The Pre-FOMC Announcement Drift and Information Leakage: Kyle Meets Macro-Finance*

Chao Ying†

May 8, 2020

Please click here for the latest version.

Abstract

This paper shows that the pre-FOMC announcement drift is largely caused by information leakage. The VIX index—a gauge for market uncertainty—decreases significantly in the 24 hours preceding FOMC announcements. The systematical reduction of market uncertainty before announcements comes from information leakage, which is associated with realizations of equity premium. To account for these dynamics, I extend Kyle’s (1985) model to the case where market makers are compensated for the risk of holding assets. Market makers update their beliefs about the pricing kernel as well as the asset value from observing cumulative trading volume, which resolves uncertainty gradually and results in an upward drift in market prices even the leaked news is on average neutral.

Keywords: Information Leakage; Pre-FOMC Announcement drift; Risk Compensation

*I am grateful to Hengjie Ai, Bob Goldstein, Xiaoji Lin, and Colin Ward for their advice. I thank Moritz Lenel and Vish Viswanathan for their helpful comments.
†University of Minnesota, Carlson School of Management, 321 19th Ave South Minneapolis, MN, 55455, USA. E-mail: yingx040@umn.edu
1 Introduction

Recent work by Lucca and Moench (2015) documents the substantial stock market return before the Federal Open Market Committee (FOMC) announcements without a major increase in conventional measures of risk. They find that the pre-FOMC drift of the S&P 500 index is on average 49 basis points during the 24-hour window preceding FOMC announcements, which corresponds to about 80% of the annual realized excess returns in the stock market. In the meantime, the return volatility and trading volume before announcements are similar to other days, which makes the pre-FOMC announcement drift more puzzling and provides a notable challenge to the existing asset pricing models.

In this paper, I show that the pre-FOMC announcement drift is a result of information leakage. Empirically, the CBOE VIX index decreases significantly and systemically during the 24-hour window before the announcements, which implies the reduction of market uncertainty preceding FOMC news. However, the FOMC members refrain from discussions of monetary policy information in the week before FOMC meetings, and, more importantly, in the 24-hour pre-FOMC window over which is associated with substantial returns. Therefore, the systematic reduction of aggregate market uncertainty has to come from the information leakage, where the trading of insiders reveals information to the market gradually before announcements.\(^1\)

To support the information leakage channel, I provide more empirical evidence through two different classifications of FOMC announcements. First, by sorting the FOMC days via their reduction of uncertainty during the 24-hour window before announcements into three groups, I find different patterns of pre-FOMC returns. The high-reduction group has a sharper drift than the average FOMC results, which is associated with an average return of 85 basis points during this window. By contrast, for the low-reduction group, the uncertainty is increased instead of being reduced before announcements, which corresponds to an average -20 basis points pre-FOMC return. Second, the introduction of press conferences on half of the announcements

---

\(^1\)Hu, Pan, Wang, and Zhu (2020) document the reduction of uncertainty before announcements through the VIX index. However, in contrast to leaks, they assume all investors observe it and interpret the source of uncertainty as the variance of returns, which is orthogonal to asset fundamentals.
since April 2011 leads to two classes of FOMC announcements: with press conferences and without press conferences.\(^2\) For the group with press conferences, the uncertainty reduces significantly before announcements with an average 15 basis points pre-FOMC return. However, for the group without press conferences, the reduction of uncertainty, and the return before announcements are not significantly different from zero. Both classifications imply that not all FOMC days are the same, and the pre-announcement drift is associated with the reduction of uncertainty before announcements.

To quantitatively account for the uncertainty reduction and the pre-FOMC returns, I extend Kyle’s (1985) model to the case where market makers require a risk compensation for news related to fundamentals. To introduce the macroeconomic conditions into market makers’ pricing decisions, I develop a continuous-time equilibrium model in which the aggregate economic growth is driven by a latent state variable and an i.i.d. component (short-run shocks). The investors cannot observe the latent variable directly. They update their beliefs from observing the aggregate output. The fed has private information about the true value of the latent growth variable, which should be revealed through periodic FOMC announcements. However, some insiders know private information one day before announcements. They start to trade to maximize the conditional expected gains, understanding the order affects the price.

The competitive market makers set the price, which equals the marginal utility weighted payoffs through the stochastic discount factor (SDF) determined from the above economy. Since the SDF applies extra discounting to payoffs positively correlated with utility, the asset market requires a premium for such payoffs relative to risk-free returns. In other words, market makers are compensated for the risk of holding assets because they may see a decline in the value of a security after it has been purchased from a seller and before it is sold to a buyer. Therefore, based on the observed cumulative order flow, market markers update the estimation of asset payoffs as well as the SDF simultaneously. The uncertainty of underlying fundamentals is resolved gradually during this process, which is associated with the realizations of equity premium before announcements.

\(^2\)From January 2019, the Chairman of the Federal Reserve holds a press conference after each meeting.
Since the information from trading resolves the market makers’ uncertainty, the price dynamic with respect to market makers’ perspective is a submartingale instead of a martingale in the standard continuous time Kyle model.\(^3\) The submartingale has a deterministic growth rate, which is the negative covariance between SDF and the asset value, i.e., the announcement premium per unit of time. Given the realization of equity premium, market makers would expect the insiders have a positive order rate instead of 0. The expected order rate increases in the concavity of the SDF, the prior variance of the growth rate upon announcements, and the volatility of uninformed trading volume. This implies market makers expect that insiders choose to trade more aggressively when (1) market makers require a higher risk compensation for uncertainty; (2) the underlying fundamental is noisier; (3) uninformed volume is higher, and price impact is lower. The expected order rate converges to zero when market makers are risk neutral, as in Back (1992).

Furthermore, given uncertainty is not always reduced before some FOMC meetings, I generalize the model that insiders may not be better informed, and market makers have to update their estimate of the probability that insiders have private information. Also, market makers update their beliefs of asset payoffs and the SDF, which are solved by a nonlinear filtering technique. The model yields a nonlinear pricing rule, which drives price volatility, market depth, and price response to be stochastic. More importantly, the price dynamic is a submartingale with a positive and time-varying growth rate. The expected order rate is positive, which converges to zero when the market makers are risk neutral and converge to the benchmark when the market makers always think insiders are better informed. If the market makers’ prior is that insiders are always not better informed, there is no pre-announcement drift since no uncertainty is resolved.

Related literature

The paper relates to several strands of the literature. First, it relates to the literature that studies the impact of asymmetric information on asset prices and price impacts such as Kyle (1985),

\(^3\)See Back (1992) and Collin-Dufresne and Fos (2014).
Back (1992), and Li (2013); however, there are two important differences. First, the above papers consider risk-neutral market makers who are not compensated for the risk of holding assets. The price dynamic in these papers is a martingale, which implies there is no pre-announcement drift when the leaked news is on average neutral. However, in the meantime, many empirical works of the literature treat the positive abnormal return in the pre-announcement period as evidence of information leakage, such as Collin-Dufresne and Fos (2015). My paper extends the literature to the case where market makers require a risk compensation for news related to fundamentals. It not only explains the pre-FOMC announcement drift, but also provides a theoretical foundation to other empirical works which conclude information leakage results in the pre-events positive abnormal returns. Second, the previous theoretical work finds the expected order rate from insiders is zero from market makers’ perspective. However, in my paper, the market makers expect a positive trading volume from insiders due to the realizations of a substantial amount of pre-announcement premium. They think the insiders would trade in a way to chase the positive excess return.

My paper is related to the broader literature on macro-finance, which addresses the link between asset prices and economic fluctuations. Many works of literature highlight the importance of macroeconomic conditions, such as Chen (2010) and Kuehn and Schmid (2014) on credit spread, Chernov, Schmid, and Schneider (2016) and Tourre (2017) on international finance, and Dvaid, Schmid, and Zeke (2018) on capital misallocation. However, how asymmetric information affects the asset prices is largely ignored in this literature caused by the technical challenge. This paper fills this gap between the macro-finance literature and the asymmetric information literature. There is a pre-announcement drift if and only if market makers require a risk compensation for news related to fundamentals.


5 While Subrahmanyam (1991) considers risk-averse market makers under CARA utility in a one-period Kyle model, there is no pre-announcement drift since the fundamental risk in the news is not priced due to CARA utility.

6 See reviews in Cochrane (2017).

7 The information channel I emphasize is consistent with recent work by Nakamura and Steinsson (2018). They provide empirical evidence and develop a theoretical model to show that Fed announcements affect beliefs. 
My paper contributes to the macroeconomic announcements literature both empirically and theoretically. Lucca and Moench (2015) document the substantial stock market return during the 24-hour period preceding FOMC announcements. Hu, Pan, Wang, and Zhu (2019) provide evidence that the option-implied variance increases before announcements and drops afterward and attribute the FOMC announcement premium to heightened stock market uncertainty. Vissing-Jorgensen (2019) provides a history of leak discussions in FOMC documents for the period 1948-2013 in order to show that the FOMC itself expresses frequent concerns about leaks. The two different classifications of FOMC meetings in this paper support the information leakage channel through asset-market-evidence. Theoretically, Ai and Bansal (2018) demonstrate that, the substantial equity market returns realized on FOMC announcement days imply that preferences must satisfy generalized risk sensitivity in a representative agent economy. Their quantitative framework provides the macroeconomic conditions for this paper. I find the generalized risk sensitivity is necessary to account for the pre-announcement drift in the presence of asymmetric information.

