working paper
1714

Does Public News Decrease Information Asymmetries? Evidence from the Weekly Petroleum Status Report

Julio A. Crego

November 2017

CENTRO DE ESTUDIOS
MONETARIOS Y FINANCIEROS
Casado del Alisal 5, 28014 Madrid, Spain
www.cemfi.es

CEMFI Working Paper No. 1714
November 2017

Does Public News Decrease Information Asymmetries? Evidence from the Weekly Petroleum Status Report

Abstract

I argue that the arrival of a public signal, regardless of its content, can yield an increase in adverse selection costs in financial markets. To explain its occurrence, I propose a dynamic model with a public signal and risk-averse informed investors. In this set-up, the public signal induces informed investors to participate in the market as it reduces uncertainty. While it increases adverse selection costs, the increase in participation results in more informative prices. Apart from the static effects, the model's dynamics deliver testable hypotheses about price and liquidity before and after the signal's release. Using transaction-level data, I estimate the effect of the release of the Weekly Petroleum Status Report on the bid-ask spread, volume, and midpoint returns via a difference-in-difference strategy. I find that the mean bid-ask spread doubles immediately after the release and that volume increases by 32 percent. Moreover, this effect persists over time, and is independent of the report's content whereas prices react to this information immediately. Nevertheless, liquidity at the end of the trading session is not affected by the report.

JEL Codes: G12, G14.

Keywords: Public information, news release, asymmetric information, liquidity.

Julio A. Crego
Tilburg University
J.A.Crego@uvt.nl

Acknowledgement

I am indebted to my advisor Dante Amengual for his encouragment and useful comments. For helpful suggestions, I also thank Manuel Arellano, Guillermo Caruana, Julio Gálvez, Monica Martinez-Bravo, José Penalva, Rafael Repullo, and Enrique Sentana. Errors and omissions are exclusively my own responsibility.

Large market fluctuations within the trading day are usually a result of the arrival of public information which, even if occasionally might be unexpected (e.g., catastrophes and firm scandals) are typically scheduled in advance[1]. While the academic literature has established that prices react to this type of information, the effect that they have on the degree of information asymmetry or, more generally on market liquidity, is still unclear. Hence, the purpose of this paper is to empirically assess the consequences of the release of public information on liquidity, and to shed light on the mechanism at work.

A deeper understanding of the liquidity and price reactions is important to policy makers, who manage most of these releases. As highlighted by the Turner Review in the United Kingdom, numerous informational deficiencies in financial markets, both in terms of product disclosure and information reporting, may have caused various financial markets to "deviate from their rational price equilibria", ultimately leading to the financial crisis. Although this paper does not provide a welfare analysis, the policy implication that underlies much of the empirics is that of ensuring market efficiency as a crucial ingredient in the decision of releasing public information. The optimal policy, however, would depend if one needs to increase information rents to provide incentives for information acquisition, or to decrease transaction costs for uninformed investors, thereby preventing a market collapse.

Previous empirical evidence on the effect of these releases on information asymmetries provides mixed results. While bid-ask spreads and price impact rise after earnings announcements, and macroeconomic news, (e.g., Fleming and Remolona, 1999) other papers have found a decrease in these measures following news press releases (e.g., Tetlock, 2010). Similarly, theoretical models provide ambiguous results on the impact of public signals on information asymmetries, which depend on the relationship between the public signal and the private information owned by some traders. On one hand, if the public signal discloses private information, bid-ask spreads reduce as informed traders lose the informational advantage (Tetlock, 2010). On the other hand, public news might increase the bid-ask spread if public data releases do not provide the same information to every trader (Kim and Verrecchia, 1991). For instance, a group of traders might possess the

[1]As an example, the Consumer Comfort Index report is released by Bloomberg every Thursday.

skills to understand the implications of the data in terms of the fundamental value, while the remaining investors might dismiss the information as noise. This mechanism implies that extreme realizations of the public signal result in wider spreads than those close to the signal's expected value.

In this paper, I show that this prediction is inconsistent with the data in the context of the Weekly Petroleum Status (WPS) Report, a document released by the US Energy Information Administration (EIA). As reflected by the data, bid-ask spreads increase regardless of the reported change in oil inventories. I then extend Glosten and Milgrom (1985) by introducing a public signal and risk-averse informed investors, and show that the model provides predictions that are consistent with the empirical evidence. In the model, some of the risk-averse investors do not make use of their private information before release of the public signal because of the risk they face. Once it is revealed, however, risk goes down and they enter the market, increasing adverse selection costs. Consequently, the market maker widens the bid-ask spread to maintain zero profits. The mechanism hinges on the reduction in uncertainty caused by the public signal. Therefore, the volume and spreads reaction does not depend on the precise value of the signal realization. Nevertheless, the realization is important as the market maker moves prices to reflect this information.

To understand price and quote movements after the introduction of a public announcement, I compare my model with and without public news. Before the release, both models are identical since agents are not forward-looking. At the moment of the disclosure, however, spreads and volume are higher in the former model than in the latter if a public release takes place. Meanwhile, the midpoint is higher in the model with public news if the signal is good, and lower otherwise. With respect to their dynamics, the theoretical model predicts that the difference on spreads, and volume, between the two scenarios is persistent. Further, it only dissipates when the market maker learns the private information from the trades. In contrast, returns are different just at the moment of the release, consistent with the semi-strong efficient markets hypothesis (Fama, 1970).

I test the implications of the theoretical model within the context of the WPS report, which provides an empirically appropriate setting for a credible empirical exercise. This document, which is released by the 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 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 the 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. 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 this point in time, 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. This effect in returns is immediate since after 10:30 a.m. estimates become insignificant. In contrast, bid-ask spreads widen by 2.79 bps, independently of the change in inventories, and this effect is persistent. I find a similar pattern in the case of volume with an initial boost of 32% and a posterior decline. At the end of the day, however, effects become insignificant, which are in agreement with the limiting results of the model.

With respect to my empirical approach, my contribution is threefold. First, I use ultra-high frequency data. While some existing papers consider intraday data by aggregating at thirty minutes intervals, I use data at the one-minute frequency. This distinction is important as my results suggest that the announcement effect vanishes before the end of the trading session. Second, I estimate the effect of the presence of this information not only at the release time, but also before and after the information becomes public. Third, my empirical strategy relies on a difference-in-difference approach which identifies the effect, conditional on the information content. In contrast to previous literature, this method isolates the effect of the announcement from overall effects in the market, and do not require to use days with "no-surprise news" as a control. I find that spreads and volume react independently of this content; therefore, using "no-surprise news" days as a control would suggest that public news does not affect these variables. Further, the relationship between the market reaction to a public signal and its content delivers new insights on the mechanisms that can be at work.

I also contribute to the theoretical literature on how prices incorporate public information. To the best of my knowledge, my model is the first one to include risk-averse investors and a public signal in a framework with non-strategic traders. This departure from previous literature creates endogenous changes in investor composition around the public signal. This mechanism is necessary, as existing models with constant investor composition cannot explain an increment on information asymmetry which does not depend on the signal's content.

In the following two sections I review the most related literature; and I describe the institutional framework, and the data. Next, in Section 3, I present the raw data descriptively. Then, in Section 4, I introduce the model whose predictions I test in Section 5. Finally, Section 6 explores additional results and robustness checks and Section 7 concludes. Proofs of the theoretical results are gathered in Appendix A.

1 Related Literature

This paper contributes to a vast empirical and theoretical literature on the effect of public news on the market. Previous research, as I do, focuses on one specific type of news items: periodic, pre-scheduled, or unscheduled. The first group consists on public information that is released at a given day repetitively, e.g. macroeconomic news, or data on stocks of crude. Meanwhile, the second group includes publications that take place periodically but the information issuer chooses the exact date and time, usually being aware of the content. The main example, which is the focus of previous research, are earnings announcements.

The last group contains news surprises usually disclosed by media press.

Regarding periodic releases, most of the preceding literature concentrates on the effect of macroeconomic data publication on the bond and currency market. In particular, Ederington and Lee (1993), using futures data on T-bonds and exchange rates, find that prices reaction to public news is short-lived while the effect on variance is persistent. Whereas they focus on one specific maturity, Fleming and Remolona (1999) explore different maturities; they conclude that macroeconomic news affects specially medium maturity contracts. Beyond prices and volatility, Green (2004) analyzes price impact around the date at which information becomes public. He observes that the release increases price impact; furthermore, this effect is positively related to the relevance of the public information. In contrast, Pasquariello and Vega (2007) do not find any significant changes on the price impact after the news announcement, although their model predicts it should decrease. Additionally, they find that order flow is more informative when the professional forecasts dispersion is high and the public signal is noisy.

A similar strand of the literature studies the effect of these announcements on the stock market. Using a database that differentiates institutions and individual investors, Nofsinger (2001) finds that both types take long positions on large firms after good economic news, but they short after bad news. Complementary, Boyd et al. (2005) analyze the Bureau of Labor Statistics' monthly announcement of the unemployment rate. They establish that there are two opposite effects present: on the one hand, the report affects the interest rate; but on the other, it provides information about dividend growth. The net effect depends on the economic cycle. Prices react positively during expansions whereas they respond negatively during recessions.

In a similar vein, Hess et al. (2008) show that prices of commodities react differently according to the economic cycle. Precisely, during recessions, positive news lead to an increase in commodity prices while this effect is not present during expansions. Meanwhile, Roache and Rossi (2010) argue that commodities are barely affected by these news items compared with bonds or stocks. In contrast, Elder et al. (2012) using intraday data find that macroeconomic announcements influence positively volume and variance, apart from their effect on prices.

This paper is the first one to analyze the Weekly Petroleum Status Report, which presents two important features different from macroeconomic news. The first one relates to its content. While macroeconomic news can affect prices through several channels such as changing the discount rate or risk aversion, news on oil inventories are less likely to have an impact beyond the drop in uncertainty about the underlying demand of supply of oil. Further, even if the information channel is not the only mechanism, its effect can be identified comparing oil firms with the remaining ones as long as other factors affect equally both sectors. The second characteristic that demarcates this report from other macroeconomic news is the publication channel. On the one side, macroeconomic news are usually released privately to media outlets before the scheduled date under embargo agreements. Even if contracts and lock-up rooms should prevent information to be leaked, Bernile et al. (2016) and Kurov et al. (2016) show that prices react ahead of the official communication. On the other side, the US Energy Information Administration publishes the report electronically on its web page; thus, it is not subject to any leaking.

In relation to pre-scheduled news as earnings announcements, the literature starts with the seminal papers by Beaver (1968) and Ball and Brown (1968) who show that prices present a drift after these announcements. In addition, Foster et al. (1984), Bernard and Thomas (1989, 1990) and Savor and Wilson (2016) corroborates these findings and analyze their relationship with different types of risks. Related to information asymmetry, Vega (2006) determines that stocks associated with high probability of informed trading, consensus public news, and low media coverage experience low or insignificant drift after the release.

Another strand of the literature studies the trading behavior of institutions and individuals around earnings announcements (see Ke and Ramalingegowda, 2005; Vieru et al., 2006; Hirshleifer et al., 2008; Campbell et al., 2009; Kaniel et al., 2012) and determines that the order flow of both types of traders, prior to the report, predicts earnings. Particularly relevant to this paper is Kaniel et al. (2012) who decompose the order flow between liquidity provision and informed trading and establish that both components are equally important.

While these studies are silent on the bid-ask spread dynamics around earnings announcements, Lee et al. (1993) and Krinsky and Lee (1996) show that the spread widens after an earnings announcement; however, this effect is short-lived. Considering positive and negative news differently, Brown et al. (2009) show that the former decreases information asymmetry with respect to zero-surprise news, whereas the latter increases it.

Most of the aforementioned papers present empirical evidence that some investors know the content of earnings reports in advance. These findings indicate that these announcements are more related to the disclosure of previously privately held information than to the publication of new information to all investors. While both types of public signals are interesting, I focus on the latter which distinguish this paper from the literature I have just discussed. Moreover, in contrast to oil inventories, earnings reports are subject to some discretionary decision by the firm such as the publication date, the way to present the content, etc. Actually, Kross and Schroeder (1984), Begley and Fischer (1998) and Michaely et al. (2016) present evidence that firms strategically choose the timing to release earnings news. Yet, previous literature has not dealt with this possible endogeneity issue.

With regard to unscheduled news, prior research concentrates on media outlets and electronic news feeds. Liu et al. (1990) analyze the *Heard on the street* column from the Wall Street Journal and finds a significant drift on the returns of mentioned firms. Barber and Loeffler (1993) and Chan (2003) obtain similar results using the *Dartboard* column and headlines from the Dow Jones Interactive Publication Library. Additionally, Berry and Howe (1994) and Antweiler and Frank (2004) study the effect on volatility and volume; although, they obtain different results. The former estimates a positive relationship between the number of news released by Reuters' News Service and volume, but they do not find any relationship with volatility. In contrast, Antweiler and Frank (2004), using data from Yahoo document that a higher news presence entails higher volatility and volume. In a related study, Chae (2005) explores the effect on volume depending on the firm's information asymmetry before the news arrival. He shows that firms with a higher degree of asymmetry present a lower increase in volume before the release, whereas the relationship reverses after it. With respect to adverse selection, Ranaldo (2005) argues that news items causing large price disruptions enlarge spreads albeit others enhance liquidity. Meanwhile, Tetlock (2010) using a sample containing more than 2.2 million news pieces finds that price impact decreases after their publication. Contrarily, Riordan et al. (2013) identify that adverse selection increases after Reuters releases some information, especially if it conveys negative news.

My paper departs from this literature by concentrating on scheduled news. Note that, while scheduled news provide (or attempt to) additional information to the market, press releases do not produce information, but they distribute it to more investors. The consequences of both types of news might be very different.

Finally, this paper also contributes to the theoretical literature on how public information affects prices. Even if intuitively, public news do not increase the degree of information asymmetry, Kim and Verrecchia (1991, 1994, 1997) show that it can happen if informed traders own a better information processing technology than the market maker. In contrast, Tetlock (2010) formalizes the intuitive result that the existence of a public release decreases information asymmetries and consequently, price impact. These models consist of three periods: before, during, and after the announcement date. In each period, informed and uninformed investors receive a liquidity shock and decide how much to buy of the risky asset to maximize their utility function. As a result, there is a unique price, instead of one bid and one ask price. Besides, these models do not deliver any prediction about the dynamics after the announcement.

In contrast, I construct 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.[2] Regarding risk-aversion, Holden and Subrahmanyam (1994) introduce risk-averse informed traders in Kyle (1985)'s model. Nonetheless, 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.

