13) , $A^*_{1t_R} - \mu > A^*_{0t_R}$. The same proof can be easily extended to bid prices to obtain that $B^*_{1t_R} - \mu < B^*_{0t_R}$. Thus, $A^*_{1t_R} - B^*_{1t_R} > A^*_{0t_R} - B^*_{0t_R}$.

Note that the independence of μ and the limiting case follow straightforward from the proof of Proposition 4, and Lemma 3.

□

A.1 Proofs extensions

Proof. Proposition 5

Let define $h(A) = A + \sigma_\omega - 2\sigma_\omega P(\omega = \sigma_\omega|\mathcal{H}_t, d = 1) + \tilde{\mu}$; therefore, A^* is the only ask quote such that $h(A^*) = 0$. Following the proof of Proposition 1, we need to prove that $h'(A) > 0$ for all A; or equivalently $\frac{dP(\omega = \sigma_\omega|\mathcal{H}_t, d = 1)}{dA} > \frac{1}{2\sigma_\omega}$. With some algebra we get that $P(\omega = \sigma_\omega|\mathcal{H}_t, d = 1) = 1 - \frac{1}{X + \frac{1}{(1-\Pi))}}$ where $\Pi = P(\omega = \sigma_\omega|\mathcal{H}_t)$ and

$$X = \frac{2\delta\Pi F\left(\frac{\sigma_\omega + \tilde{\mu} - A}{\sigma}\right)}{(1-\Pi)(1-\delta)\left(1 - \Phi\left(\frac{A - \mathbb{E}\left(v|\mathcal{H}_t\right)}{\theta}\right)\right)}$$

A sufficient condition for an equilibrium to exist and be unique is that $\frac{dX}{dA} > 0$ for all A. This derivative is given by:

$$\frac{dX}{dA} = \frac{-2\Pi\delta(1-\Pi)}{(1-\Pi)(1-\delta)\left(1-\Phi\left(\frac{A-\mathbb{E}\left(v|\mathcal{H}_t\right)}{\theta}\right)\right)}\Delta$$

Since the first factor is positive, we can focus on Δ:

$$\Delta = \left(1-\Phi\left(\frac{A-\mathbb{E}\left(v|\mathcal{H}_t\right)}{\theta}\right)\right) f\left(\frac{\sigma_\omega+\tilde{\mu}-A}{\sigma}\right)\frac{1}{\sigma} - \phi\left(\frac{A-\mathbb{E}\left(v|\mathcal{H}_t\right)}{\theta}\right)\frac{1}{\theta}F\left(\frac{\sigma_\omega+\tilde{\mu}-A}{\sigma}\right)$$

Δ needs to be positive for every A which leads to the condition in Proposition 5.

□

Proof. Lemma 1

I focus on the ask side. $A_t = \mathbb{E}\left(v|\mathcal{H}_t, d=1\right)+\kappa$ can never be an equilibrium for $\kappa > 0$ since risk-neutral market makers would under-price.[20] Therefore, the only possible equilibrium is $A_t = \mathbb{E}\left(v|\mathcal{H}_t, d=1\right)$ which is indeed an equilibrium as it maximizes the utility of the market maker. Actually, she is indifferent between leaving the market or providing liquidity.

The same rationale follows for the bid side. Note that, while in principle both sides must be considered together, the reaction functions are independent. For instance, $A_t > \mathbb{E}\left(v|\mathcal{H}_t, d=1\right)$ and $B_t > \mathbb{E}\left(v|\mathcal{H}_t, d=-1\right)$ is not an equilibrium because traders will sell to this market maker but they will buy to any market maker that posts $A_t = \mathbb{E}\left(v|\mathcal{H}_t, d=1\right)$. As a consequence, the utility of the first market maker will be negative.

□

Proof. Proposition 6

Results are a weighted average between the baseline model, and the difference between a model without private news and one with private news but without public news.

In the case of the midpoint the proof of Proposition 2 applies. Bid-ask spreads, however, are different as the market maker learns the private information with probability ρ. In that case, after t_R:

[20] $G(c) > 0\ \forall c > 0$ ensures these agents exists.

$$A^*_{1,t} - B^*_{1,t} - (A^*_{0,t} - B^*_{0,t}) = -(A^*_{0,t} - B^*_{0,t})$$

hence

$$\Delta BidAsk^{CS} = (1-\rho)\Delta BidAsk^{B} - \rho\left(A^*_{0,t_R} - B^*_{0,t_R}\right).$$

Similarly, if ω is revealed at t_R then:

$$\mathbb{E}\left(|d^*_{1,t}|\right) - \mathbb{E}\left(|d^*_{0,t}|\right) = (1-\delta) - \left((1-\delta) + \delta\frac{1}{2}\left(F\left(\frac{\sigma_\omega - A^*_{0,t_R}}{\sigma_\epsilon + \sigma_\mu}\right) + F\left(\frac{B^*_{0,t_R} + \sigma_\omega}{\sigma_\epsilon + \sigma_\mu}\right)\right)\right)$$

for all $t \geq t_R$ which leads to:

$$\Delta Volume^{CS} = (1-\rho)\Delta Volume^{B} - \rho\frac{\delta}{2}\left(F\left(\frac{\sigma_\omega - A^*_{0,t_R}}{\sigma_\epsilon + \sigma_\mu}\right) + F\left(\frac{B^*_{0,t_R} + \sigma_\omega}{\sigma_\epsilon + \sigma_\mu}\right)\right).$$

□