The rest of the paper is organized as follows. I provide empirical evidence to support that the pre-announcement drift is a result of information leakage in Section 2. In Section 3, I present the macroeconomic environment that the investors can not observe the growth rate of output directly. Section 4 describes the insider trading before announcements, where I extend Kyle’s (1985) model of insider trading to the case where market makers require a risk compensation for news related to fundamentals. In section 5, I generalize the model that insiders may not be better informed, and market makers have to update their estimate of the probability that insiders have private information. Section 6 concludes.

---

8In addition, Bernile, Hu, and Tang (2016) and Kurov, Sancetta, Strasser, and Wolfe (2017) find evidence consistent with informed trading about 30 minutes before scheduled announcements. Cieslak, Morse, and Vissing-Jorgensen (2019) suggest that information leaks by Fed drive a biweekly pattern of high return rates. 

9Ai, Bansal, Im, and Ying (2018) and Wachter and Zhu (2018) develop quantitative models of the announcement premium with a representative agent, based on the generalized risk sensitivity in Ai and Bansal (2018). Ying (2020) shows that the needs of rebalancing out of less disagreement contributes to the huge trading volume upon announcements.
2 Empirical Evidence

In this section, I provide evidence that there is information leakage before FOMC news. The time-series pattern of information leakage during the 24-hour period preceding FOMC announcements is consistent with the pre-FOMC drift documented by Lucaa and Moench (2015). In addition, different classifications of FOMC announcements support that the information leakage mainly contributes to the pre-FOMC announcement drift.

2.1 The average cumulative VIX change and return before FOMC announcements

Figure 1: The average cumulative VIX change and return before FOMC announcements

This figure shows the average cumulative VIX Change and average cumulative return on the S&P 500 index on three-day windows from 1996 to 2019. The solid black line of the left (right) panel is the average cumulative VIX Change (average cumulative return on the SPX) from 9:30 AM ET on days prior to scheduled FOMC announcements to 4:00 pm ET on days after scheduled FOMC announcements. The gray shaded areas are pointwise 95% confidence bands around the average. The sample period is from January 1996 to December 2019. The dashed vertical line is set at 2:00 PM ET, when FOMC announcements are typically just released or 15 minutes before the release.

To capture changes in market expectations in a timely manner, I use the CBOE VIX index, which is a model-free measure of implied volatility computed from the S&P 500 index option prices. For the intraday returns, I obtain transaction-level data on S&P 500 index. The sample
period is from January 1996 to December 2019. During this period, there are in total 187 scheduled releases of FOMC statements. Except 9 of them, other releases are either around 2:15 PM ET (before April 2011) or 2:00 PM ET (after April 2011). Therefore, I follow Lucca and Moench (2015) and focus on the 2 PM-to-2 PM pre-FOMC window, which should not contain any announcement information if there is no information leakage.

Figure 1 shows the average cumulative VIX Change and average cumulative return on the S&P 500 index around FOMC announcements. The black solid line of the right panel represents the mean pointwise cumulative intraday percentage return of the SPX over a three-day window from the market open of the day ahead of scheduled FOMC meetings to the day after. As shown in Table 2, the pre-announcement drift for FOMC is on average 33.2 basis points, which is statistically significant at the 1% level. The left panel indicates the VIX decreases 0.134% with a t-stat of -2.7 during the 24-hour period preceding FOMC announcements, which is consistent with Hu, Pan, Wang, and Zhu (2019). The significant reduction of VIX index before announcements indicates the presence of information leakage, which resolves the market uncertainty systematically.

2.2 Classifications of FOMC meetings

To support the information leakage channel, I provide more empirical evidence through two different classifications of FOMC announcements. First, sorting the FOMC days by their reduction of uncertainty during the 24-hour window before announcements into three groups, I find different patterns of pre-FOMC returns. As shown in Table 2, the high-reduction group has a sharper drift than the average FOMC results, which is associated with an average return of 92.5 basis points during this window. By contrast, for the low-reduction group, the uncertainty is increased instead of reduced before announcements, which corresponds to an average -7.6 basis points pre-FOMC return.

8 of the 9 exceptions are released around 12:30 PM ET from April 2011 to December 2012. Another exception happened at 11:30 AM ET on March 26, 1996 because of the Chairman’s other duties. The results hold robustly without these releases.

The pre-announcement reduction of VIX accounts for around 60% of the total reduction around FOMC meetings.
Figure 2: Classifications of FOMC meetings: sort on the reduction of uncertainty

This figure shows the average cumulative return on the S&P 500 index on three-day windows, which is sorted on the reduction of uncertainty before announcements from 1996 to 2019.

Second, the introduction of press conferences on half of the announcements since April 2011 leads to two classes of FOMC announcements: with press conferences and without press conferences.\textsuperscript{12} For the group with press conferences, the uncertainty reduces significantly before announcements with an average 15 basis points pre-FOMC return. However, for the group without press conferences, the reduction of uncertainty and the return before announcements are not significantly different from zero. Both classifications imply that not all FOMC days

\textsuperscript{12}From January 2019, the Chairman of the Federal Reserve holds a press conference after each meeting.
are the same, and the pre-announcement drift is associated with the reduction of uncertainty before announcements.

Figure 3: Classifications of FOMC meetings: press conferences

This figure shows the average cumulative return on the S&P 500 index on three-day windows with and without press conferences from April 2011 to December 2019.

Table 3 shows results for regressing the changes in VIX ($\Delta VIX$) on the cumulative excess returns on the S&P500 ($Cum. Return$),

$$Cum. \ Return_t = \alpha + \beta \Delta VIX_t + \varepsilon_t,$$

where both $\Delta VIX_t$ and $Cum.\ Return_t$ are calculated from 2 p.m on pre-announcement date
to 2 p.m on announcement date windows, and $t$ represents each FOMC announcement. As shown in Table 3, on average when VIX decreases 1 basis points on before FOMC news, the cumulative return increases 0.911 basis points. Table 4 indicates that the change of VIX before announcements itself can explain a large fraction of the pre-announcement drift.

3 The Macroeconomic Conditions

In this section, I introduce the macroeconomic environment, including technology and information process, which determines how aggregate risks and risk prices change, especially upon FOMC announcements. Later on in section 4, I add the key elements in the microstructure literature, including the inside trader, liquidity traders and market makers.

3.1 Physical setup of the model

There are a large number of identical infinitely lived households in the economy. I assume that the consumption of the representative agent, $C_t$, follows

$$\frac{dC_t}{C_t} = x_t dt + \sigma dB_{C,t}, \quad (1)$$

where $x_t$ is a continuous-time AR(1) process (an Ornstein-Uhlenbeck process) unobservable to the agent in the economy. The law of motion of $x_t$ is

$$dx_t = a_x (\bar{x} - x_t) dt + \sigma_x dB_{x,t}. \quad (2)$$

The standard Brownian motions $B_t$ and $B_{x,t}$ in equations (1) and (2), respectively, are independent.

At time 0, the agent’s prior belief about $x_0$ can be represented by a normal distribution with mean $m_0$ and variance $\zeta_0$. Although $x_t$ is not directly observable, the agent can use two sources of information to update beliefs about $x_t$. First, the realized consumption path contains information about $x_t$, and second, at pre-scheduled discrete time points $T, 2T, 3T, \cdots$, additional signals about $x_t$ are revealed through announcements. For $n = 1, 2, 3, \cdots$, I denote
$s_n$ as the signal observed at time $nT$ and assume $s_n = x_{nT} + \varepsilon_n$, where $\varepsilon_n$ is i.i.d. over time, and normally distributed with mean zero and variance $\sigma^2_S$.

Given the information structure, the posterior distribution of $x_t$ is Gaussian and can be summarized by its first two moments. I define $\hat{x}_t = E_t[x_t]$ as the posterior mean and $q_t = E_t [(x_t - \hat{x}_t)^2]$ as the posterior variance, respectively, of $x_t$ given information up to time $t$. For $n = 1, 2, \cdots$, at time $t = nT$, the agent updates his beliefs using Bayes’ rule:

$$\hat{x}_{nT}^+ = q_{nT}^+ \left[ \frac{1}{\sigma^2_S} s_n + \frac{1}{q_{nT}^-} \hat{x}_{nT}^- \right]; \quad \frac{1}{q_{nT}^+} = \frac{1}{\sigma^2_S} + \frac{1}{q_{nT}^-},$$

where $\hat{x}_{nT}^+$ and $q_{nT}^+$ are the posterior mean and variance after announcements, and $\hat{x}_{nT}^-$ and $q_{nT}^-$ are the posterior mean and variance before announcements, respectively. A special case is that the announcements can completely reveal the information about $x_t$, which means, $\sigma^2_S = 0$.