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

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 disaggregated data, and several tables which contain the most granular information. All these documents are available online and they are easily machine readable as the format is constant across weeks.[3]

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 1 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 on oil stocks on its lag and week-of-the-year dummy variables. The regression residuals are my measure of news which do not contain the predictable component due to autocorrelation or seasonality.

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

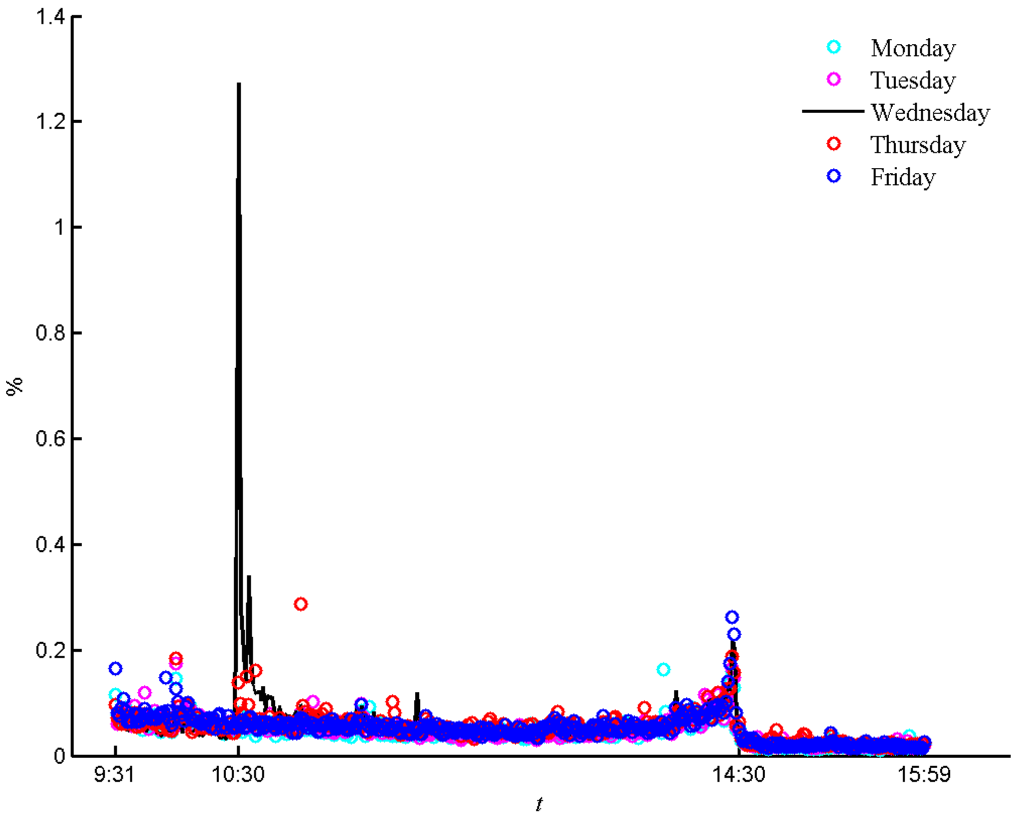

Figure 1: Intraday Variance Decomposition. This figure present the variance of the nearest-maturity future by minute normalized to sum to 100% every week. Overnight returns are not considered.

Furthermore, to ease the interpretation, I normalize it to have standard deviation equal to one. I label the resulting variable as ΔInv.[4] 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 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 construct an index of oil use. Using the Bureau of Economic Analysis Input-Output tables from 1999, I obtain how many dollars of this commodity a particular industry needs to produce one dollar of

[4]Results along the paper are maintained if I consider raw inventory variations.

total output. In general, industries require between three and ten cents on 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 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.[5] Regarding data processing, I apply the filters proposed by Holden and Jacobsen (2014) and I aggregate the variables inside a 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. Further, to prevent minutes with low activity to drive the results, I attach to each observation a weight proportional to the square root of the number of operations. Nevertheless, results are robust to the weighting scheme, including equal weights.

From the final sample, I construct three main variables: midpoint returns, proportional effective bid-ask spread (hereafter 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.

Table 1 presents summary statistics of the aforementioned variables between 10:00 a.m. and 11:00 a.m. It shows that oil firms are similar to non-oil firms; however, they

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

	Return (bps)		Spread (bps)		Volume (thousands)	
	Oil	Non-Oil	Oil	Non-Oil	Oil	Non-Oil
Mean	-0.050	-0.025	3.460	4.865	34.443	69.297
Standard dev.	12.085	13.774	2.694	18.723	42.814	280.042
Median	0.000	0.000	2.772	3.364	22.300	15.700
25%	-5.184	-4.975	1.785	2.167	7.826	3.925
75%	5.168	4.930	4.285	5.531	45.200	53.109
N	390,028	4,164,529	390,028	4,164,529	390,028	4,164,529

Table 1: Summary statistics. The table describes summary statistics of the main market variables divided in two groups: oil and non-oil firms. Oil firms are defined as those who require more than \$0.2 of crude oil to produce \$1 of output. The sample is described in Section 2. It consists of 50 randomly chosen firms stratified in two volume buckets from January 2007 to June 2013.

tend to have lower spread and volume. Additionally, I observe that volume is highly volatile in non-oil firms as they are a heterogeneous group. Regarding the composition, the oil sector contributes with 390,028 observations to the sample, which accounts for 9% of the sample (four firms).

3 Empirical Facts

The informational advantage some traders have usually relates to firm idiosyncrasies such as future investments, or the state of current projects. In contrast, oil inventories provide information relevant at the aggregate level, as they reflect global demand and supply. Thus, we can interpret the petroleum status report as a public signal independent of the private information. In this case, the market maker and the informed investors learn exactly the same. As a result, standard information-based models predict that spreads and volume are unaffected by the signal.

Figure 2 shows the mean spread and volume on Wednesdays around 10:30 a.m., for oil firms. We observe that spreads double at the time of the release, and then decrease to a level around 30% higher than the initial one. Similarly, volume is 40% higher than the initial level some seconds from the announcement, after an initial increase of more than 100%.

These results contradicts the outstanding theoretical literature, unless there is another

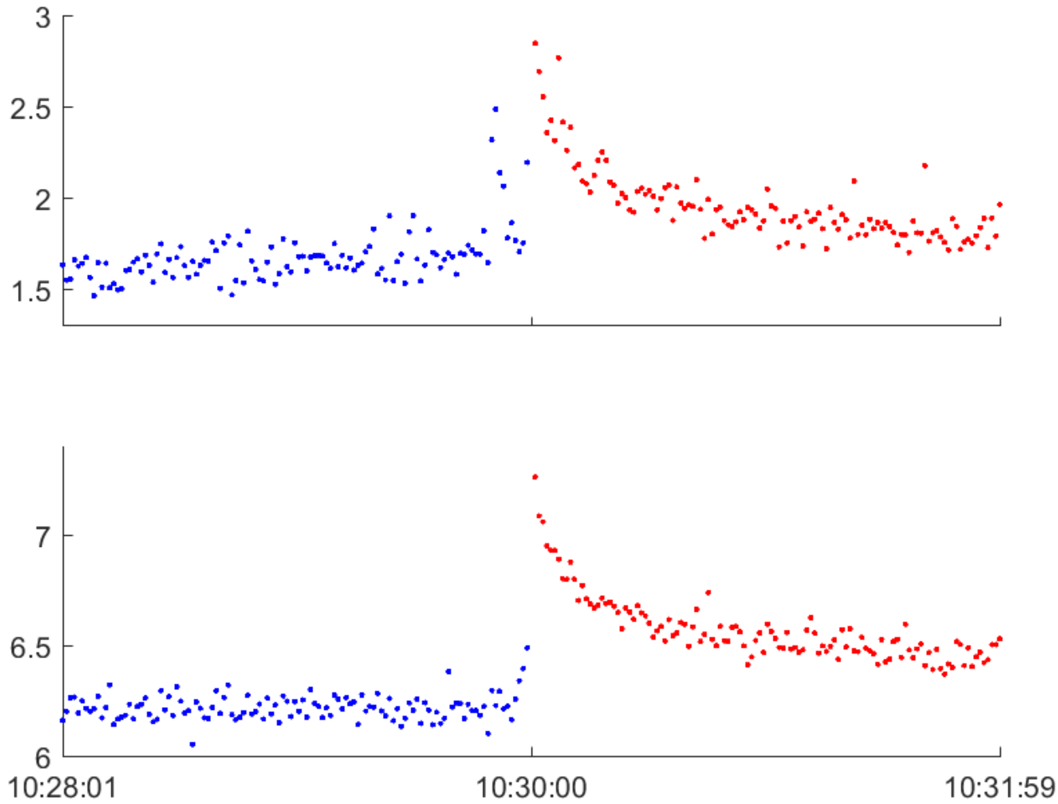

Figure 2: Spread and Volume Around 10:30 a.m. This figure presents the mean spread (top), and volume (bottom) for each second, around 10:30 a.m. The sample is described in Section 2, and this graph only considers Wednesdays and oil firms.

layer of asymmetry. For instance, Kim and Verrecchia (1997) consider that informed investors can extract information from the public signal whereas the market maker cannot. Although this characterization is reasonable to analyze earnings announcements, it is hard to justify in the context of the oil inventory release, since its layout and information content are clear. Yet, I can test these models as they provide a sharp prediction: the increase on spread and volume depends on the realization of the signal, e.g. the change on inventories. In particular, it should be positively correlated with its absolute value.[6]

In Figure 3, I divide the sample into days with an extreme inventories increase ($\Delta Inv >$ .43), an extreme decrease ($\Delta Inv < -.43$), and a small change ($-.43 < \Delta Inv < .43$).[7] If

[6] Kim and Verrecchia (1991) predict that the increase in volume is proportional to the signal realization.

[7] ΔInv has mean zero and standard deviation equal to one. I select the threshold such that each group has the same number of observations. The jump at 10:30 and the pattern is robust to different

informed investors learn more about the public signal, the increase at 10:30 a.m. should be bigger in the first two groups. In contrast, the three columns are identical. This lack of dependence would also be present if the measure of news (ΔInv) is noisy and does not provide information, even if traders have asymmetric technologies. In this situation, price movements at 10:30 a.m. should not be related to ΔInv. Nevertheless, Figure 4 shows that prices and change on inventories are negatively correlated. The correlation equals -0.28, and it is significant at the 1% level. In conclusion, the data does not support the mechanism proposed by Kim and Verrecchia (1997).

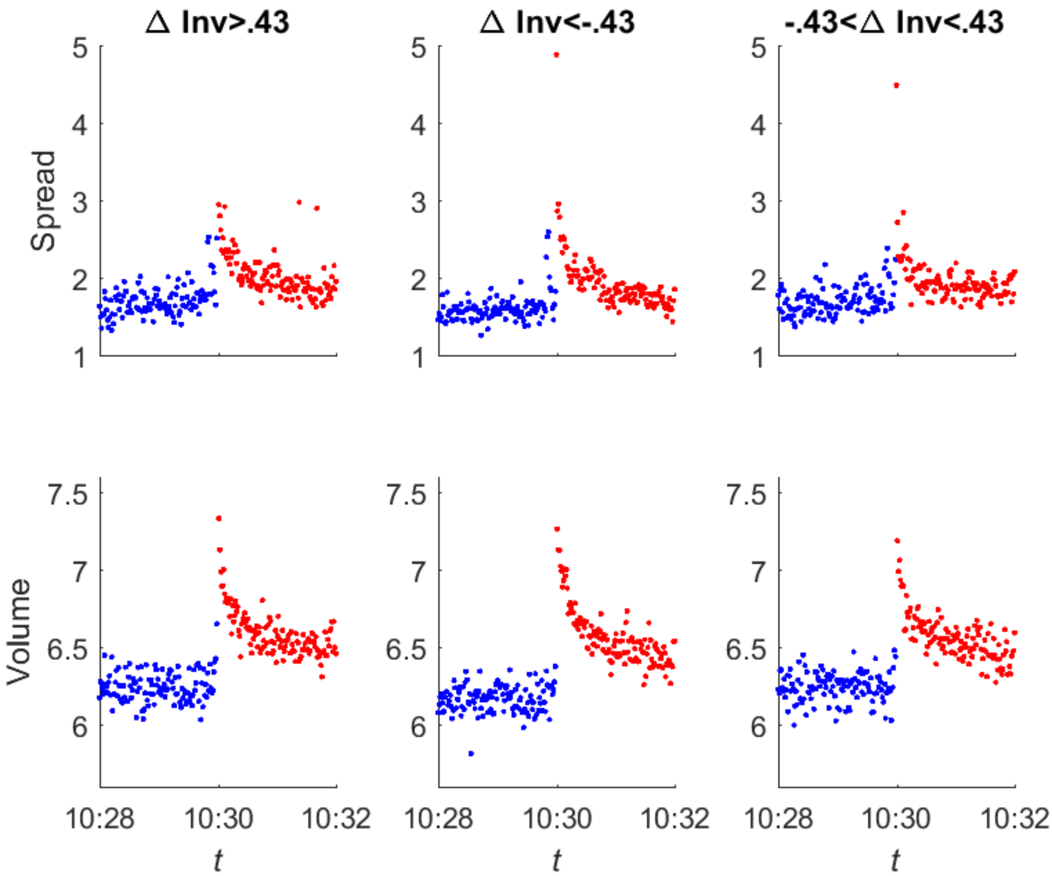

Figure 3: Spread and Volume Around 10:30 a.m. This figure presents the mean spread (top), and volume (bottom) for each second, around 10:30 a.m. In the first column I plot the mean across days in which ΔInv is exceeds .43 while the middle column present the means when this measure is less than -.43. Similarly, the mean of the remaining days is presented in the right column. The sample is described in Section 2, and this graph only considers Wednesdays and oil firms.

thresholds (.75,1.65,2). The dispersion of the points, however, changes across graphs as each mean is estimated with a different number of observations.

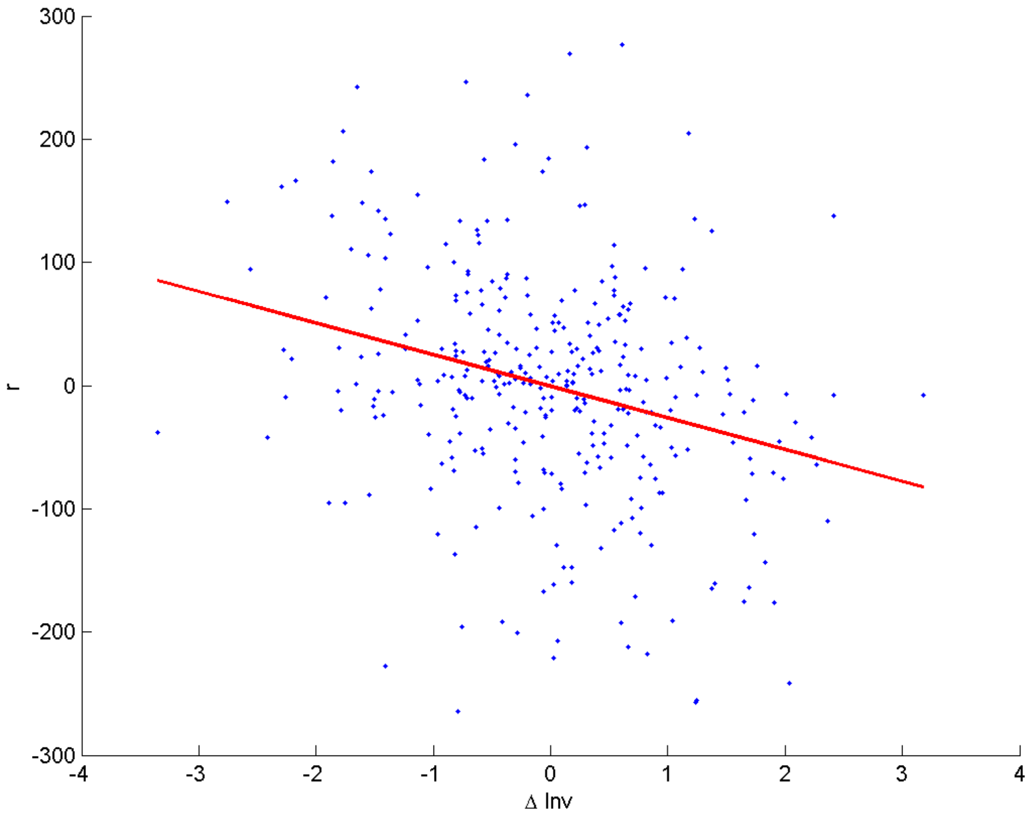

Figure 4: Returns at 10:31 a.m. and Inventory Changes. This figure plots the return between 10:31 a.m. and 10:30 a.m, against the change on inventories. The red line depicts the fitted value of a linear regression between these two variables.

Some times during the trading day are special such as 11:30 when the London stock exchange closes. A similar event at 10:30 a.m. can justify the extreme changes we observe in the data. To rule this explanation, and provide statistical significance to previous results, I estimate the following regression:

$$y_{i,t} = \mu + \delta_t + \alpha_0 TP_t + \alpha_1 \Delta Inv_t + \alpha_2 Wed_t + \alpha_3 Wed_t \cdot \Delta Inv_t + \\ (\beta_0 \Delta Inv_t + \beta_1 Wed_t + \beta_2 Wed_t \cdot \Delta Inv_t) \cdot TP_t + u_{i,t}. \quad (1)$$

I use the sample from 10:25 a.m. to 10:28 a.m. as the *before* period ($TP = 0$), and 10:30 a.m. and 10:31 a.m. as two different *after* periods ($TP = 1$). I divide the *after* period to distinguish the temporary hike we observe just at the release time, and the more stable effect. The specification takes into account intraday dynamics by comparing Wednesdays ($Wed = 1$) with the remaining weekdays ($Wed = 0$). In addition, I include month-year

	Spread (bps)		Volume (logs)	
	10:30	10:31	10:30	10:31
$Wed \cdot TP$	1.424***	0.770***	0.406***	0.335***
	(0.335)	(0.239)	(0.0874)	(0.0748)
$Wed \cdot TP \cdot \Delta Inv$	0.0719	0.0358	0.0854	0.0659
	(0.221)	(0.221)	(0.0707)	(0.0635)
Observations	32,467	32,467	32,467	32,467
$\beta_1 = \beta_2 = 0$ (p-value)	0	0	0	0

Table 2: Oil Firms Reaction. I restrict the sample to those firms who need more than 0.2$ of oil to produce 1$ of output. The table presents the estimates for the model in Equation (1). The first row indicates the dependent variable. The second row defines the *after* period. The third and fifth row are the estimates of β_1 and β_2, respectively. Standard errors are clustered at the monthly level and they are presented in parenthesis. Significance levels are indicated by *** (1%), ** (5%) and *(10%).

fixed effects (δ_t) to control for low frequency movements. Since ΔInv varies on a weekly basis, and to account for some correlation across weeks, I cluster standard errors at the monthly level.

The main coefficients of interest are β_1 and β_2. The first one measures the effect of the release itself; that is, if the information on the report does not deviate from its expectation. Meanwhile, the second estimate captures the response to the content. Under the assumption of independence between the public and private signals, seminal market microstructure papers predict that $\beta_1 = \beta_2 = 0$. [8] Meanwhile, if agents differ on their information processing technologies, we expect $\beta_1 > 0$ and $\beta_2 < 0$. In Table 2, I present the results for the subsample of oil firms. The estimates confirm the results from the previous graphs. Precisely, the spread increases by 1.42 bps at 10:30 a.m. whereas it is 0.77 bps higher at 10:31 than before the announcement. Regarding volume, it boosts by 40.6% and then decreases up to a level 33.5% higher than the one prior to the release. These effects do not depend significantly, on the report's content.

An effect public news has that is independent of their content is a reduction in uncertainty. If this mechanism is the main driver of spread changes, the report should not affect non-oil firms, as the stock of oil is not informative about their fundamental value.

[8] For instance this prediction is obtained if you include a public signal in the models proposed by Glosten and Milgrom (1985), Kyle (1985), Easley and O'Hara (1987), Easley et al. (1997)

	Spread (bps)		Volume (logs)	
	10:30	10:31	10:30	10:31
$Wed \cdot TP$	0.150	0.408	0.044	0.025
	(1.586)	(1.049)	(0.039)	(0.038)
$Wed \cdot TP \cdot \Delta Inv$	0.574	-0.846	-0.006	-0.003
	(1.582)	(0.934)	(0.034)	(0.042)
Observations	347,046	346,979	347,046	346,979
$\beta_1 = \beta_2 = 0$ (p-value)	0.411	0.648	0.334	0.693

Table 3: Non-oil Firms Reaction. I restrict the sample to those firms who need less than 0.2$ of oil to produce 1$ of output. The table presents the estimates for the model in Equation (1). The first row indicates the dependent variable. The second row defines the *after* period. The third and fifth row are the estimates of β_1 and β_2, respectively. Standard errors are clustered at the monthly level and they are presented in parenthesis. Significance levels are indicated by *** (1%), ** (5%) and *(10%).

Table 3 presents the estimates of Equation 1 for non-oil firms. In line with the information channel, neither volume, nor spreads of these firms react to the release.

The uncertainty faced by informed investors and the market maker is irrelevant in previous information-based models, as most of them consider risk-neutral investors. On the other hand, inventory models take into account risk-aversion, but they predict a negative relationship between fundamental volatility and spreads. In this regard, I propose a model in which there is asymmetric information, and traders are risk averse which can explain these facts.

4 Model

Previous literature focuses on the effect of a public signal on the market through information instead of changes in investor composition. Equivalently, these papers assume that the probability of a trade from an informed agent is the same before and after the public signal. In this section, I develop a model in which some informed investors do not trade before the public signal because of risk aversion. After the realization of the signal, which eliminates part of the uncertainty, they enter the market. As a consequence, the bid-ask spread increases, even if the market maker and the informed trader obtain the same information from the public signal. In addition, volume also increases after the revelation of the public signal, regardless of its content.

4.1 Set-up

The set-up of the model is similar to Glosten and Milgrom (1985) and Easley et al. (1997). There is an asset being traded with common value $v = \omega + \mu + \varepsilon$. ω is the part of the asset value that is subject to asymmetric information, and μ is the content of the public signal, e.g. the change in inventories. Traders do not know the public information until time $t = t_R$, while some of them observe ω from $t = 0$. Additionally, ε is a noise with zero mean and variance σ_ε that reflects the part of the fundamental on which the signals do not provide information. Following Easley and O'Hara (1987), ω can take two values: σ_ω with probability $\frac{1}{2}$ and $-\sigma_\omega$ otherwise. The assumption of a binary distribution simplifies the model since, under risk neutrality, agents who observe ω always trade in equilibrium. Therefore, investors decision to trade after the signal realization, and not to do so before, is completely due to risk aversion.[9]. Analogously, μ has zero mean, and variance given by σ_μ without specifying any particular distribution.

There are two types of traders in the market: a mass δ of noisy traders who buy and sell randomly with probability 1/2, and a mass $1 - \delta$ of informed traders who know the realization of ω and maximize their expected utility. From now on, I assume that both types of agents are in the market ($0 < \delta < 1$). At each time t, one random trader arrives at the market. Once she arrives; she buys, sells or does not trade, and she leaves the market. This assumption is common to the aforementioned seminal papers to preclude strategic timing, which complicates the analysis. 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.

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 $p(-1)$ is the bid price. Note that, as it is common in the literature, informed

[9] If ω is normally distributed, some informed traders would not trade because they have a high risk aversion or because the value of ω is low. In this case, the conclusions of the model are maintained under a wide range of parametric values but most of the expressions cannot be obtained in closed form.

investors trade, at most, one unit.[10] Moreover, informed investors are heterogeneous in their risk aversion. To model this heterogeneity, I consider that γ is randomly distributed with c.d.f $F(\cdot)$. Further, to ease the exposition, I consider that $F(\cdot)$ is differentiable and strictly increasing with support equal to $\mathbb{R}^+$. While there is no direct evidence on informed investors' preferences, Koijen (2014) shows that fund managers present heterogeneous 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 willingness to take 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. Precisely, if market makers are heterogeneous in their risk aversion, and there is a positive mass of them that are risk neutral, those who are risk averse would exit the market. As a consequence, in equilibrium, the relevant market maker is risk neutral.

I consider the model in discrete time, such that an investor arrives at each period $t = \{0, 1..., T\}$. The timing inside a period is as follows: firstly, and only in period t_R, the public information is revealed. Secondly, the market maker posts an ask and a bid price. Finally, a trader arrives at the market and buys, sells or leaves. Afterwards, the next period starts.

The set-up of the model closely relates to the aforementioned papers. 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.

4.2 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

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

given by:

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

where I do not consider the case in which $B^* > A^*$, as it does not occur in equilibrium.[11] 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, conditional on the whole sequence of transactions up to time t, $\mathcal{H}_t$. In the reference models, the number of sales and buys are a sufficient statistic for the whole sequence. In contrast, the order of events is important in my model, as a buy after a sale -low price- is more likely to come from an informed agent than a buy after another buy -high price. Nevertheless, similar to previous literature, in equilibrium, prices and transactions convey the same information. As a result, 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 $d_t = 1$ ($d_t=-1$).[12] Importantly, Π_t^+ can easily be obtained recursively, even if it depends on the whole sequence of buys and sales. Actually price dynamics result from this function of history, as time does not affect prices directly. More specifically, prices increase after a buy, and decrease after a sale. No-trades, however, push prices in a different direction depending on the potential information rents. If the distance between A^* and σ_ω is smaller than the one between B^* and $-\sigma_\omega$, it is more likely that an investor does not trade if $\omega=\sigma_\omega$. As a consequence, the market maker updates her belief about the fundamental, and quotes, upwards.

[11] Along the paper, I use stars to represent equilibrium quantities and actions.

[12] To avoid extra notation, I denote by d, the current order (d_t), and the reaction of informed investors (d_i).

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.

4.3 Equilibrium with Public News

An equilibrium in this model is given by the triples $\{d_i^*(A_t, B_t), A^*(\mathcal{H}_t), B^*(\mathcal{H}_t)\}_0^{t_R-1}$, and $\{d_i^*(A_t, B_t, \mu), A^*(\mathcal{H}_t, \mu), B^*(\mathcal{H}_t, \mu)\}_{t_R}^T$ where d_i^*, A^*, and B^* are the agents' best responses. The main difference with the previous case is that these responses depend on μ after t_R. Proposition 1 describes the equilibrium prices and strategies. Additionally, Corollary 1 ensures that the spread is 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 $t < t_R$,

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

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

If $t \geq t_R$,

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

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

where Π_t^+ (Π_t^-) *are the market maker beliefs that* $\omega = \sigma_\omega$ *conditional on all the trades up to* t *and* $d_t = 1$ $(d_t = -1)$.

Corollary 1.

$$\sigma_\omega > A_t^* > B_t^* > -\sigma_\omega \ \forall t$$

Figure 5 sketches the main mechanism of he model, when $\omega = \sigma_\omega$. Before the signal realization, informed agents face significantly more risk. In this situation, traders above $\underline{\gamma}$ leave the market without trading. After the signal, however, traders between $\underline{\gamma}$ and $\overline{\gamma}$, as well as those below $\underline{\gamma}$, transact. As a consequence, adverse selection costs rise which leads the market maker to set a wider spread.

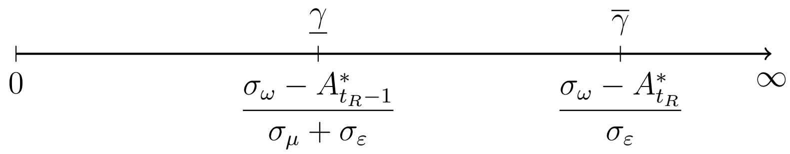

Figure 5: Mechanism. This figure sketches the main mechanism in the model when $\omega = \sigma_\omega$.

Regarding price informativeness, as the market maker learns from previous trades, information rents reduce. Thus, less informed investors participate, reducing the information content of trades. Nonetheless, the market maker learns also from the no-trade periods which provide much more information when prices are close to one of the possible values of ω. As a result, in this model price informativeness can increase at the same time as spreads decrease

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 decrease volume since noise traders always transact. Instead, an increment in σ_ε decreases spreads and volume since informed agents are less likely to participate.

4.4 Model Predictions

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

Along the paper I consider three different variables: 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 $m_t = \frac{A_t+B_t}{2}$.[13] 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 afterwards. 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. Thus, it has been widely used in the literature as a measure of information asymmetry. Proposition 3 states that as soon as μ becomes common knowledge, the bid-ask spread increases. Besides, this rise is independent of μ as it is a result of a composition effect. If the informed investors do not know μ they are less willing to act on their information since it entails a high risk. Once they know μ these investors start trading. As a result, the market maker needs to widen the spread to break even in expectation. After the release of the signal, the market maker continues learning about ω from trades. Hence, the bid-ask spread decreases and converges to the one in the model without news. In this model, the limiting spread is zero, but it could be positive by including some additional friction such as transaction costs.