In the interior of $(nT, (n+1)T)$, the agent updates his beliefs based on the observed consumption process using the Kalman-Bucy filter:

$$d\hat{x}_t = a_x \left[ \bar{x} - \hat{x}_t \right] dt + \frac{q(t)}{\sigma} d\tilde{B}_{C,t},$$

where the innovation process, $\tilde{B}_{C,t}$ is defined by $d\tilde{B}_{C,t} = \frac{1}{\sigma} \left[ \frac{dC_t}{C_t} - \hat{x}_t dt \right]$. The posterior variance, $q(t)$ satisfies the Riccati equation:

$$dq(t) = \left[ \sigma^2_x - 2a_x q(t) - \frac{1}{\sigma^2} q^2(t) \right] dt.$$  

3.2 Preferences and the SDF

I assume that the representative agent is endowed with a Kreps-Porteus preference with risk aversion $\gamma$ and intertemporal elasticity of substitution $\psi$. In continuous time, the preference is represented by a stochastic differential utility, which can be specified by a pair of aggregators $(f, A)$ such that in the interior of $(nT, (n+1)T)$,

$$dV_t = [-f(C_t, V_t) - \frac{1}{2} A(V_t) ||\sigma_V(t)||^2] dt + \sigma_V(t) dB_t$$
I adopt the convenient normalization $A(v) = 0$ and denote $\tilde{f}$ the normalized aggregator. Under this normalization, $\tilde{f}(C,V)$ is:

$$
\tilde{f}(C,V) = \frac{\beta}{1 - 1/\psi} \frac{C^{1-1/\psi} - ((1 - \gamma) V)^{1-1/\psi}}{((1 - \gamma) V)^{1-1/\psi} - 1}.
$$

(7)

The case of $\psi = 1$ is obtained as the limit of (7) with $\psi \to 1$:

$$
\tilde{f}(C,V) = \beta V \left[ (1 - \gamma) \ln C - \ln [(1 - \gamma) V] \right].
$$

Because announcements typically result in discrete jumps in the posterior belief about $x_t$, the value function is typically not continuous at announcements. Given our normalization of the utility function, for $t = nT$, the pre-announcement utility and post-announcement utility are related by:

$$
V_t^{-} = E_t^{-} \left[ V_t^{+} \right],
$$

where $E_t^{-}$ represents expectation with respect to the pre-announcement information at time $t$. In what follows, I assume $\gamma > \frac{1}{\psi}$ so that the above preference satisfies generalized risk sensitivity in Ai and Bansal (2018).

In the above setup, I can show that the value function of the representative agent takes the form

$$
V (\hat{x}, t, C_t) = \frac{1}{1 - \gamma} H (\hat{x}, t) C_t^{1-\gamma},
$$

for some twice continuously differentiable function $H (\hat{x}, t)$. The HJB equation and the corresponding boundary conditions for $H (\hat{x}, t)$ can be found in Appendix. Given the utility of the representative agent, the state price density, denoted $\{\pi_t\}_{t=0}^{\infty}$ can be characterized by the following theorem.

**Lemma 1.** For $n = 1, 2, 3 \cdots$, in the interior of $((n-1)T, nT)$, $\pi_t$ is a continuous diffusion process with the law of motion

$$
\frac{d\pi_t}{\pi_t} = -r (\hat{x}, t) dt - \sigma_\pi (\hat{x}, t) d\tilde{B}_t,
$$
where \( r(\hat{x}, t) \) is the instantaneous risk-free interest rate and \( \sigma_\pi(\hat{x}, t) \) is the market price of risk. At announcements, \( t = nT \), \( \pi_t \) is discontinuous, and the A-SDF under IES=1 is given by

\[
m^*_t = \frac{\pi^+_t}{\pi^-_t} = \frac{H(\hat{x}^+_t, t^+)}{E^-_t[H(\hat{x}^+_t, t^+)]}.
\]

with

\[
H(\hat{x}^+_t, t^+) \begin{cases} 
\approx e^{\frac{1}{\alpha_x + \beta} \hat{x}^+_t + \mathcal{H}(t^+)} & \text{when } \psi \neq 1 \\
= e^{\frac{1}{\alpha_x + \beta} \hat{x}^+_t + \mathcal{H}(t^+)} & \text{when } \psi = 1
\end{cases}
\]

where \( h = \frac{\gamma - \frac{1}{\psi}}{\alpha_x + \beta} \), which is positive if and only if \( \gamma > \frac{1}{\psi} \).  

Given the above A-SDF, any asset with value \( A(\hat{x}^+_t, t^+) \) upon announcements \( t = nT \) will be valued by weighting the payoffs through investors’ future marginal utility and taking expectations:

\[
P(\hat{x}^-_t, t^-) = E\left[\frac{H(\hat{x}^+_t, t^+)}{E^+_t[H(\hat{x}^+_t, t^+)]} A(\hat{x}^+_t, t^+) | \hat{x}^-_t, t^-\right], \quad t = nT.
\]

When there is no information leakage, the information of the FOMC meetings is only revealed upon announcements, which results in non-trivial reductions of uncertainty, and are associated with realizations of a substantial amount of equity premium after announcements as in Ai and Bansal (2018) and Ai, Bansal, Im and Ying (2019). To capture the pre-FOMC announcement drift, I introduce the information leakage and insider trading into this framework.

---

13 For asset prices, I don’t need to solve the functional form of \( \mathcal{H}(t^+) \) since it is deterministic.

14 Ai and Bansal (2018) assumes that all agents in the financial market know the information before FOMC announcements at a specified rate exogenously. If this is true, we would also see the pre-FOMC announcement drift in other financial markets due to the complete market assumption in Ai and Bansal (2018). However, Lucca and Moench (2015) and Brooks, Katz and Lustig (2018) find no such effect in U.S. Treasury securities and money market futures. Besides, this is also not consistent with the empirical facts upon FOMC announcements. For example, there are monetary policy surprises (Nakamura and Steinsson (2018)) and huge trading volume as well as realized volatility (Lucca and Moench (2015), Bollerslev, Li, and Xue (2018) and Ying (2020)) after announcements.
4 Insider information

I extend Kyle’s (1985) model (in the continuous-time formulation given by Back (1992)) to allow that market makers are compensated for the risk of holding assets, where the market makers estimate the value of the risky asset as well as the A-SDF simultaneously before announcements.

4.1 Model setting

Inside traders in the stock market observe the signal of announcements $s_n = x_{nT} + \varepsilon_n$ before announcements, which implies they know the underlying value of the $A(\hat{x}_{nT}, nT)$ earlier than other investors in the market. In addition to the insiders, there are liquidity traders who have random, price-inelastic demands. All orders are market orders and are observed by all market makers. Denote by $Z_t$ the cumulative orders of liquidity traders through time $t$. The process $Z$ is assumed to be a Brownian motion independent of $\varepsilon_n$, which has mean zero and variance $\sigma_z^2$ (per unit of time). Let $X_t$ denote the cumulative orders of the informed trader and set $Y = X + Z$.

Without loss of generality, the insider traders know the information one day before announcement, i.e. at $t = nT - 1$. I assume the consumption is constant on the day before announcement.\footnote{Ai and Bansal (2019) have the same assumption to characterize their main result.} The variations in aggregate consumption in one day is unlikely to account for the huge announcement premium. It implies there is no more information which can reveal the latent growth variable $\theta_t$ until the signal upon announcements.

At the beginning of $nT - 1$, the posterior of distribution of $x$ for the market makers is summarized by $\{\hat{x}_{nT-1}, Q_{nT-1}\}$. Motivated by the macroeconomic conditions defined in last section, the market makers have the following A-SDF at time $nT - 1$:

$$m^{\star}_{nT-1, nT} = \frac{\pi_{nT}}{\pi_{nT-1}} = \frac{H(\hat{x}_{nT}, nT)}{E_{nT-1}[H(\hat{x}_{nT}, nT)]},$$

and later on they may update the estimate of the A-SDF based on the observed cumulative order flow before announcements.
The market makers, who are competitive, set the price at time \( t \in [nT - 1, nT] \) as

\[
P_t = E \left[ \frac{H(\hat{x}_{nT}, nT)}{E[H(\hat{x}_{nT}, nT)] | \mathcal{F}_t^Y} A(\hat{x}_{nT}, nT) | \mathcal{F}_t^Y \right].
\] (9)

Since the inside trader knows the true value of \( s_n = x_{nT} + \varepsilon_n \) at time \( nT - 1 \), there is no uncertainty of the underlying fundamental to him before announcement since he knows \( \hat{x}_{nT} \). This implies the A-SDF of the inside trader is 1.\(^{16}\) Given the perfect knowledge of the terminal value \( A(\hat{x}_{nT}, nT) \), the insider maximizes the expectation of his terminal profit:

\[
J(nT - 1, P_{nT-1}, A(\hat{x}_{nT}, nT)) = \max_{\theta_t \in \mathcal{A}} E \left[ \int_{nT-1}^{nT} (A(\hat{x}_{nT}, nT) - P_t) \theta_t dt | \mathcal{F}_{nT-1}^Y, A(\hat{x}_{nT}, nT) \right]
\] (10)

where I denote by \( \mathcal{F}_t^Y \) the information filtration generated by observing the entire past history of aggregate order flow \( Y \) (which I denote by \( Y^t = \{Y_s\}_{s \leq t} \)) and the posterior distribution of \( x_{nT-1} \sim N(\hat{x}_{nT-1}, Q_{nT-1}) \) at time \( nT - 1 \). In addition, the insider knows the actual value of the stock upon announcements, and, of course, his own trading. Following Back (1992), I assume that the insider chooses an absolutely continuous trading rule \( dX_t = \theta_t dt \) that belongs to an admissible set \( \mathcal{A} = \{ \theta \text{ s.t. } E \left[ \int_{nT-1}^{nT} \theta_s^2 ds \right] < \infty \} \). Therefore, the dynamics of aggregate order flow \( Y \) is

\[
dY_t = \theta_t dt + dZ_t.
\]

### 4.2 The equilibrium

An equilibrium is a price process and an admissible trading strategy, \((P_t, \theta_t)\), that satisfy the market maker’s rationality condition (9) while solving the insider’s optimality condition (10).