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,

$$lim_{T\to\infty} A^*_{1,T} - B^*_{1,T} = lim_{T\to\infty} A^*_{0,T} - B^*_{0,T} = 0 \tag{2}$$

[13] In the empirical part, the return is defined as the difference in logs. While theoretical results do not change if we consider that all variables are in logarithmic form, expressions become more convoluted.

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 afterwards.

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:

$$lim_{T\to\infty}\mathbb{E}\left(|d^*_{1,T}|\right) = lim_{T\to\infty}\mathbb{E}\left(|d^*_{0,T}|\right) = (1-\delta) \tag{3}$$

To show the different hypotheses graphically, I plot the dynamics of the three relevant quantities in Figure 6. I consider a χ^2_3 distribution for the risk aversion coefficient and some specific values for the parameters. Specifically, I set the proportion of informed traders (δ) to 0.4. I fix the variance of the private information, the public information, and the residual noise ($\sigma^2_\omega, \sigma_\mu, \sigma_\varepsilon$) to $0.5, 1$ and 0.25 respectively, and the release time (t_R) to 31. Moreover, I present the results for two different values of μ; 1 and -1. Since results depend on the whole path of buys and sales, the figures show the mean across $10,000$ Monte Carlo simulations. To compare the two scenarios, I use the same realization of all random variables in the model with and without news, in each simulation.

The top graph shows the midpoint returns for the two different values of μ. We observe that returns are zero at every point in time, except at the release time when they equal μ. The middle plot indicates that the bid-ask spread increases at t_R, and declines afterwards. Moreover, the lines for the different values of μ overlap, indicating that the whole dynamics are independent of the content. We recognize the same pattern for volume in the bottom graph.

5 Empirical Analysis

To test the model hypotheses, I follow 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

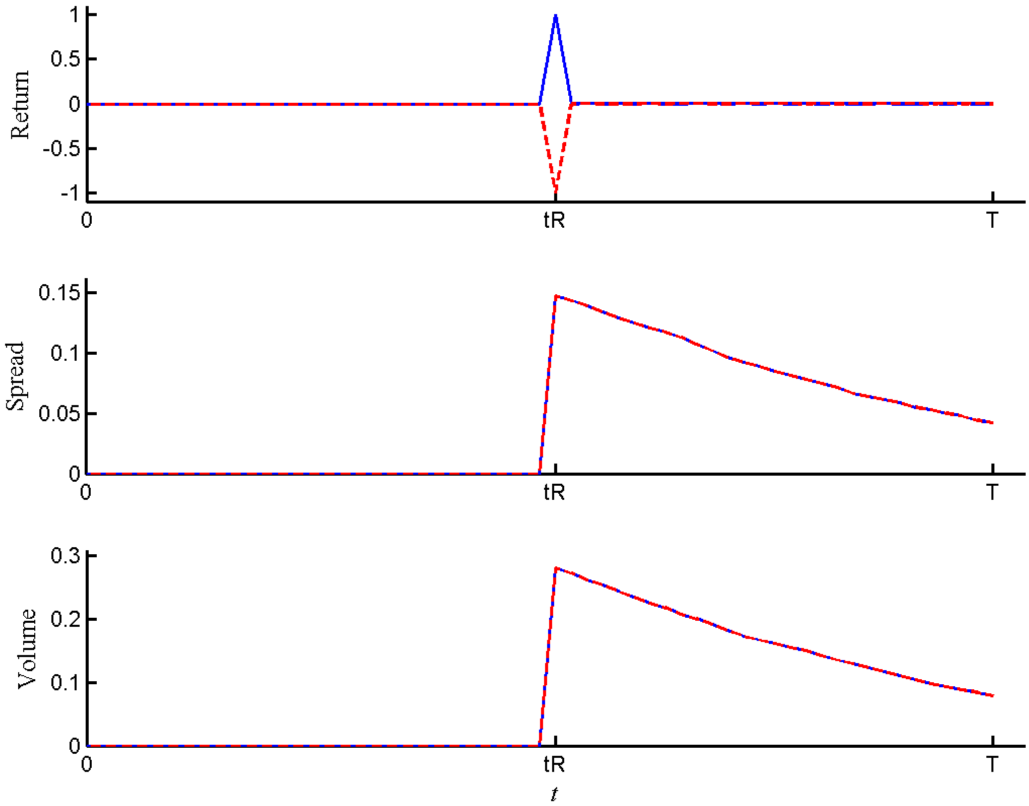

Figure 6: Theoretical Predictions. This figure depicts the difference between a day with news and one without. To obtain the different dynamics, I consider $\gamma_i \sim \chi^2_3$, $\delta = 0.4$, $\sigma^2_\omega = 0.5$, $\sigma_\mu = 1$, and $\sigma_\varepsilon = 0.25$. Moreover, the initial time is $t = 10{:}00$, the release time $t_R = 10{:}30$, and the period length is one minute. The red line represent the case when $\mu = -1$ whereas the blue line corresponds to $\mu = 1$.

Wednesdays with the other weekdays. To be precise, I estimate the following equation,

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

where Oil_i is a dummy variable that takes value one if the firm belongs to the oil sector and zero otherwise.

The main advantage of this empirical approach is that it provides estimates of the effect using the cross-section and interday dimensions without relying on the intraday variation. Accordingly, I estimate the effect of the inventory release minute by minute. The result of this strategy is the full intraday dynamics of a public announcement. To identify the effect of interest, I make two main assumptions: First, non-oil firms are not affected by the oil report. Although I cannot test the validity of this assumption, Table 3 suggests that it is likely to hold. Second, the presence of the report release only affects Wednesdays. In Section 6.6 I relax this assumption, yet results are unaltered.

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. In order to characterize the complete dynamics, I estimate Equation (4) for every minute independently. Hence, observations are at the daily-firm level. To focus the attention around 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 imply 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 7 I depict the estimates of γ_2 and their 5% confidence intervals. They confirm the previous hypotheses. Precisely, prices rise immediately by 5.5 bps, if inventories drop by a one standard deviation, but they remain constant afterwards. 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 also differentiate the Weekly Petroleum Status Report from other announcements. For example, prices present a drift after earnings announcements (see Ball and Brown, 1968; Beaver, 1968); however, they do not after the release of this report. Likewise, whereas prior literature finds evidence of anticipation to macro-news releases (see Bernile et al., 2016; Kurov et al., 2016, and references therein); we do not observe prices react to inventory data before its disclosure. Additionally, this evidence supports the assumption that informed investors lack prior information about oil stocks. Otherwise, these investors would buy before an

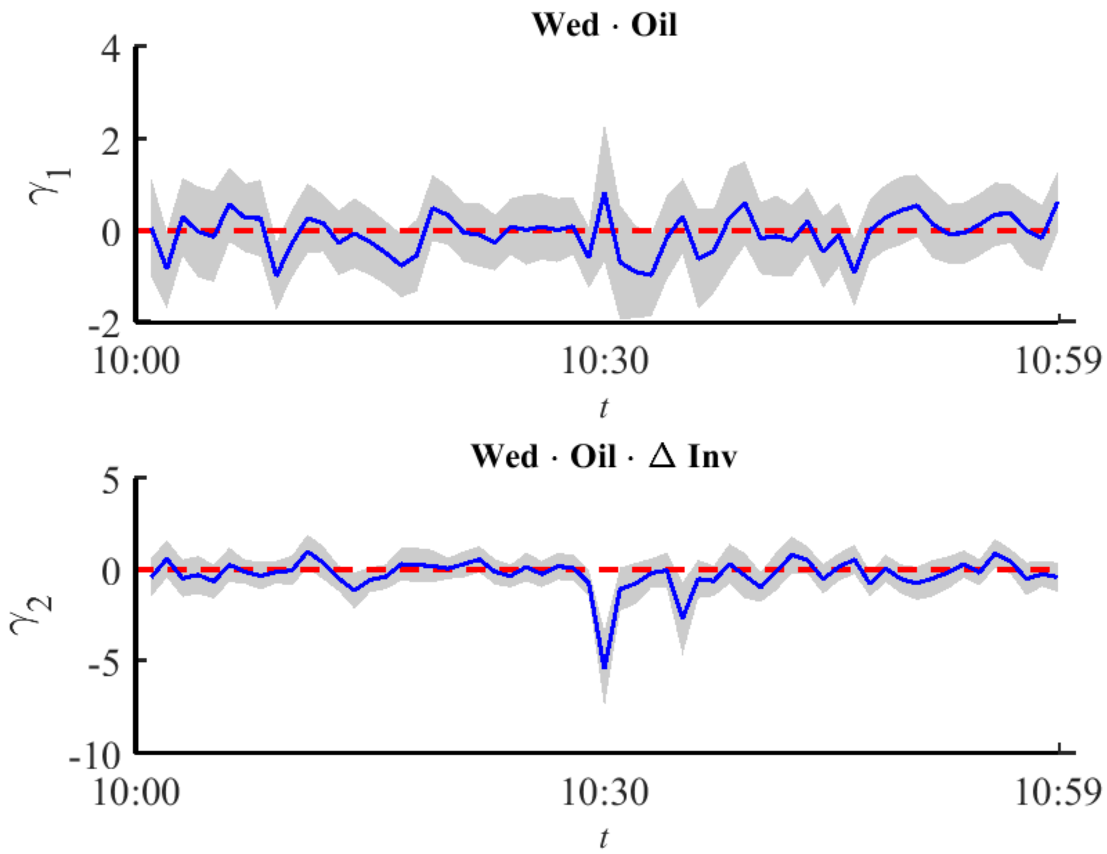

Figure 7: Estimates Return. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (4) using midpoint returns as a dependent variable. I estimate the coefficients independently for every minute exploiting variation across days. The upper plot corresponds to γ_2, while the lower one refers to γ_1. I describe the sample in Section 2. The red dashed line indicates the zero.

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 on inventories and returns before the release, which is not consistent with the data.

Regarding the spread, the model predicts an increase after the release, independently on the information content ($\gamma_2^{10:30} > 0$). Moreover, this effect lasts for some periods as the market maker learns about the private signal although, it finally vanishes. Accordingly, 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 8 shows that the spread increases at 10:30 a.m. by 2.79 bps and it slowly decreases, becoming insignificant after thirty minutes. At the same time, the model predicts that the content of the report is irrelevant for the spread. The lower plot confirm that oil inventory changes do not affect the spread.

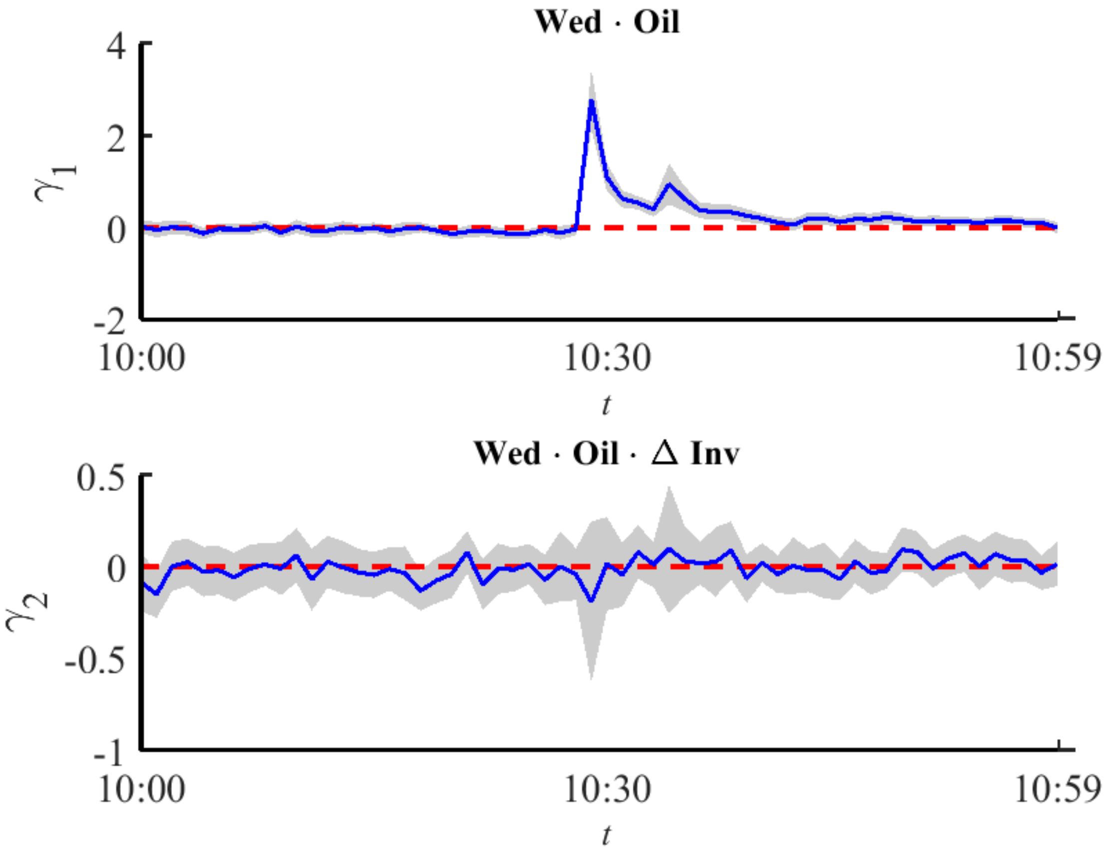

Figure 8: Estimates Spread. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (4) using proportional effective bid-ask spreads as a dependent variable. I estimate the coefficients independently for every minute exploiting variation across days. The upper plot corresponds to γ_2, while the lower one refers to γ_1. I describe the sample in Section 2. The red dashed line indicates the zero.

With respect to volume, the mechanism behind the model implies a strong comovement between the effect on this variable and the spread. Figure 9 present the estimates of Equation (4) with volume as a dependent variable. We observe a similar pattern to the one in the analogue figure for the spread (Figure 8). Precisely, volume rises by 32% at the moment of the announcement and it decreases afterwards. 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.[14] Regarding the report's content, it does not affect the increment in volume, or the posterior dynamics, as the model suggests.