There are two main differences comparing to the standard microstructure literature (such as, Kyle (1985) and Back (1992)). The first difference is that, for the market makers, instead of

\(^{16}\)This can be shown directly through equation (8) where I take the expectation under the inside trader’s information at time \( nT - 1 \). Equivalently, I can assume the inside trader is risk neutral around announcements.
only estimating the value of the risky asset, they also update the A-SDF simultaneously before announcements based on the observed cumulative order flow. I will show later this difference is crucial, which implies the price \( P_t \) is a submartingale instead of a martingale given the market makers’ information. Motivated by this property that the market cares about the discounted present value, i.e. the produce of A-SDF and the value of the risky asset upon announcements jointly, I define

\[
V_t = P_t E \left[ H (\hat{x}_{nT}, nT) \mid \mathcal{F}^Y_t \right] = E \left[ H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \mid \mathcal{F}^Y_t \right], \tag{11}
\]

as a key state variable (instead of \( P_t \) itself) to capture the inside trader’s trading policy.

The second difference is that for the inside trader, in addition to the terminal value \( A (\hat{x}_{nT}, nT) \), he has another advantage that he knows the underlying fundamental \( \hat{x}_{nT} \). This implies he behaves like a risk-neutral agent where A-SDF is always 1. His profit is determined by \( \{ A (\hat{x}_{nT}, nT), P_t \} \) instead of \( \{ H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT), V_t \} \).\(^{17}\) The inconsistency of the discount rate between the inside trader and the market makers makes the characterization of the equilibrium challenging.

Before I characterize the equilibrium, let me first state the following lemma where I follow the literature to assume the value of \( A (\hat{x}_{nT}, nT) \) follows a log-normal distribution. More specifically, I specify \( \ln A (\hat{x}_{nT}, nT) = c\hat{x}_{nT} + a(nT) \), where \( c > 0 \) so that the asset value is increasing in the underlying growth rate.\(^{18}\)

\begin{lemma}
\forall t \in [nT - 1, nT], if the insider adopts a trading strategy of the form given in \end{lemma}

\(^{17}\)When the market makers are risk neutral as in the standard literature, the inside trader and market makers have the same goal \( \{ A (\hat{x}_{nT}, nT), P_t \} \).

\(^{18}\)One example of \( A (\hat{x}_t, t) \approx e^{\frac{x_t + \phi}{\sigma} \hat{x}_t + a(t)} \) is through the stock which has the claim to the following dividend process:

\[
\frac{dD_t}{D_t} = [\bar{x} + \phi (x_t - \bar{x})] dt + \phi \sigma \mathrm{d}B_{C,t}, \tag{12}
\]

where we allow the leverage parameter \( \phi \geq 1 \) so that dividends are more risky than consumption, as in Bansal and Yaron (2004). In addition, the shock \( dB_{D,t} \) is independent of \( dB_{C,t} \) and \( dB_{x,t} \). The proof is in the appendix.
\[
\theta_t = \frac{c (\hat{x}_{nT} - \hat{x}_{nT-1}) + \frac{h}{c} \sigma_v^2}{(nT - t) \lambda} - \frac{Y_t}{nT - t}
= \frac{\log [A(\hat{x}_{nT}, nT)] - \mu_P + \frac{h}{c} \sigma_v^2}{(nT - t) \lambda} - \frac{Y_t}{nT - t},
\]

where \( \lambda = \frac{\sigma_v}{\sigma_z} \) and \( \mu_P = cx_{nT-1} + N(nT) \), then the market maker’s estimation on \( H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT) \) given by \( V_t = E[ H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT) | \mathcal{F}^Y_t ] \) has the following dynamics

\[
\frac{dV_t}{V_t} = \frac{c - h}{c} \lambda \left[ dY_t - \frac{1}{\lambda} \frac{h}{c} \sigma_v^2 dt \right].
\]

It starts at \( V_{nT-1} = e^{\mu_V + \frac{1}{2} \left( \frac{c - h}{c} \right)^2 \sigma_v^2} \) with

\[
\mu_V = (c - h) x_{nT-1} + H(nT) + N(nT); \quad \sigma_v^2 = c^2 [Q_{nT-1} - Q_{nT}].
\]

Denote \( \Lambda_t = E[ H(\hat{x}_{nT}, nT) | \mathcal{F}^Y_t ] \) as the market maker’s estimation on \( H(\hat{x}_{nT}, nT) \). It has the following dynamics

\[
\frac{d\Lambda_t}{\Lambda_t} = \frac{-h}{c} \lambda \left[ dY_t - \frac{1}{\lambda} \frac{h}{c} \sigma_v^2 dt \right],
\]

which starts at \( \Lambda_{nT-1} = e^{\mu_{\Lambda} + \frac{1}{2} \left( \frac{h}{c} \right)^2 \sigma_v^2} \) where \( \mu_{\Lambda} = -hx_{nT-1} + H(nT) \).

Define an adjusted order flow \( \hat{Y}_t \) as

\[
\hat{Y}_t \equiv Y_t - \int_{nT-1}^t E \left[ \theta_s | \mathcal{F}^Y_s \right] ds = Y_t - \frac{1}{\lambda} \frac{h}{c} \sigma_v^2 \left[ t - (nT - 1) \right],
\]

where

\[
E \left[ \theta_t | \mathcal{F}^Y_t \right] = \frac{1}{\lambda} \frac{h}{c} \sigma_v^2
\]

is the market makers’ expectation of the inside trader’s order rate. In addition, \( \hat{Y}_t \) is a Brownian Motion with instant variance \( \sigma_v^2 \) with respect to the market maker’s filtration \( \mathcal{F}^Y_t \).
The market makers’ estimation on $H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT)$ and $H(\hat{x}_{nT}, nT)$ are functions of the current adjusted cumulative order flow $\hat{Y}_t$. From the definition of the price:

$$P_t = E[H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT) | F^Y(t)] = \frac{V(t, \hat{Y}_t)}{A(t, \hat{Y}_t)},$$

it implies the price rule is also a function of the current adjusted order flow. Then, given the make makers’ pricing rule, $P(t) = P(t, \hat{Y}_t)$, the informed trader chooses the order rate to maximize her trading profit. That is,

$$J(t, y, A(\hat{x}_{nT}, nT)) = \max_{\theta \in \mathcal{A}} E \left[ \int_t^{nT} \left( A(\hat{x}_{nT}, nT) - P(s, \hat{Y}_s) \right) \theta_s ds | \hat{Y}_t = y, A(\hat{x}_{nT}, nT) \right]$$

subject to

$$d\hat{Y}_t = \left[ \theta_t - \frac{1}{\lambda} c \frac{\sigma_v^2}{\sigma_z} \right] dt + dZ_t.$$

The equilibrium is characterized by the following theorem:

**Theorem 1.** There exists an equilibrium where the price process $P_t$ and optimal strategy of the insider is $\forall t \in [nT - 1, nT]$,

$$\frac{dP(t, \hat{Y}_t)}{P(t, \hat{Y}_t)} = \lambda d\hat{Y}_t + \frac{h}{c \sigma_v^2} dt$$

$$\theta(t, \hat{Y}_t) = \frac{\log [A(\hat{x}_{nT}, nT)] - \mu_P}{(nT - t) \lambda} - \frac{\hat{Y}_t}{nT - t} + \frac{1}{\lambda} \frac{h}{c} \sigma_v^2$$

with $P_{nT-1} = e^{\mu_P - \frac{1}{2} \left( \frac{h}{c} \right)^2 \sigma_v^2 + \frac{1}{2} \left( \hat{Y}_{nT-1} \right)^2 \sigma_v^2}$ and $\lambda = \frac{\sigma_v}{\sigma_z}$, where

$$\mu_P = cx_{nT-1} + N(nT), \quad \sigma_v^2 = c^2 [Q_{nT-1} - Q_{nT}].$$

The maximized expected profit of the insider is

$$J(t, P(t, \hat{Y}_t), A(\hat{x}_{nT}, nT)) = \frac{1}{2} \sigma_z \sigma_v (nT - t) A(\hat{x}_{nT}, nT)$$

$$+ \frac{P(t, \hat{Y}_t) - A(\hat{x}_{nT}, nT) + A(\hat{x}_{nT}, nT) \left[ \log A(\hat{x}_{nT}, nT) - \log P(t, \hat{Y}_t) \right]}{\lambda}.$$
Further, with respect to inside trader’s filtration, $P \left( t, \hat{Y}_t \right)$ converges almost surely to $A(\hat{x}_{nT}, nT)$ at time $nT$. With respect to the market makers’ filtration, both pricing rule $P \left( t, \hat{Y}_t \right)$ and price-response function $P_y \left( t, \hat{Y}_t \right)$ are submartingales with a deterministic growth rate $\frac{h}{c} \sigma_v^2$ per unit of time.