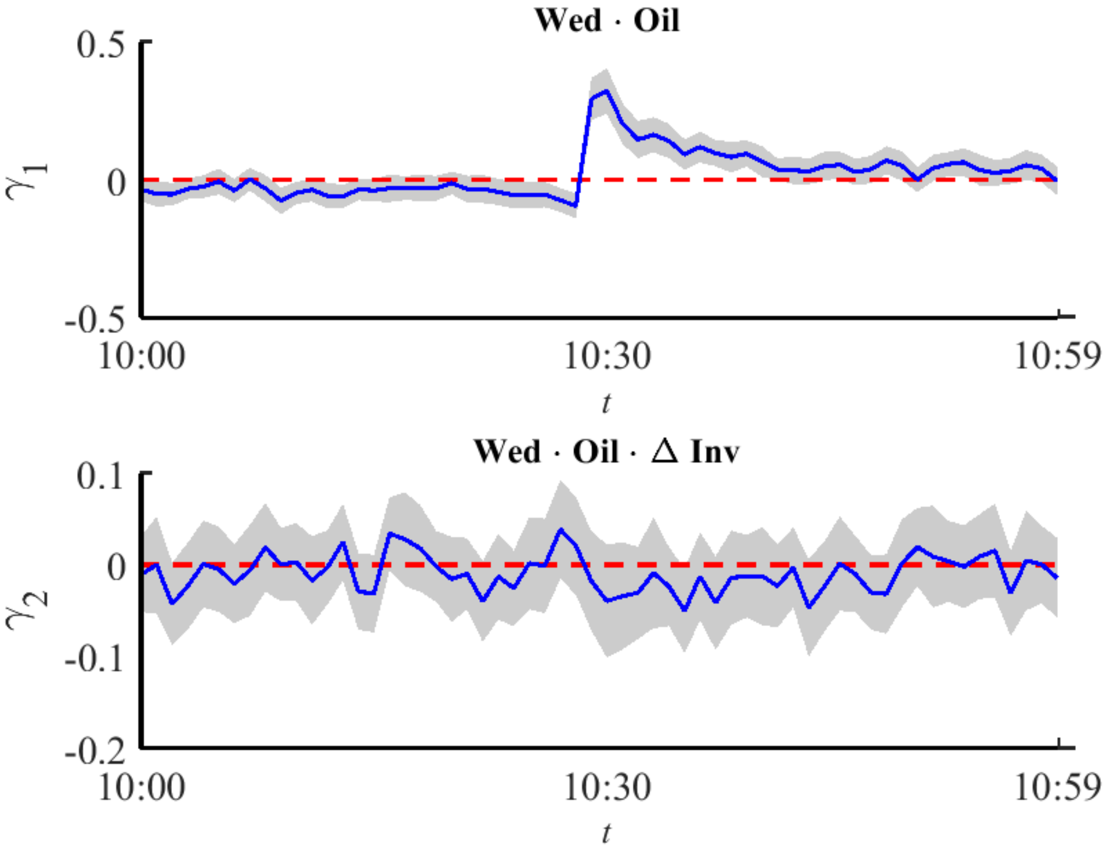

Figure 9: Estimates Volume. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (4) using number of transactions as a dependent variable. I estimate the coefficients independently for every minute exploiting variation across days. The upper plot corresponds to γ_2, while the lower one refers to γ_1. I describe the sample in Section 2. The red dashed line indicates the zero.

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 of returns, but it takes some minutes in terms of volume and spreads. However, this evidence only captures the short-term. To test the limiting results, in Figure 10 I plot the estimate for the last hour of the trading day. The graphs confirm the dissipation of the effects, as estimates are not significant in all cases.

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

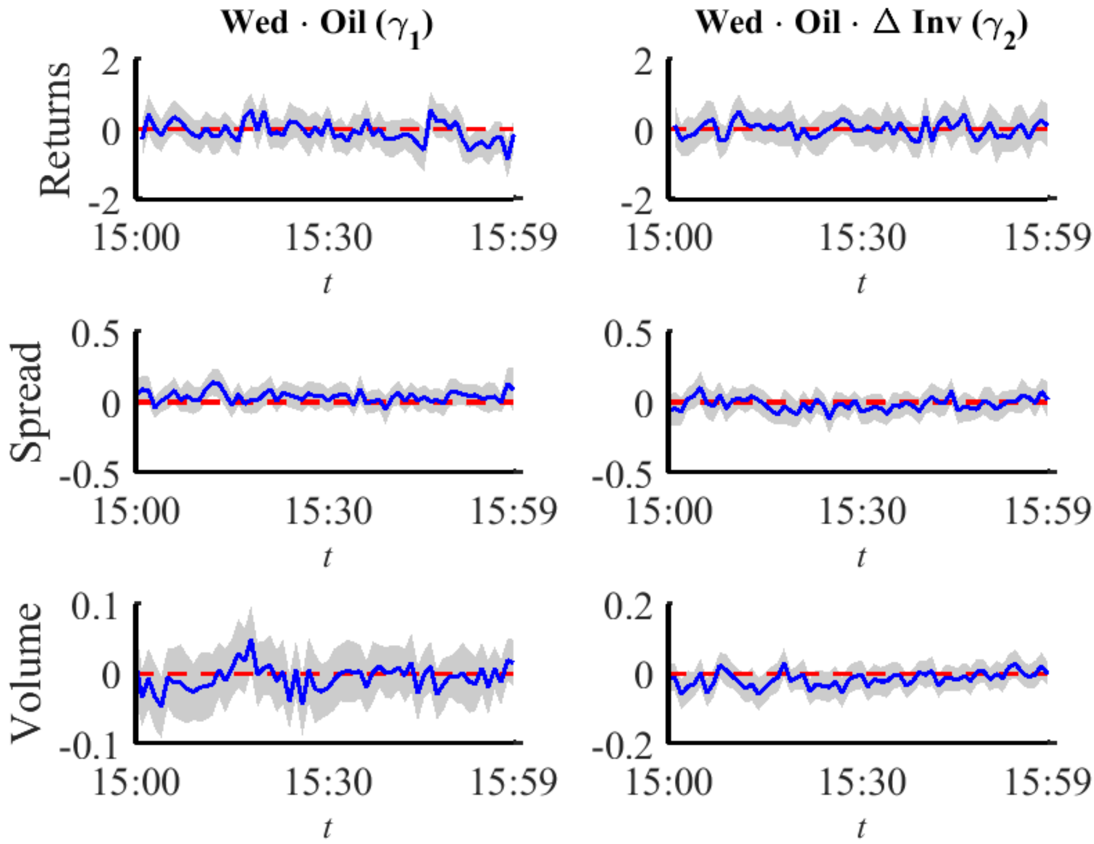

Figure 10: Afternoon Sample. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (4). Each row consider a different dependent variable, from top to bottom: midpoint return, proportional effective bid-ask spread, and number of transactions. I estimate the coefficients independently for every minute exploiting variation across days. The left-hand side plots correspond to γ_2, while the right-hand side ones refers to γ_1. I describe the sample in Section 2. The red dashed line indicates the zero.

In sum, the empirical evidence supports the theoretical model. Nevertheless, there might be interesting hypotheses that I leave outside the model for simplicity, such as the symmetry of the signal, or a constant level of noise traders before and after the release. I explore these possibilities empirically in the next section. I provide several robustness checks to ensure that the identification assumptions are reasonable. Finally, results are maintained using a different sample on stocks and also using data on the ETF market.

6 Additional Results

6.1 Asymmetric and non-linear reaction to news

Along the paper, I assume a symmetric effect of the report's information. 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 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; instead, optimism does not have an effect. Similarly, Akhtar et al. (2011) finds that Australian stocks only react to negative consumer sentiment announcements. 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) find evidence supporting this theory in the case of earning 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.

In Figure B.1, 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. [15]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 is similar to the baseline case.

[15]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.

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.[16]. Then, I estimate the following regression:

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

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 B.2 shows that results remain unchanged under a binary identification. Further, I cannot reject that the effect is symmetric. This evidence reinforce the previous findings on the independence between the report content and the reaction of volume and spread.

6.2 Increase of noise traders

I propose a model in which volume increases because informed traders become more active in the market. However, additional noise traders might enter the market to rebalance their portfolios after the public signal. In fact, Vieru et al. (2006) and Kaniel et al. (2012) show that individual investors trade after earnings announcements in the opposite direction to the news' content. Contrarian trading is consistent with investors adjusting their portfolios to maintain a fixed proportion of the capital in each asset.

If portfolio rebalancing causes the observed increase in market activity after the report release, the increment in volume should be proportional to the absolute change in prices; thus, to the absolute change in inventories. To explore this alternative explanation, I estimate the following model:

$$y_{i,t} = \mu + \delta_t + \theta_0 Oil_i + \theta_1 \left|\Delta Inv_t\right| + \theta_2 Wed_t + \theta_3 Wed_t \cdot \left|\Delta Inv_t\right| + \\ \left(\gamma_0 \left|\Delta Inv_t\right| + \gamma_1 Wed_t + \gamma_2 Wed_t \cdot \left|\Delta Inv_t\right|\right) \cdot Oil_i + \varepsilon_{i,t}. \quad (7)$$

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

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. In contrast, if the findings from the baseline specification are due to portfolio adjustments, the estimate of γ_1 would become insignificant when volume is the dependent variable.

Figure B.3 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. A possible explanation is that investors rebalance their portfolios at a much lower frequency. Alternatively, they might not trade just after the publication of the report to avoid high spreads.

6.3 Different weekdays as controls

To estimate the dynamic effect of the Weekly Petroleum Status Report, I rely on the parallel trend assumption. In other words, I assume that any differences between oil and non-oil firms on Mondays, Tuesdays, Thursdays and Fridays remain constant on Wednesdays, except for the data publication. However, there might be other factors that affect these days at 10:30. Actually, on Thursdays the Energy Information Administration publishes a similar report on gas which is very likely to affect the oil market and specially oil firms. Even if Figure 1 suggest that oil prices do not move abnormally the remaining weekdays, it is possible that my results rely on just one weekday.

To tackle this issue, in Figures B.4, B.5 and B.6, I present the estimates of γ_2 and γ_1 on Equation (4), but restricting the control group to one specific day. For instance, the first row compares Wednesdays with Mondays, whereas the second one compares them with Tuesdays. We observe that the effect is present in all specifications. Besides, the magnitudes are identical with the exception of Thursdays.

This evidence indicates that Mondays, Tuesdays and Fridays have parallel trends between them. It is still possible that the presence of the report release affects these weekdays identically. I address this concern in Section 6.6.

6.4 OHLC dataset

Since my available dataset only includes four oil firms, there might be a concern that the results are specific to these firms. Accordingly, I consider a supplementary dataset made up of one minute OHLC data from January 2005 to February 2015 on 265 firms.[17] The dataset includes 235 very liquid US stocks, and the 30 Dow Jones components; thus, the sample overweights energy firms (22 out of 265). Regarding the time period, the initial date makes sure that the report is published on Wednesdays at 10:30 a.m., as there is no information available prior to that date.

While this dataset has the main advantage of being a long and wide panel data, it has a drawback as it does not include information on bid-ask spreads. In fact, it only contains information on the open, close, high and low transaction prices inside a minute, as well as the number of shares traded. Nonetheless, Corwin and Schultz (2012) show that the difference between high and low (HL) is a valid proxy of bid-ask spreads. They develop a method to transform HL into spreads, using the fact that while the variance of return is proportional to the time length, the bid-ask spread is not. In their paper, they assume a constant variance across two consecutive time periods. If this assumption holds, the difference-in-difference strategy would actually capture the effect on spreads without the need of further transformation. Instead, if this assumption does not hold, and the report affects the fundamental volatility the firm's fundamental value; then, the estimate would capture the effect on variance. To see this, consider the difference of HL between two consecutive periods:

$$\mathbb{E}\left[\Delta HL_t\right] = 2\kappa_1(\sigma_t^2 - \sigma_{t-1}^2) + 4\kappa_2\left[\sigma_t f(S_t) - \sigma_{t-1} f(S_{t-1})\right] + \left[f(S_t) - f(S_{t-1})\right]$$

where σ is the fundamental volatility, $f(S) = log(2+S) - log(2-S)$, S is the quoted spread and, κ_1 and κ_2 are constants. Note that under the null hypothesis that the bid-ask spread does not change ($S_t = S_{t-1}$), ΔHL_t is equal to a linear combination of the difference in variance and the difference in volatility. To absorb these terms, in all the specifications

[17] OHLC stands for open, high, low and close. It is a dataset available at different providers, I obtained it from Pi Trading.

I include the square of the open to close return as a control. Therefore, although the quantitative results lack of interpretation, the qualitative results are meaningful.

Table 4 provides some summary statistics of this database. We observe that the sample is more homogeneous and it contains five times more observations. Moreover, volume is lower than in the main dataset because this one includes years with lower market activity.

	Return (bps)		Spread (bps)		Volume (thousands)	
	Oil	Non-Oil	Oil	Non-Oil	Oil	Non-Oil
Mean	-0.032	-0.061	14.744	16.022	24.576	22.229
Standard dev.	13.690	14.030	14.695	13.689	44.765	29.525
Median	0.000	0.000	10.556	12.492	8.825	12.852
25%	-5.670	-6.484	5.492	7.098	2.932	5.100
75%	5.632	6.408	18.948	20.562	24.400	27.800
N	30,610,000	3,077,000	30,610,000	3,077,000	30,610,000	3,077,000

Table 4: Summary statistics 1m. The table describes summary statistics of the main market variables divided in two groups: oil and non-oil firms. Oil firms are defined as those who require more than 0.2$ of crude oil to produce 1$ of output. The sample is described in Section 6.4. Spread refers to the difference between the highest and lowest price inside a minute.

Figure B.7 shows the estimates of Equation (4) using this data. We observe that results are very similar in magnitude and dynamics to the baseline specification, even if both samples are different. Precisely, returns decrease by 6.6 bps after a 1 standard deviation increment on inventories and, volume increases by 41%.

6.5 Oil market

I use stocks' data in my baseline specification because it provides cross-sectional variation, even within the treatment group. However, it is relevant to understand the effect the Weekly Petroleum Status Report has on the oil market. To pursue this aim, I obtain OHLC data for the most liquid oil ETF (USO) and gold ETF (GLD) at the minute level. The dataset is analogous to the one I describe in the previous section in terms of stocks.

I rely on ETFs instead of futures or other derivatives because they track the actual spot price. Moreover, other derivatives present within month variability due to maturity days that might affect the results. Nevertheless, results are almost identical using the closest-to-maturity futures, even if smaller in magnitude. On the other hand, I consider

GLD as a control for two main reasons: First, both ETFs are similar in terms of liquidity. Secondly, oil is not a substitute or complement of crude; instead, it is an investment asset.

As in the baseline specification I estimate the following regression:

$$y_{ETF,t} - y_{GLD,t} = \mu + \delta_t + \gamma_0 \Delta Inv_t + \gamma_1 Wed_t + \gamma_2 Wed_t \cdot \Delta Inv_t + \varepsilon_t$$

minute by minute. Since I lack cross-sectional variation, I work directly in differences. This empirical strategy is more restrictive as the month fixed effects capturethe effect on the difference of factors without variation within a month. Meanwhile, I cluster standard errors at the month level to capture any remaining time series correlation.