Last, the market depth, as defined by $P \left( t, \hat{Y}_t \right)$, is constant and the same as the standard Kyle model.

The key difference comparing to Kyle is that the expected order rate from the informed under the market makers’ conditional information set $\mathcal{F}_t^Y$ is not zero. That is,

$$E \left[ \theta_t | \mathcal{F}_t^Y \right] = \frac{1}{\lambda} \frac{h}{c} \sigma_v^2 = h \sigma_z \sqrt{Q_{nT-1} - Q_{nT}} > 0,$$

when $\gamma > \frac{1}{\psi}$ since the trading can reduce the uncertainty. When the agent is risk-neutral to future news, the model goes back to Kyle where the expected order rate is zero.

Given this equilibrium, I can explicitly characterize the pre-FOMC announcement drift and the implied variance reduction at each time $t \in [nT - 1, nT]$:

**Proposition 1.** The unconditional expected return for any $t \in [nT - 1, nT]$ is

$$\log E \left[ \frac{P_t}{P_{nT-1}} | \mathcal{F}_{nT-1}^Y \right] = \frac{h}{c} \sigma_v^2 (t - (nT - 1)) = hc [Q_{nT-1} - Q_{nT}] (t - (nT - 1)) \quad (22)$$

$$= \frac{\gamma - \frac{1}{\psi} c}{a_x + \beta} [Q_{nT-1} - Q_{nT}] (t - (nT - 1)), \quad (23)$$

which implies the pre-FOMC announcement drift grows at a constant rate $\frac{\gamma - \frac{1}{\psi} c}{a_x + \beta} [Q_{nT-1} - Q_{nT}]$.

The reduction of the uncertainty at time $t$ comparing to $nT - 1$ is

$$\text{Var} \left[ \ln P_{nT} | \mathcal{F}_t^Y \right] - \text{Var} \left[ \ln P_{nT} | \mathcal{F}_{nT-1}^Y \right] = -c^2 [Q_{nT-1} - Q_{nT}] [t - (nT - 1)] .$$

Thus, the uncertainty reduces at a constant rate $c^2 [Q_{nT-1} - Q_{nT}]$ per unit of time.
This figure shows the average implied variance change before announcements and the pre-FOMC announcement drift in the model. The blue dashed line captures the results in the benchmark that market makers are risk averse to the fundamentals. The red dashed line shows the results when $h = \frac{1}{\psi}$, i.e., the Kyle model where market makers are risk neutral to the fundamentals. The parameters are reported in Table 1.

Though the pre-FOMC announcement drift depends on the curvature of continuation utility function, the reduction of the uncertainty only depends on the variance of asset value. This gives me a clear discipline of the two important parameters: $c$ and $h$. Proposition 2 compares my model with the original Kyle’s model, where the market maker is risk neutral to the future news, i.e., $\gamma = \frac{1}{\psi}$.
This figure shows the average implied variance change before announcements and the pre-FOMC announcements drift in the model. The blue dashed line capture the results in the benchmark that market makers are risk averse to the fundamentals. The red dashed line shows the results when $h = 1/\psi$, i.e., the Kyle model where market makers are risk neutral to the fundamentals. The parameters are reported in Table 1.

**Proposition 2.** The pre-FOMC announcement premium would be zero if the market makers’ are risk neutral to future news, i.e when $h = 0$. Besides, for any terminal value $A(\hat{x}_{nT}, nT)$, under the same realization of $Z_t$, the difference between the price $P_t$ when $h > 0$ and the price $P_{t}^{Kyle}$ when $h = 0$ decreases at a deterministic rate:

$$\log P_{t}^{Kyle} - \log P_t = hc \left[ Q_{nT-1} - Q_{nT} \right] \left[ nT - t \right];$$

the difference between the inside trader’s trading strategy $\theta_t$ when $h > 0$ and the inside trader’s
trading strategy $\theta^Kyle_t$ when $h = 0$ (i.e., $\gamma = \frac{1}{\psi}$) is constant:

$$\theta_t - \theta^Kyle_t = \frac{1}{\lambda c} \sigma_v^2 h.$$ 

$\forall t$, the reduced uncertainty for both economics are the same.

This proposition highlights the main mechanism of the pre-announcement drift in the paper: the market makers are compensated for the risk of holding assets. Therefore, the information leakage before announcement result in non-trivial reductions of uncertainty, and are associated with realizations of a substantial amount of equity premium.

**Proposition 3.** The expected profits of insiders at time $nT - 1$ is

$$E \left[ J (nT - 1, P_{nT-1}, A (\hat{x}_{nT}, nT)) \right] = \frac{P_{nT-1} - e^{\mu P + \frac{1}{2} \sigma_v^2} + c + h}{\lambda} \sigma_v e^{\mu P + \frac{1}{2} \sigma_v^2},$$

which is an increasing in the curvature of continuation utility $h = \gamma - \frac{1}{\psi}$.

Figure 6: Expected profits

This figure shows the expected profits insiders is increasing in the curvature of continuation utility $h = \gamma - \frac{1}{\psi}$ and in the risk exposure $c$.

This proposition indicates that the literature underestimates insiders’ profits using the traditional Kyle model, especially when the risk exposure of risky asset is large. When the maker
makers are risk averse to future news, insiders have another advantage that their A-SDF is 1 since there is no aggregate uncertainty to them.

5 The inside trader may not know the inside information

From the empirical section, not all the FOMC announcements are the same. Motivated by this fact, I extend the above model by assuming the insider trader may or may not be informed of the signal $s_n$ before announcements.\footnote{This part is based on Li (2013), which extends Kyle (1985) to study the insider trading with uncertain informed trading. He still assumes that market makers and inside traders are risk-neutral.} In the meantime, the market makers are not sure whether the inside trader has observed the signal or not. The market makers share a common belief that such an event, in which the inside trader observes this information earlier than the public, occurs with a probability $\pi_0$ at time 0.\footnote{I can interpret the information asymmetry to be severe in a particular market if this probability is high.} This implies, in addition to the value of the risky asset and the A-SDF, the market maker also have to update their estimate of the probability that the inside trader has private information.

5.1 Model setting

Let $X_\delta(t)$ denote the net orders from the inside trader. Then the total cumulative order flow $Y(t)$ can be expressed as

$$Y(t) = X_\delta(t) + Z(t),$$

where $\delta$ is an indicator function, which is equal to 1 if the informed trader has information and is equal to 0 otherwise. By observing this order flow, the market makers update their estimates about the probability that the inside trader possesses private information and the value of the risky security.

Let $\mathcal{F}_0(t) = \mathcal{F}^Y(t) \times \{\delta = 0\}$ under the hypothesis $\delta = 0$ and $\mathcal{F}_1(t) = \mathcal{F}^Y(t) \times \{\delta = 1\}$ under the hypothesis $\delta = 1$. If the inside trader does not have any private information ($\delta = 0$), then he has the same A-SDF as the market maker, which is defined in equation (8). Therefore,
the best estimate of the security’s value is

\[
\hat{v}^* \equiv E \left[ \frac{H (\hat{x}_{nT}, nT)}{E[H (\hat{x}_{nT}, nT | \mathcal{F}_{nT-1})]} A (\hat{x}_{nT}, nT) | \mathcal{F}_{nT-1} \right]
\]

\[
= \frac{E[H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) | \mathcal{F}_{nT-1}]}{E[H (\hat{x}_{nT}, nT | \mathcal{F}_{nT-1})]} \equiv \bar{V}
\]

where

\[
\bar{V} = E[H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) | \mathcal{F}_0 (t)] ; \quad \bar{\Lambda} = E[A (\hat{x}_{nT}, nT) | \mathcal{F}_0 (t)]
\]

is the estimate of \(H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT)\) and \(A (\hat{x}_{nT}, nT)\) respectively when the inside trader does not have any private information.

Let

\[
\pi (t) = E [\delta | \mathcal{F} (t)] ; \quad V^* (t) = E \left[ \frac{H (\hat{x}_{nT}, nT)}{E[H (\hat{x}_{nT}, nT | \mathcal{F}_1 (t))]} A (\hat{x}_{nT}, nT) | \mathcal{F}_1 (t) \right] = \frac{V(t)}{\Lambda(t)}
\]

be the estimate of the probability that the inside trader has private information at time \(t\) and the value estimate of the risky security at time \(t\) conditional on \(\delta = 1\), respectively, where I define

\[
V (t) = E[H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) | \mathcal{F}_1 (t)] ; \quad \Lambda (t) = E[A (\hat{x}_{nT}, nT) | \mathcal{F}_1 (t)]
\]

With the uncertainty of \(\delta\), market makers estimate the discounted value under the information structure \(\mathcal{F}_1 (t)\) and estimate the probability that the strategic trader has observed private information under the information structure \(\mathcal{F}^Y (t)\).