In terms of the model, we can interpret oil prices as a firm with very low idiosyncratic volatility (σ_ε) relative to the overall volatility. Consequently, the theoretical prediction is that the announcement should affect strongly the three variables considered compared to the case of individual firms.

Figure B.8 presents the estimates of γ_2 (left) and γ_1 (right) for the three main market variables. The first row shows that returns decrease by 10 bps after a 1 standard deviation increase in inventories. This magnitude is 1.5 times the one for stocks. At the same time, spreads rise by 36 bps instead of 10.3 bps. Meanwhile, volume triples at the time of the release while it boosts by 41% in the case of equity.

6.6 Before and after 2005

The identification in Section 5 relies on the assumption that the existence of the report only affects Wednesdays. One possibility to relax it is to compare the baseline results before and after the EIA starts to release the report at 10:30 on Wednesdays. Unfortunately, the exact date of implementation is not available. However, before 2005 I cannot find any report or news article about oil inventories at the time the data is released nowadays. Therefore, I consider January 2005 as the implementation point. If this date is not the implementation point, we should expect similar results before and after. In contrast, Figure B.9 shows that there are not any effects before 2005. While this evidence supports my assumption about the implementation date, it is only definitive in as far as there is no other factor that affects oil firms on Wednesdays differently in the two periods. Yet, a misspecification of the implementation time would drive estimates towards zero.

To compare both periods, I estimate the following regression Equation using the OHLC dataset:

$$\begin{aligned} y_{i,t} = \mu + \delta_t + \theta_0 Oil_i + \theta_1 \Delta Inv_t + \theta_2 Wed_t + \theta_3 Wed_t \cdot \Delta Inv_t + \\ (\kappa_1 \Delta Inv_t + \kappa_2 Wed_t + \kappa_3 Wed_t \cdot \Delta Inv_t) \cdot Oil_i \\ (\gamma_0 \Delta Inv_t + \gamma_1 Wed_t + \gamma_2 Wed_t \cdot \Delta Inv_t) \cdot Oil_i \cdot \mathbb{1}\{year_t > 2004\} + \varepsilon_{i,t} \end{aligned} \quad (8)$$

where $\mathbb{1}\{year_t > 2004\}$ takes value 1 if the observation takes place on 2005 or afterwards. As in the baseline case, I carry the estimation minute by minute using variation across firms and days. Note that I add a new layer of comparison; in particular, before and after the implementation. Therefore the possible effect the report could have on other weekdays besides Wednesday does not threaten the identification.

I present the estimates in Figure B.10. We can observe that the effect is very similar to the baseline case. Precisely, returns decrease by 6.6 bps after a 1 standard deviation build-up while the report information does not affect spreads and volume. In contrast, returns do not react to the release itself whereas the high-low spread increases by 11 bps and volume rises by 47%. Likewise, the results depict very similar dynamics. The effect on prices is immediate whereas volume and spread present persistent changes.

7 Conclusions

Most market microstructure models predict that spreads would not be expected to increase under the presence of more public information. Nevertheless, I present empirical evidence of an increment on spreads right after the Weekly Petroleum Status Report. Prior research attributes this increase to the existence of asymmetries in another dimension, such as information processing capabilities. While these asymmetries may be important, they cannot explain other empirical findings. Precisely, they predict that the reaction of volume and spreads depends on the signal realization. However, I find strong evidence that this reaction is unrelated to the content of the report.

I construct and test a model in which informed investors are risk-averse. This additional feature induces an increase on spread and volume, as we observe in the data. These results indicate that it is not enough to homogenize the reports' layout or the communication channel. Yet, adverse selection costs might increase. Instead, the model suggests that agencies should release the information when private information is less important, such as the end of the day, if they want to reduce he impact on spreads.

Besides its implications around a public announcement, risk aversion generates other empirical predictions that I leave for further research. For instance, in contrast to most previous literature, bid and ask prices in my model are not first-order Markovian. In particular, a sale after a buy -when prices are higher- is more informative than if it follows another sale. This prediction is the result of the endogenous participation of informed traders. When prices are high, the information rents to buy are small; thus, if a trader buys, she is likely to be a noisy trader.

This paper also presents some limitations. In the model, public and private signals are independent. This assumption is reasonable for announcements related to macroeconomic variables, or other global factors, which are unlikely to correlate with firm specific private information. In other contexts, in which both signals are highly dependent, the predictions of the model might reverse. Understanding the role of this dependence is left for further research.

References

S. Akhtar, R. Faff, B. Oliver, and A. Subrahmanyam. The power of bad: The negativity bias in australian consumer sentiment announcements on stock returns. *Journal of Banking & Finance*, 35(5):1239–1249, 2011.

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.

J. Begley and P. E. Fischer. Is there information in an earnings announcement delay? *Review of Accounting Studies*, 3(4):347–363, 1998.

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.

J. H. Boyd, J. Hu, and R. Jagannathan. The stock market's reaction to unemployment news: Why bad news is usually good for stocks. *The Journal of Finance*, 60(2):649–672, 2005.

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. Y. Campbell, T. Ramadorai, and A. Schwartz. Caught on tape: Institutional trading, stock returns, and earnings announcements. *Journal of Financial Economics*, 92(1): 66–91, 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.

S. A. Corwin and P. Schultz. A simple way to estimate bid-ask spreads from daily high and low prices. *The Journal of Finance*, 67(2):719–760, 2012.

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.

E. F. Fama. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417, 1970.

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.

D. A. Hirshleifer, J. N. Myers, L. A. Myers, and S. H. Teoh. Do individual investors cause post-earnings announcement drift? direct evidence from personal trades. *The Accounting Review*, 83(6):1521–1550, 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.

R. Kaniel, S. Liu, G. Saar, and S. Titman. Individual investor trading and return patterns around earnings announcements. *The Journal of Finance*, 67(2):639–680, 2012.

B. Ke and S. Ramalingegowda. Do institutional investors exploit the post-earnings announcement drift? *Journal of Accounting and Economics*, 39(1):25–53, 2005.

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

O. Kim and R. E. Verrecchia. Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics*, 17(1-2):41–67, 1994.

O. Kim and R. E. Verrecchia. Pre-announcement and event-period private information. *Journal of Accounting and Economics*, 24(3):395 – 419, 1997.

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.

W. Kross and D. A. Schroeder. An empirical investigation of the effect of quarterly earnings announcement timing on stock returns. *Journal of Accounting Research*, 22 (1):153–176, 1984.

A. Kurov, A. Sancetta, G. Strasser, and M. H. Wolfe. Price drift before us macroeconomic news: private information about public announcements? 2016.

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.

R. Michaely, A. Rubin, and A. Vedrashko. Further evidence on the strategic timing of earnings news: Joint analysis of weekdays and times of day. *Journal of Accounting and Economics*, 62(1):24 – 45, 2016.

J. R. Nofsinger. The impact of public information on investors. *Journal of Banking & Finance*, 25(7):1339–1366, 2001.

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.

S. K. Roache and M. Rossi. The effects of economic news on commodity prices. *The Quarterly Review of Economics and Finance*, 50(3):377–385, 2010.

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.

M. Vieru, J. Perttunen, and H. Schadewitz. How investors trade around interim earnings announcements. *Journal of Business Finance & Accounting*, 33(1-2):145–178, 2006.

A Proofs

Proof. Proposition 1

Since the market maker problem is symmetric, I focus 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. The zero profit condition is given by:

$$A_t^* = \mathbb{E}\left[v|d=1\right] = \mathbb{E}\left[\omega|d=1\right]$$

where the last equality comes from independence between ω, μ and ε. Using the distribution of ω, we can rewrite this condition as:

$$g(A_t^*) \equiv A_t^* + \sigma_\omega - P(\omega = \sigma_\omega|d=1, \mathcal{H}_t)2\sigma_\omega - \tilde{\mu} = 0 \tag{9}$$

where

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

It is trivial to show that $g'(A) \geq 0\ \forall A$, $g(\tilde{\mu} - \sigma_\omega) \leq 0$ and $g(\tilde{\mu} + \sigma_\omega) \geq 0$. Therefore, an equilibrium exists and it is unique. Moreover,

$$\tilde{\mu} - \sigma_\omega \leq A_t^* \leq \tilde{\mu} + \sigma_\omega \tag{10}$$

□

Proof. Corollary 1

Note that in equilibrium, under the assumption that $0 < \delta < 1$,

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

Therefore, the inequalities in Equation (10) are strictly satisfied. We can follow a similar argument, and show that these inequalities hold for the bid price.

Besides, to complete the proof, I need to show that $B_t^* < A_t^*$. Given Equation (9) and the analogue one for the bid side of the market, it is easy to show that:

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

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 $A^*(\mathcal{H}_t) = -B^*(-\mathcal{H}_t)$. 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 1.

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

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

$$lim_{T\to\infty}\, P(N_B > N_S|\omega = \sigma_\omega) = 1 \text{ and } lim_{T\to\infty}\, P(N_B > N_S|\omega = -\sigma_\omega) = 0. \tag{11}$$

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

Lastly, the results in Equation (11) 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:

$$F\left(\frac{\sigma_\omega + \mu - A^*_{1t}}{\sigma_\varepsilon}\right) > F\left(\frac{\sigma_\omega + \mu - A^*_{0t}}{\sigma_\varepsilon + \sigma_\mu}\right) \tag{12}$$

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:

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

It is trivial to show that

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

Since $g(\cdot)$ is increasing and $g(A^*_{1t_R}) = 0$ then, $\bar{A} > A^*_{1t_R}$. Therefore, (12) is satisfied. A similar argument follows for the bid side of the market to show that:

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

To prove that the increase does not depend on μ is enough to show that volume is not affected by this variable. From Equation (9), we know that $A^*_{1t} - \tilde{\mu}$ does not depend on μ. As a consequence, volume, defined as the right hand side of inequality (12), does not depend on μ.

The limiting result is trivial using Lemma 1 □

Proof. Proposition 3

With some algebra, it can be shown that:

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

therefore, given (12) , $A^*_{1t_R} - \mu > A^*_{0t_R}$. 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 1.

□

B Additional Figures

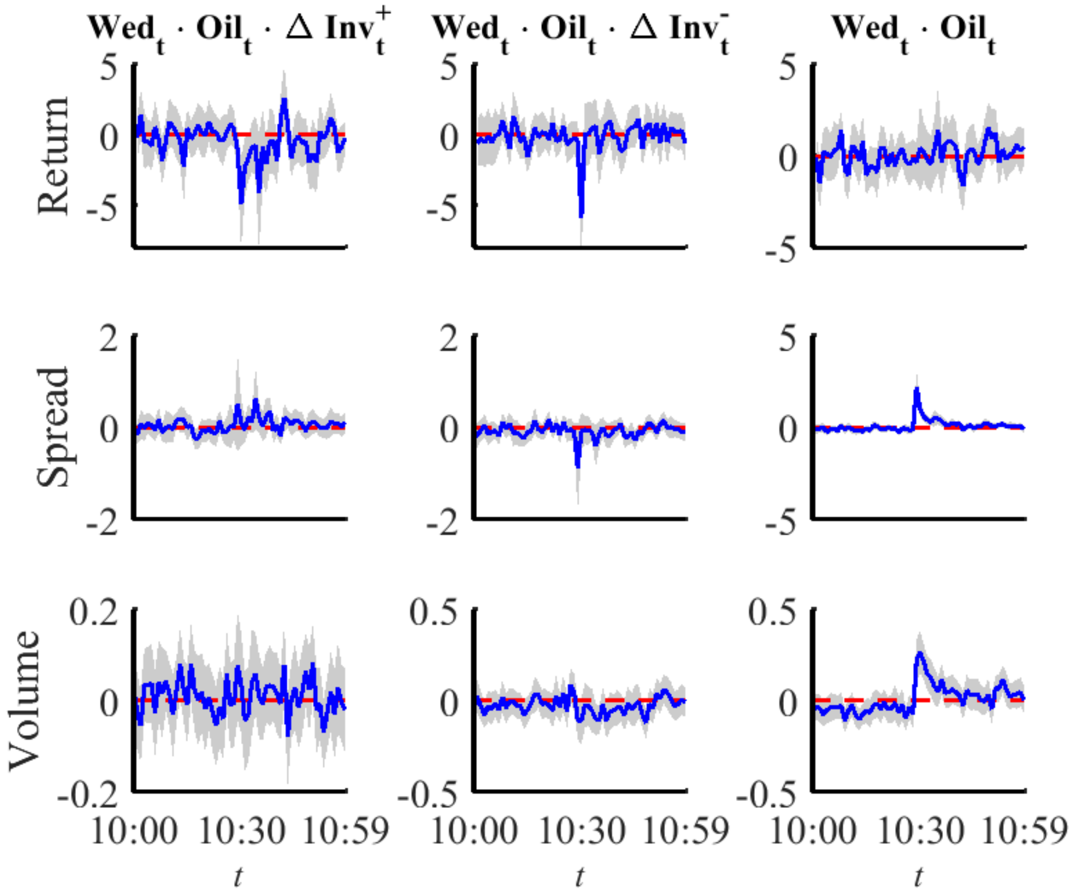

Figure B.1: Asymmetric Reaction. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (5) Each row consider a different dependent variable, from top to bottom: midpoint return, proportional effective bid-ask spread, and number of transactions. I estimate the coefficients independently for every minute exploiting variation across days. The left-hand side plots correspond to γ_2^+, and the middle ones to γ_2^-; meanwhile, the right-hand side ones refers to γ_1. I describe the sample in Section 2. The red dashed line indicates the zero.

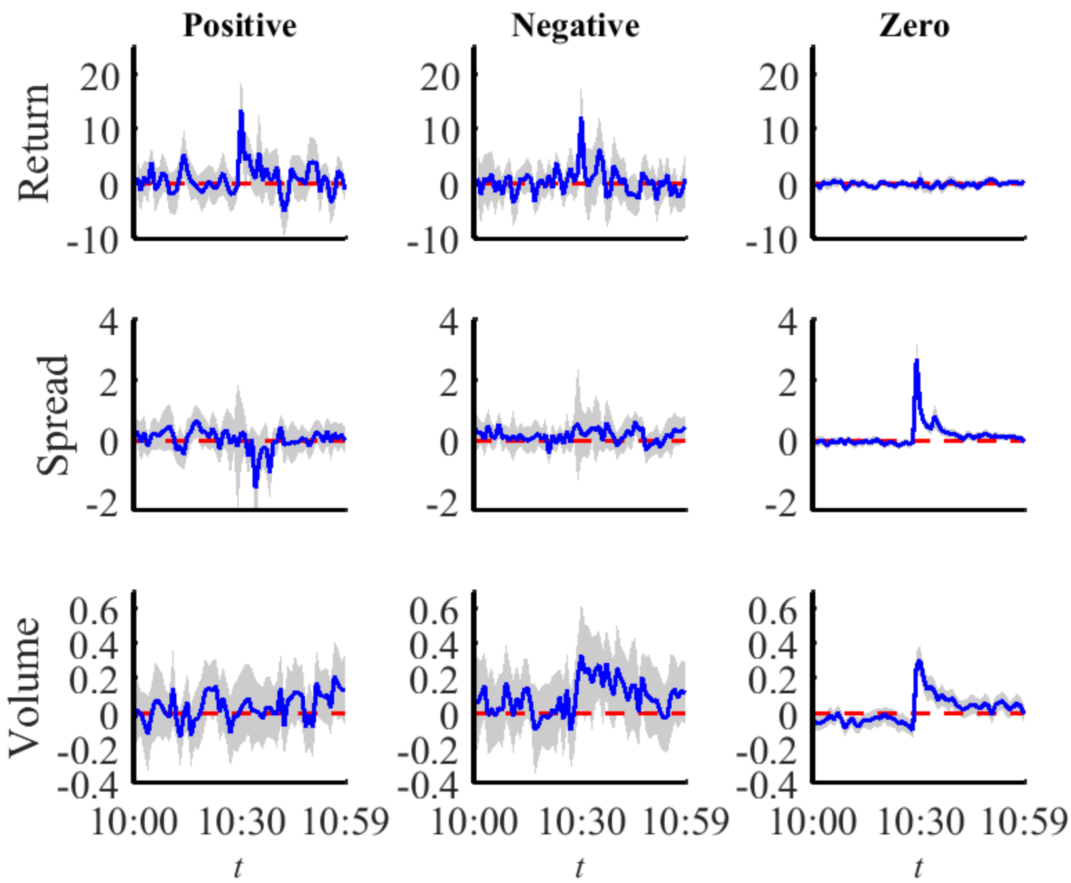

Figure B.2: Binary Specification. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (6) Each row consider a different dependent variable, from top to bottom: midpoint return, proportional effective bid-ask spread, and number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

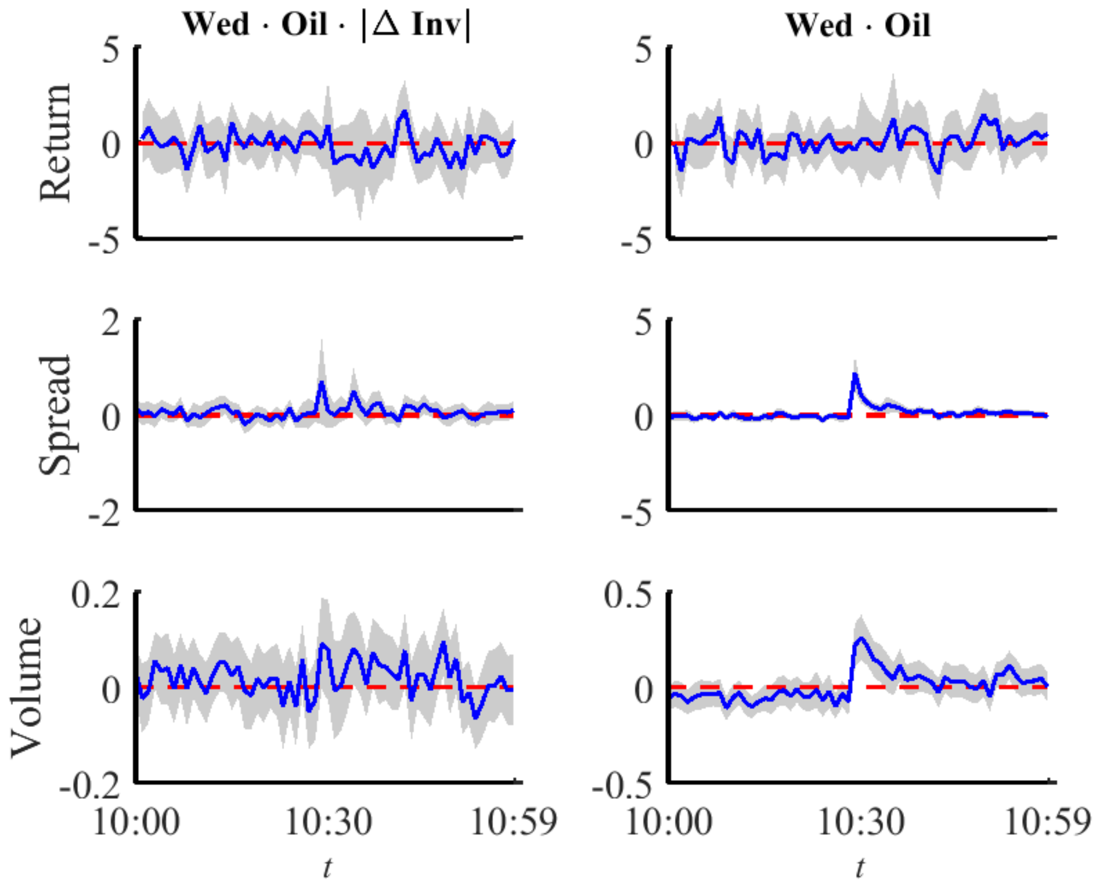

Figure B.3: More Noise Traders. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (7). Each row consider a different dependent variable, from top to bottom: midpoint return, proportional effective bid-ask spread, and number of transactions. I estimate the coefficients independently for every minute exploiting variation across days. The left-hand side plots correspond to γ_2, while the right-hand side ones refers to γ_1. I describe the sample in Section 2. The red dashed line indicates the zero.

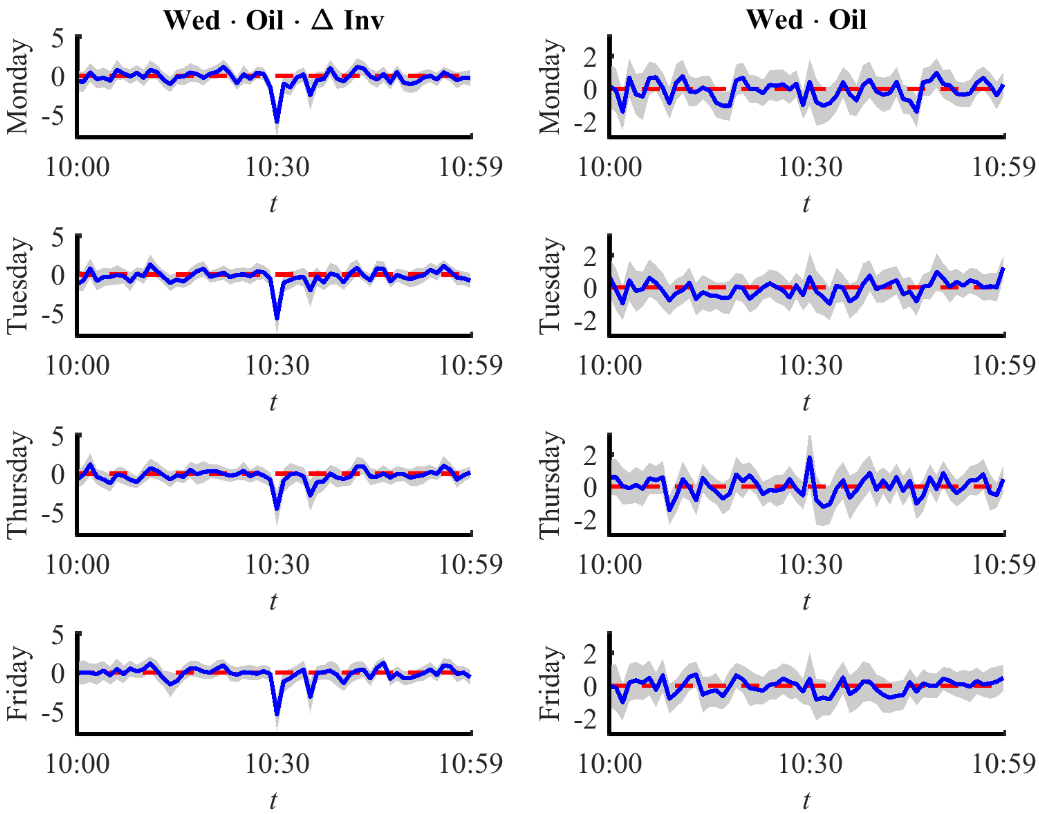

Figure B.4: Different weekday as a control (Returns). This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (7), using midpoint returns as a dependent variable. I restrict the sample to Wednesdays and another weekday as a control. This control day is Monday in the first row and Tuesday, Thursday and Friday the following rows. I estimate the coefficients independently for every minute exploiting variation across days. The left-hand side plots correspond to γ_2, while the right-hand side ones refers to γ_1. I describe the sample in Section 2. The red dashed line indicates the zero.

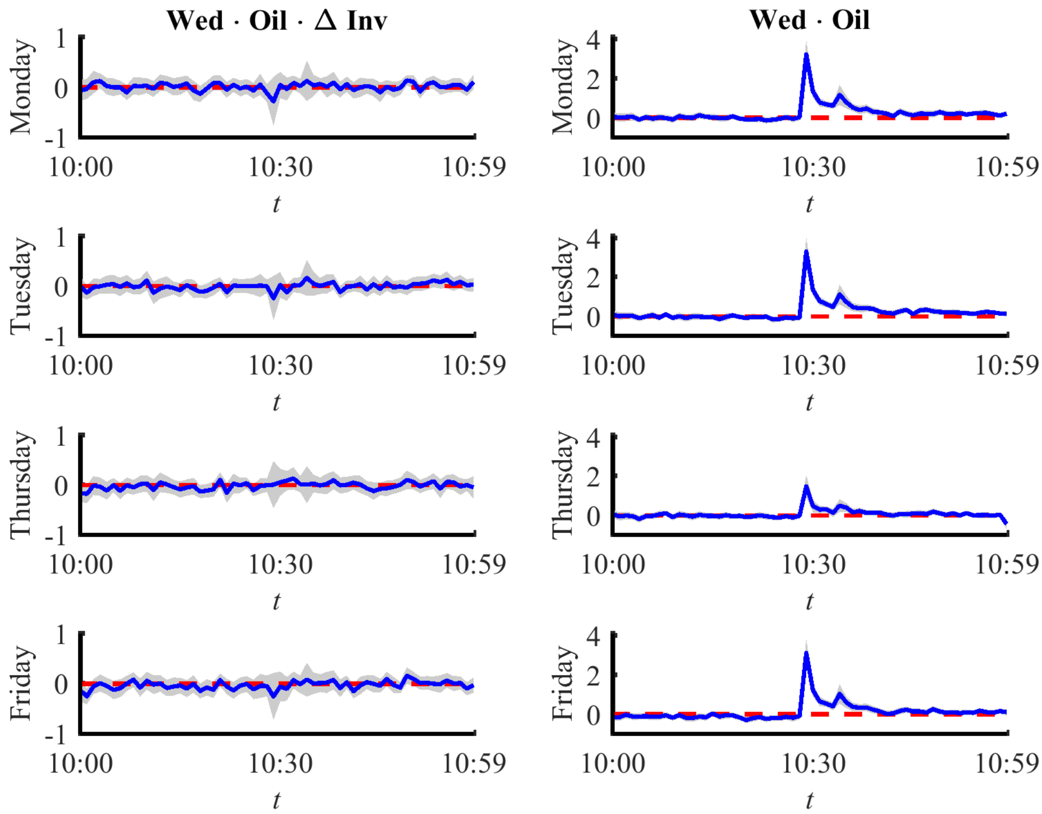

Figure B.5: Different weekday as a control (Spread). This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (7), proportional effective bid-ask spreads as a dependent variable. I restrict the sample to Wednesdays and another weekday as a control. This control day is Monday in the first row and Tuesday, Thursday and Friday the following rows. I estimate the coefficients independently for every minute exploiting variation across days. The left-hand side plots correspond to γ_2, while the right-hand side ones refers to γ_1. I describe the sample in Section 2. The red dashed line indicates the zero.

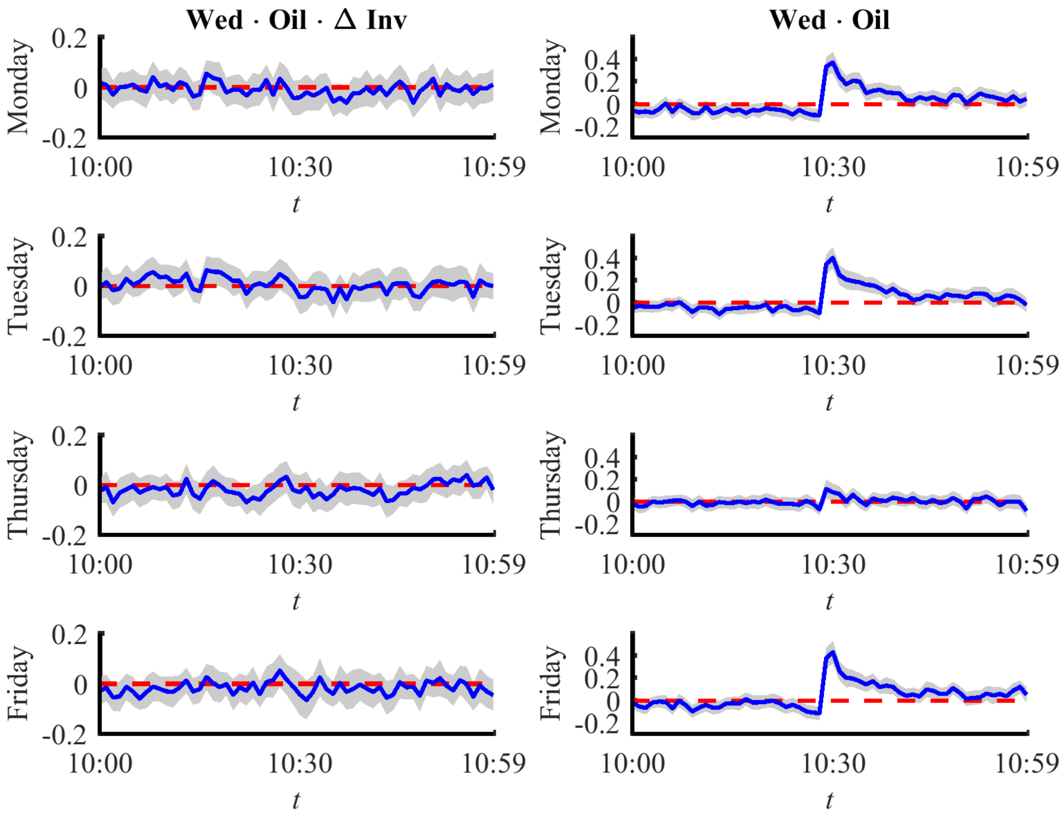

Figure B.6: Different weekday as a control (Volume). This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (7), using number of transactions as a dependent variable. I restrict the sample to Wednesdays and another weekday as a control. This control day is Monday in the first row and Tuesday, Thursday and Friday the following rows. I estimate the coefficients independently for every minute exploiting variation across days. The left-hand side plots correspond to γ_2, while the right-hand side ones refers to γ_1. I describe the sample in Section 2. The red dashed line indicates the zero.