Given these two estimates, the market makers, set the price

\[
P (t) = E \left[ \frac{H (\hat{x}_{nT}, nT)}{E[H (\hat{x}_{nT}, nT | \mathcal{F} (t))]} A (\hat{x}_{nT}, nT) | \mathcal{F} (t) \right]
\]

\[
= \frac{E[H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) | \mathcal{F} (t)]}{E[H (\hat{x}_{nT}, nT | \mathcal{F} (t))]} \\
= \frac{\pi (t) V (t) + (1 - \pi (t)) \bar{V}}{\pi (t) \Lambda (t) + (1 - \pi (t)) \bar{\Lambda}}.
\]

(24)
Note that the market makers set the price
\[ P(t) = E \left[ \frac{H(\hat{x}_{nT}, nT)}{E[H(\hat{x}_{nT}, nT) | \mathcal{F}_1(t)]} A(\hat{x}_{nT}, nT) | \mathcal{F}_1(t) \right] \]
when \( \pi(t) = 1 \). This goes back to the case in the previous section.

Following Back (1992) and Li (2013), I assume that the insider chooses an absolutely continuous trading rule \( \theta(t, V) \) that belongs to an admissible set \( \mathcal{A} = \{ \theta \text{ s.t. } E \left[ \int_{nT}^{nT} \theta^2(t, V) \, ds \right] < \infty \} \).

When the strategic trader has no information other than what the market makers have, her order rate becomes \( \theta(t, \bar{V}) \). Given this trading strategy, the cumulative flow is
\[ Y(t) = \int_0^t \theta(s, V) \, ds + Z(t). \quad (25) \]

I impose the following restriction on the market makers’ value estimates of the underlying value conditional on \( \delta = 1, V(t) \):
\[ E \left[ \int_{nT-1}^{nT} V^2(s) \, ds \right] < \infty \]
This restriction implies that the pricing rule defined by equation (24) satisfies
\[ E \left[ \int_{nT-1}^{nT} P^2(s) \, ds \right] < \infty \]
which is sufficient to rule out the so-called doubling strategy that the informed trader could use.

### 5.2 The equilibrium

An \emph{equilibrium} is an quadruple \((X_0, X_1, P, \Pi)\) such that

1. both \( X_0 \) and \( X_1 \) are the optimal trading strategies of the inside trader when she has not or has observed private information, respectively, given \( P(t) \) and \( \Pi \);
2. \( P(t) = \frac{\prod(t)V(t)+\pi(t)V}{\prod(t)\Lambda(t)+\pi(t)\Lambda} \) is the stock price at time \( t \), where \( V(t) \) and \( \Lambda(t) \) are the market makers’ value estimates of the risky security and SDF conditional on \( \delta = 1 \), and \( \prod(t) = \pi(t) \), is the market makers’ probability estimates that the inside trader has private information, given the insider trader’s trading strategies \( X_0 \) and \( X_1 \).

From the market makers’ point of view, the cumulative order flow has two possible interpretations:

\[
dY(t) = \begin{cases} 
\theta(t, V) dt + dZ(t) & \text{if the insider is informed;} \\
\theta(t, \bar{V}) dt + dZ(t) & \text{if the insider is not informed.}
\end{cases}
\] (26)

Given the observation of the cumulative order flow, the market makers update the probability that the strategic trader has private information, the SDF as well as the value of the security conditional on the insider is informed. These estimates are done by solving a nonlinear filtering problem. The equilibrium is characterized in Theorem. I provide the related proof in Appendix B.

**Figure 7: Model implications: uncertainty and return**

This figure shows the average implied variance change before announcements and the pre-FOMC announcement drift in the model where the priors follow a uniform distribution. The blue dashed line captures the results in the benchmark that market makers are risk averse to the fundamentals. The red dashed line shows the results when \( h = \frac{1}{\psi} \), where market makers are risk neutral to the fundamentals. The parameters are reported in Table 1.
Theorem 2. \( \forall t \in [nT - 1, nT] \), the pricing rule is

\[
P(t, y) = P_{nT-1} \frac{\Pi(t, y) e^{\frac{c-h}{c}y - \frac{1}{2} \left( \frac{c-h}{c} \right)^2 \sigma_z^2 (t-(nT-1))} + 1 - \Pi(t, y)}{\Pi(t, y) e^{\frac{c-h}{c}y - \frac{1}{2} \left( \frac{c-h}{c} \right)^2 \sigma_z^2 (t-(nT-1))} + 1 - \Pi(t, y)}
\]

where \( P_{nT-1} = e^{\frac{c-h}{c} \bar{y} - \frac{1}{2} \left( \frac{c-h}{c} \right)^2 \sigma_z^2} \) and \( \lambda = \frac{\sigma_z}{\sigma_v} \). The market makers’ probability estimate \( \Pi(t, y) \) given by

\[
\Pi(t, y) = \frac{\pi_0 \exp \left( \frac{1}{2 \sigma^2} \left( \frac{y - \bar{y}}{nT - t} \right)^2 + \frac{1}{2} \ln \left( nT - t - \frac{\bar{y}^2}{2 \sigma^2} \right) \right)}{1 - \pi_0 + \pi_0 \exp \left( \frac{1}{2 \sigma^2} \left( \frac{y - \bar{y}}{nT - t} \right)^2 + \frac{1}{2} \ln \left( nT - t - \frac{\bar{y}^2}{2 \sigma^2} \right) \right)},
\]

where \( h(y) = \exp \left( \mu \nu + \frac{c-h}{c} \lambda y \right) \) and \( y \) represents the adjusted order flow \( \hat{Y}_1(t) \) and \( \bar{y} = h^{-1}(\bar{V}) = \frac{c-h}{2\lambda} \frac{\sigma^2}{\sigma_v} \) since \( \bar{V} = e^{\mu \nu + \frac{1}{2} \left( \frac{c-h}{c} \right)^2 \sigma_z^2} \).

The adjusted order flow \( \hat{Y}_1(t) \) follows

\[
d\hat{Y}_1(t) = \frac{h^{-1}\left( H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT) \right) - \hat{Y}_1(t)}{nT - t} \, dt + dZ(t)
\]

\[
= \frac{(\log \left[ H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT) \right] - \mu \nu) / \left( \frac{c-h}{c} \lambda \right) - \hat{Y}_1(t)}{nT - t} \, dt + dZ(t)
\]

where \( \hat{Y}_1(nT - 1) = 0 \).

The total trading volume is

\[
Y(t) = \hat{Y}_1(t) + \Theta(t, \hat{Y}_1(t))
\]

under

\[
\Theta(t, \hat{Y}_1(t)) = \left[ 1 - \Pi(t, \hat{Y}_1(t)) \right] \frac{[\hat{Y}_1(t) - \bar{y}]}{nT - t} + \bar{\theta}(t, \hat{Y}_1(t)).
\]

Here \( \bar{\theta}(t, \hat{Y}_1(t)) \) is the expected order rate from the inside trader under market makers’ un-
conditional) information.

\[
E \left[ \theta \left( t, \hat{Y}_1 (t), H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \right) \mid \mathcal{F}_t^Y \right] = \tilde{\theta} \left( t, \hat{Y}_1 (t) \right)
\]

\[
= \frac{1}{\lambda c^2 \sigma_t^2} \Pi \left( t, \hat{Y}_1 (t) \right) E \left( t, \hat{Y}_1 (t) \right) - \Pi \left( t, \hat{Y}_1 (t) \right) \left( 1 - \Pi \left( t, \hat{Y}_1 (t) \right) \right) \frac{\hat{Y}_1 (t) - \bar{y}}{nT - t} \left( E \left( t, \hat{Y}_1 (t) \right) - 1 \right)
\]

where \( E \left( t, \hat{Y}_1 (t) \right) = e^{-\frac{h}{c} \lambda \hat{Y}_1 (t) - \frac{1}{2} (\frac{h}{c})^2 \sigma_t^2}. \) The inside trader’s trading strategy follows

\[
\theta \left( t, \hat{Y}_1 (t), H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \right) = \bar{\tilde{\theta}} \left( t, \hat{Y}_1 (t) \right) + \frac{\left( \log \left[ H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \right] - \mu_Y \right) / (\frac{c-h}{c}) \lambda - \bar{y} - \Pi \left( t, \hat{Y}_1 (t) \right) \left[ \hat{Y}_1 (t) - \bar{y} \right]}{nT - t}.
\]

Given this equilibrium, I can explicitly characterize the pre-FOMC announcement drift and the implied variance reduction, where the distribution of \( \pi_0 \) is characterized by a smooth cumulative distribution function \( G (\pi) \) for all \( \pi \in (0, 1) \) at time 0.

**Proposition 4.** Given a smooth distribution of the priors \( G (\pi_0) \), the premium realized before announcements is

\[
\log E_{nT-1} \left[ \frac{P_{nT}}{P_{nT-1}} \right] = \eta c \gamma - \frac{a_x}{\beta} Q_T - Q_T
\]

where \( \eta \) is the fraction that inside traders observe the signal. The reduction of the uncertainty at time \( t \) comparing to \( nT - 1 \) is

\[
\text{Var} \left[ \ln P_{nT} \mid \mathcal{F}_t^Y \right] - \text{Var} \left[ \ln P_{nT} \mid \mathcal{F}_{nT-1}^Y \right] = -E \left[ \Pi \left( t, \hat{Y}_1 (t) \right) \mid \mathcal{F}_t^Y \right] c^2 [Q_{nT-1} - Q_{nT}] [t - (nT - 1)],
\]

which implies the uncertainty reduction before announcements is

\[
\text{Var} \left[ \ln P_{nT} \mid \mathcal{F}_{nT}^Y \right] - \text{Var} \left[ \ln P_{nT} \mid \mathcal{F}_{nT-1}^Y \right] = -\eta c^2 [Q_{nT-1} - Q_{nT}] [t - (nT - 1)],
\]

Proposition 4 states the reduction of uncertainty would be \( c^2 [Q_{nT-1} - Q_{nT}] \) just before announcements when insider is better informed. Otherwise, the reduction of uncertainty is
zero when the insider is not better informed. Under this case, all the uncertainties are resolved after FOMC announcements. Figure 7 shows the uncertainty reduction and realized return under the case that market makers’ priors follow a uniform distribution. Instead of a linear relationship as in the benchmark, the functions are convex in time $t$ since market makers update the probability $\pi_t$ is a convex function in $t$ when insiders are better informed after they observe more information as time goes by. Figure 8 captures the uncertainty reduction and return upon announcements in the probability that insiders are better informed. If the insiders are more likely to be informed, the uncertainty reduction and realized return upon announcements would be higher.

**Proposition 5.** Suppose the measure of the inside trader who observes the signal at time $nT - 1$ is strictly positive, the pre-announcement drift is zero if the market makers are risk neutral to future news. If the insider trader does not observe the signal, there is no pre-announcement drift and all the premium is realized upon announcements. In the meantime, there is no reduction of uncertainty before announcements.

Figure 8: Uncertainty reduction and return upon announcements

This figure shows the average implied variance change before announcements and the pre-FOMC announcements drift in the model. The black dashed line capture the theoretical expressions in proposition 1, which perfectly identical to the simulations through the law of motions in theorem 1. The parameters are reported in Table 1.
6 Further implications and related literature

So far, I establish that the resolution of uncertainty via a news leak results in the pre-FOMC announcement drift in the stock market even the leaked news is on average neutral. Since the drift is driven by the reduction of uncertainty instead of on average positive monetary policy surprises, the realized pre-FOMC market return is not correlated with monetary policy shock and post-announcement return.

The information channel I emphasize is consistent with recent work by Nakamura and Steinsson (2018), who study Federal Reserve announcements affect beliefs not only about monetary policy but also about economic fundamentals. In addition to the change of the current federal funds rate, the FOMC announcements highlight the economic conditions as well as the future path of the federal funds rate.

As shown in section 5, the model only requires that the insiders know policy decisions for some FOMC announcements instead of all of them. This is consistent with Figure 2 that uncertainty reduction only happens before some FOMC announcements, which are associated with substantial pre-FOMC returns. The model of this paper is supported by Cieslak, Morse, and Vissing-Jorgensen (2019), who find a substantial list of leaks from the Fed. Besides, Vissing-Jorgensen (2019) provides a history of leak discussions in FOMC documents for the period 1948-2013 in order to show that the FOMC itself expresses frequent concerns about leaks.

6.1 Other explanations

Hu, Pan, Wang, and Zhu (2020) and Laarits (2019) contribute the pre-FOMC drift to uncertainty reduction before FOMC news in a representative-agent framework. There are two main differences comparing to this paper: (1) The market news carries two different types of risks, and only one type of risk is resolved before FOMC announcements. (2) All investors observe the resolved information at the same time, i.e., there is no asymmetric information caused by leakage.\footnote{The two risks are different in these two papers. In Hu, Pan, Wang, and Zhu (2020), the uncertainty about the potential magnitude of the news’ market impact is resolved before announcements, while the risk associated...} In addition to the pre-FOMC drift, their model also predicts substantial returns...
after announcements when the other risk is resolved. However, the post-FOMC announcement return is not significantly different zero, as documented by Lucca and Moench (2015) and Hu, Pan, Wang, and Zhu (2020). The pre-FOMC part of their models share the same time-series and cross-sectional predictions with this paper that the drift is increasing in the uncertainty reduction of the announcements and the beta of stocks.

7 Conclusion

In this paper, I show that information leakage contributes to the FOMC-announcement drift. The significant reduction of VIX index before announcements indicates the market uncertainty is resolved symmetrically through information leakage. To account for the dynamics of uncertainty and returns jointly, I extend Kyle’s (1985) model of insider trading to the case where market makers require a risk compensation for news related to fundamentals. The trading of insiders resolves uncertainty gradually, which results in an upward drift in market prices even the leaked news is on average neutral. Furthermore, given uncertainty is not always reduced before some FOMC meetings, I generalize the model that insiders may not be better informed, and market makers have to update their estimate of the probability that insiders have private information. This model sheds light on the puzzle that pre-FOMC returns exist on equities but not in fixed income assets since market makers’ priors may vary in the types of markets.

Laarits (2019) assume there are two types of announcements the Fed will make, which will reveal either monetary policy stance or long-term growth expectations. All investors learn the type of announcements before FOMC news, which resolves part of the uncertainty.
APPENDICES

The following appendices provide details of the proof in section 3 and 4. Appendix A contains all the proofs for the economy that the insider is always informed. Appendix B provides the details for the economy that market makers are uncertain about whether the insider is informed or not. Appendix C shows the related proof of the macroeconomic environment.

Appendix A  Proof of Theorem 1

The proof is in several steps. First, I prove Lemma 2, which shows if the market maker conjectures that the insider’s trading strategy follows equation (13), the price rule would be a function of the adjusted order flow $\hat{Y}_t$. Second, I characterize the price dynamics as a function of $\hat{Y}_t$ under the same trading strategy stated in Lemma 3. Third, I establish the optimality of this trading strategy through a verification proof.

A.1 Step 1: Market Maker’s Updating

Here is the proof of Lemma 2.

Proof. The conjectured trading strategy (13) implies that

$$\theta_t = \frac{\log [A(\hat{x}_{nT}, nT)] - \mu_P + \frac{h}{c} \sigma_v^2}{(nT - t) \lambda} - \frac{Y_t}{nT - t}$$

$$= \frac{\frac{c-h}{c} \left\{ \log [A(\hat{x}_{nT}, nT)] - \mu_P + \frac{h}{c} \sigma_v^2 \right\}}{(nT - t) \left( \frac{c-h}{c} \lambda \right)} - \frac{Y_t}{nT - t}$$

$$= \left( \frac{\log [H(\hat{x}_{nT}, nT)] A(\hat{x}_{nT}, nT) - \left( \mu_V - \frac{h(c-h)}{c^2} \sigma_v^2 \right)}{nT - t} \right) / \left( \frac{c-h}{c} \lambda \right) - Y_t,$$

where the last equality comes from $H(\hat{x}_{nT}, nT) = e^{-h\hat{x}_{nT} + \mathcal{H}(nT)}$, $A(\hat{x}_{nT}, nT) = e^{\hat{x}_{nT} + N(nT)}$ and the definition of $\mu_V, \mu_P$ in equation (15) and (20).

Therefore, the aggregate trading volume follows

\[ (A1) \]
\[
dY_t = \theta_t dt + dZ_t = \left( \log \left[ H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT) \right] - \left( \mu_V - \frac{h}{c^2} \sigma_v^2 \right) \right) / \left( \frac{c}{c^2} \lambda \right) - Y_t \\
\frac{nT - t}{dt} + \sigma_z dB_t
\]

(A2)

where \(dZ_t = \sigma_z dB_t\) and \(Y_{nT-1} = 0\). Now let me define the observation and innovation process. Set \(Y^*_{nT-1} = 0\) and

\[
dY^*_t = \frac{1}{\sigma_z} \left( dY_t + \left( \mu_V - \frac{h}{c^2} \sigma_v^2 \right) / \left( \frac{c}{c^2} \lambda \right) + Y_t \right) dt
\]

the last equality comes from \(\lambda = \frac{\sigma_v}{\sigma_z}\). Because \(Y_t\) are observable to market makers, \(Y^*\) is also observable. The corresponding innovation process is given by

\[
 dB^*_t = \frac{\log \left[ H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT) \right]}{c \sigma_v (nT - t)} dt + dB_t
\]

where

\[
\hat{v}_t = E \left[ \log \left[ H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT) \right] | \mathcal{F}_t^Y \right].
\]

(A3)

Therefore, the Kalman-filter equation implies

\[
d\hat{v}_t = \sum_{v,t} dB^*_t
\]

(A4)

where \(\sum_{v,t}\) is the conditional variance of \(\log \left[ H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT) \right] \) given market marker’s information (on the filtration \(\mathcal{F}_t^Y\)), i.e.

\[
\sum_v (t) = \text{Var} \left[ \log H(\hat{x}_{nT}, nT) v(\hat{x}_{nT}, nT) | \mathcal{F}_t^Y \right].
\]

(A5)