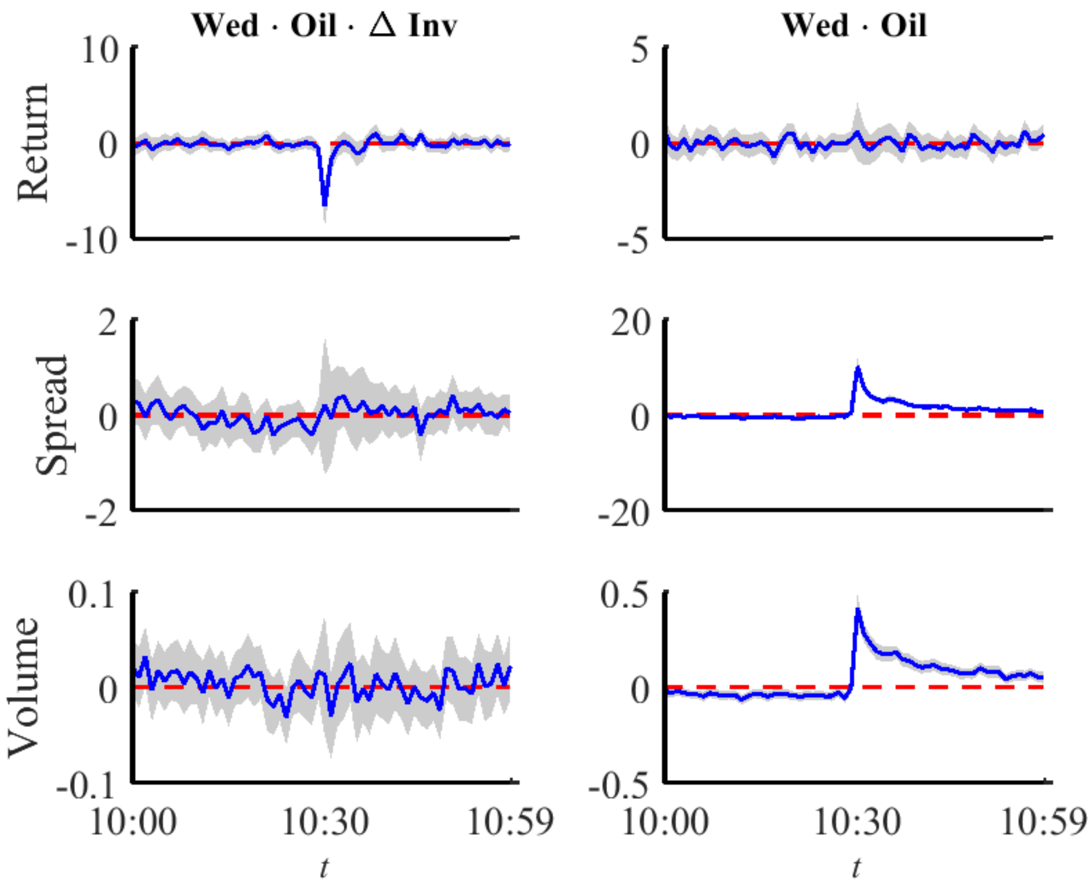

Figure B.7: Estimates Post-2005. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (4). Each row consider a different dependent variable, from top to bottom: midpoint return, high-low spread, and number of transactions. I estimate the coefficients independently for every minute exploiting variation across days. The left-hand side plots correspond to γ_2, while the right-hand side ones refers to γ_1. I describe the sample in Section 6.4 and I restrict it to observations after year 2004. The red dashed line indicates the zero.

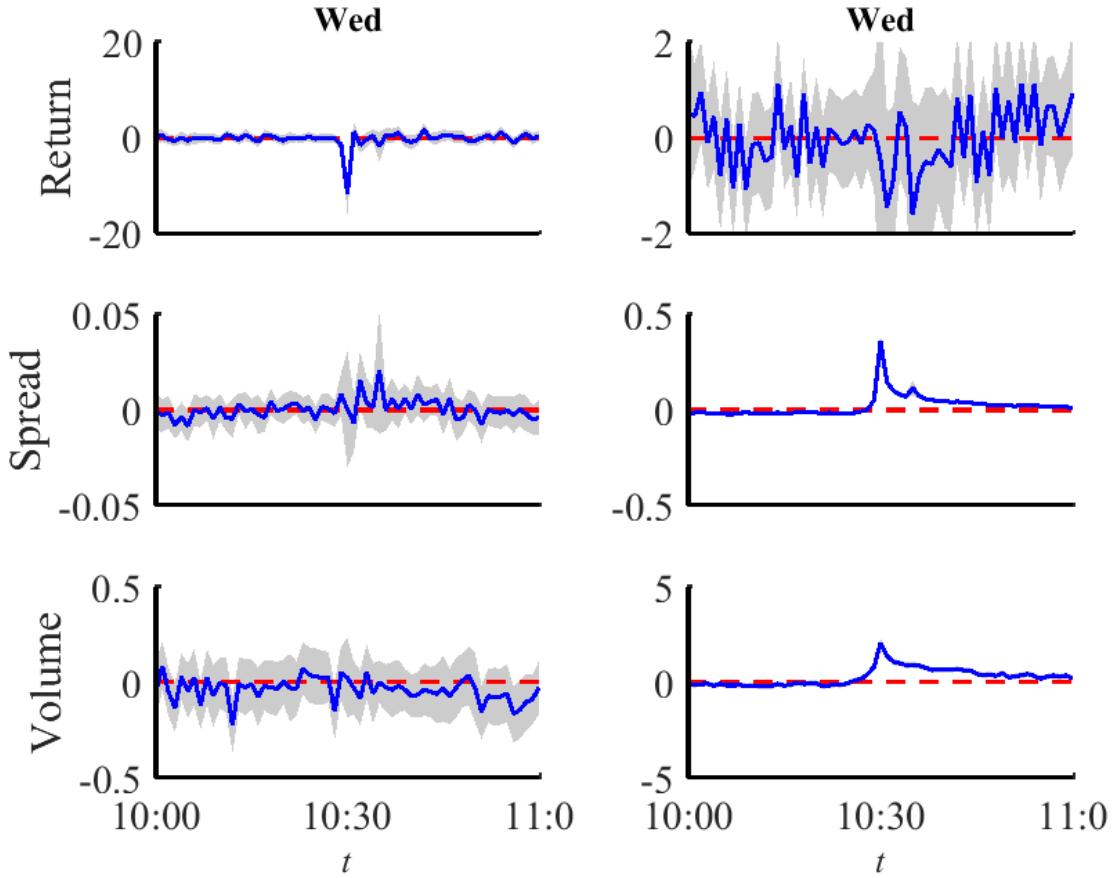

Figure B.8: Oil Market. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (4). Each row consider a different dependent variable, from top to bottom: midpoint return, high-low spread, and number of transactions. I estimate the coefficients independently for every minute exploiting variation across days. The left-hand side plots correspond to γ_2, while the right-hand side ones refers to γ_1. I describe the sample in Section 6.5. The red dashed line indicates the zero.

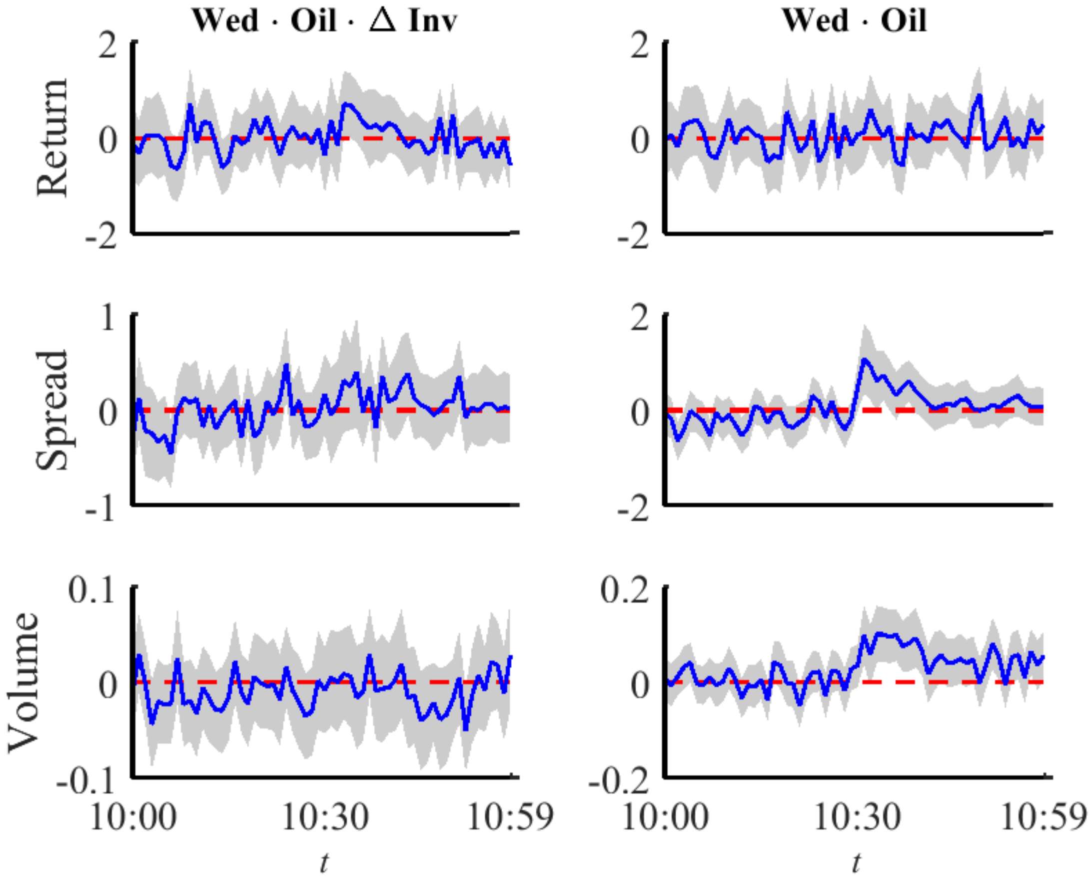

Figure B.9: Estimates Pre-2005. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (4). Each row consider a different dependent variable, from top to bottom: midpoint return, high-low spread, and number of transactions. I estimate the coefficients independently for every minute exploiting variation across days. The left-hand side plots correspond to γ_2, while the right-hand side ones refers to γ_1. I describe the sample in Section 6.4 and I restrict it to observations before year 2004. The red dashed line indicates the zero.

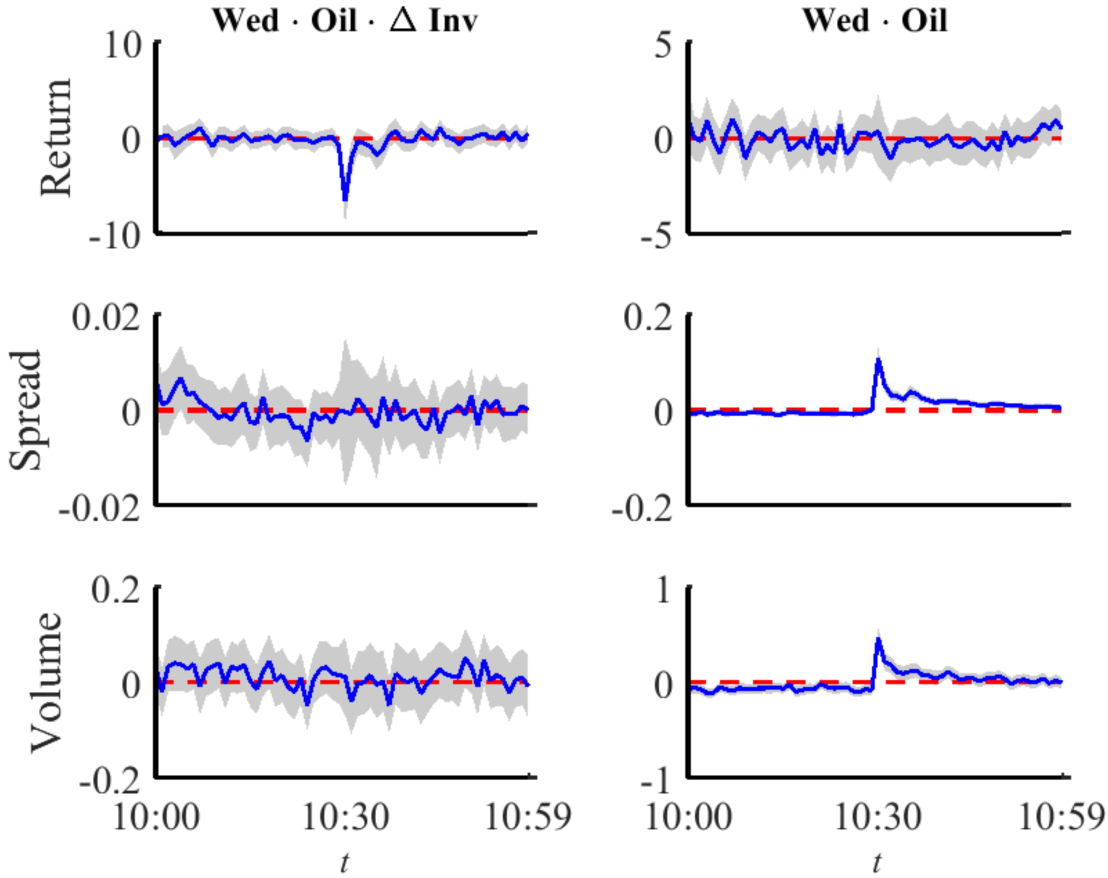

Figure B.10: Triple-diff.Estimates Post-2005. This figure plots the estimates (blue line) and their 95 % confidence interval (grey area) of the model described in equation (8). Each row consider a different dependent variable, from top to bottom: midpoint return, high-low spread, and number of transactions. I estimate the coefficients independently for every minute exploiting variation across days. The left-hand side plots correspond to γ_2, while the right-hand side ones refers to γ_1. I describe the sample in Section 6.4. The red dashed line indicates the zero.