The Kalman-filter equation also implies the dynamics of the posterior variance:

\[
\frac{1}{\sum_{v,t}} = \frac{1}{\sum_{v,0}} + \int_{nT-1}^t \frac{1}{(nT - s)^2} (\frac{c-h}{c})^2 \sigma_v^2 ds = \frac{1}{(\frac{c-h}{c})^2 \sigma_v^2} + \frac{t - (nT - 1)}{(nT - t)(\frac{c-h}{c})^2 \sigma_v^2} = \frac{1}{(nT - t)(\frac{c-h}{c})^2 \sigma_v^2}
\]

(A6)
which implies
\[ \sum_{v,t} = \left( \frac{c-h}{c} \right)^2 \sigma_v^2 (nT - t) \]  
(A7)

and the filtering equation (A4) is
\[ d\hat{v}_t = \frac{c-h}{c} \sigma_v dB_t^* = \frac{\log \left[ H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \right] - \hat{v}_t}{nT - t} dt + \frac{c-h}{c} \sigma_v dB_t \]

Define an adjusted order flow \( \hat{Y}_t \) as
\[ \hat{Y}_t \equiv Y_t - \int_{nT-1}^t \left( \frac{1}{\lambda c} \sigma_v^2 \right) ds = Y_t - \frac{1}{\lambda c} \sigma_v^2 [t - (nT - 1)]. \]  
(A8)

From the aggregate trading volume (A2), the adjusted order flow follows
\[ d\hat{Y}_t = dY_t - \frac{1}{\lambda c} \sigma_v^2 dt \]
\[ = \left( \log \left[ H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \right] - \left( \mu_V - \frac{h(c-h)}{c^2} \sigma_v^2 \right)/ \left( \frac{c-h}{c} \right) \right) \frac{dt}{nT - t} + \sigma_z dB_t - \frac{1}{\lambda c} \sigma_v^2 dt \]
\[ = \left( \log \left[ H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \right] - \left( \mu_V - \frac{h(c-h)}{c^2} \sigma_v^2 \right)/ \left( \frac{c-h}{c} \right) \right) \frac{dt}{nT - t} + \sigma_z dB_t - \frac{1}{\lambda c} \sigma_v^2 dt \]
\[ = \left( \log \left[ H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \right] - \mu_V / \left( \frac{c-h}{c} \right) \right) \frac{dt}{nT - t} + \sigma_z dB_t. \]  
(A9)

This implies
\[ \frac{c-h}{c} \lambda d\hat{Y}_t = \left[ \log \left[ H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \right] - \left[ \mu_V + \frac{c-h}{c} \lambda \hat{Y}_t \right] \right] \frac{dt}{nT - t} + \sigma_z dB_t. \]  
(A10)

Since \( \hat{v}_{nT-1} = E \left[ \log H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) | \mathcal{F}_{nT-1} \right] = \mu_V \) and \( \hat{Y}_{nT-1} = 0, \)
\[ d\hat{v}_t = \frac{c-h}{c} \sigma_v dB_t^* = \frac{\log \left[ H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \right] - \hat{v}_t}{nT - t} dt + \frac{c-h}{c} \sigma_v dB_t \]
\[ \equiv \frac{c-h}{c} \lambda d\hat{Y}_t = \frac{c-h}{c} \lambda \left[ d\hat{Y}_t - \frac{1}{\lambda c} \sigma_v^2 dt \right] \]  
(A12)
From the filtering theory, $dB^*_t$ is a standard Brownian Motion with respect to market maker's filtration. Therefore, the adjusted order flow $\hat{Y}_t$ is a Brownian Motion with instant variance $\sigma_v^2$ under $\mathcal{F}_t^Y$. This implies

$$E \left[ \theta_t | \mathcal{F}_t^Y \right] = \frac{1}{\lambda_c} \sigma_v^2$$

is the market makers’ expectation of the inside trader’s order rate, which is strictly positive when $\gamma > \frac{1}{\psi}$ since the trading can reduce the uncertainty. When the agent is risk-neutral to future news, the model goes back to Kyle where the expected order rate is zero.

The market makers’ prior belief about $\log [H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT)]$ at time $nT - 1$ is represented by a normal distribution. The Kalman filter implies the posterior distribution of $\log [H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT)]$ under $\mathcal{F}_t^Y$ is also Gaussian, which is summarized by the posterior mean $\hat{v}_t$ and the posterior variance $\sum_{v,t}$. Therefore, the market makers’ estimation on $H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT)$ is

$$V_t = E \left[ H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) | \mathcal{F}_t^Y \right]$$

$$= E \left[ e^{\log H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT)} | \mathcal{F}_t^Y \right]$$

$$= e^{\hat{v}_t + \frac{1}{2} \sum_{v,t}} = e^{\hat{v}_t + \frac{1}{2} (nT-t)(\frac{nT}{\hat{\lambda}})^2 \sigma_v^2}$$

(A14)

Applying Ito’s Lemma, I find

$$\frac{dV_t}{V_t} = \frac{1}{V_t} \left[ V_t d\hat{v}_t + \frac{1}{2} V_t (d\hat{v}_t)^2 - \frac{1}{2} \sigma_v^2 V_t dt \right]$$

$$= d\hat{v}_t \quad (A12)$$

$$= \frac{c - \frac{h}{c}}{c} \lambda d\hat{Y}_t$$

(A15)

$\forall t \in [nT - 1, nT]$, I define $\Lambda^*_t$ as the posterior mean of $\log H (\hat{x}_{nT}, nT)$ under market makers’
Applying Ito's Lemma,

$$d\Lambda_t = -h c - \hat{v}_t + \frac{c H(nT) + h N(nT)}{c - h} \frac{d\Lambda_t}{\Lambda_t} \tag{A16}$$

with $\Lambda_{nT-1}^* = -h \hat{x}_{nT-1} + \mathcal{H}(nT)$. Similarly, for the posterior variance of $\log H(\hat{x}_{nT}, nT)$ under market makers' information,

$$\Sigma_{\Lambda^*, t} = \text{Var} \left[ \log H(\hat{x}_{nT}, nT) \right] \bigg| \mathcal{F}_t^Y \tag{A6} \Rightarrow \left( \frac{h}{c - h} \right)^2 \Sigma_{v, t} = \left( \frac{h}{c} \right)^2 (nT - t) \sigma_v^2$$

The Kalman filter implies the posterior distribution of $\log H(\hat{x}_{nT}, nT)$ under $\mathcal{F}_t^Y$ is also Gaussian, which is summarized by the posterior mean $\Lambda_t^*$ and the posterior variance $\Sigma_{\Lambda^*, t}$, which implies

$$\Lambda_t = E \left[ H(\hat{x}_{nT}, nT) \big| \mathcal{F}_t^Y \right] = E \left[ e^{\log H(\hat{x}_{nT}, nT)} \big| \mathcal{F}_t^Y \right] = e^{\Lambda_t^* + \frac{1}{2} \Sigma_{\Lambda^*, t}} = e^{\Lambda_t^* + \frac{1}{2} \left( \frac{h}{c} \right)^2 (nT - t) \sigma_v^2}$$

From Ito's Lemma,

$$\frac{d\Lambda_t}{\Lambda_t} = \frac{1}{\Lambda_t} \left[ \Lambda_t d\Lambda_t^* + \frac{1}{2} \Lambda_t (d\Lambda_t^*)^2 - \frac{1}{2} \left( \frac{h}{c} \right)^2 \sigma_v^2 \Lambda_t dt \right] = d\Lambda_t^* = -h c - h \hat{v}_t = -h c \lambda d\hat{Y}_t. \tag{A17}$$
Therefore, both $V_t$ and $\Lambda_t$ are functions of the current adjusted cumulative order flow $\hat{Y}_t$. From the following definition

$$P_t = \frac{E \left[ H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \mid \mathcal{F}_t^Y \right]}{E \left[ H (\hat{x}_{nT}, nT) \mid \mathcal{F}_t^Y \right]} = \frac{V \left( t, \hat{Y}_t \right)}{A \left( t, \hat{Y}_t \right)},$$

the price rule is also a function of the current adjusted order flow $P \left( t, \hat{Y}_t \right)$.

Furthermore, from equation (A9), the process $\hat{Y}_t$ is a Brownian bridge with instantaneous variance $\sigma^2_z$ with respect to the insider’s filtration, terminating at $(\log [H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT)] - \mu_v) / (c-hc \lambda)$ (Karatzas and Shreve (1987)). It satisfies $\dot{Y}_t \rightarrow (\log [H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT)] - \mu_v) / (c-hc \lambda)$ with probability 1 as $t \rightarrow nT$. The distribution of a Brownian bridge are the same as a Brownian motion conditional on the terminal value being known (Karatzas and Shreve (1987)). The terminal value of $\hat{Y}_t$ is the random variable $(\log [H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT)] - \mu_v) / (c-hc \lambda)$, which is normally distributed with mean zero and variance $\sigma^2_z$ and is independent of $Z$. Hence, the distribution of $\hat{Y}_t$, unconditional on the terminal value or $Z$ (i.e. from market makers’ filtration), are the distribution of a Brownian motion with variance $\sigma^2_z$. This is consistent with what I get from the filtering theory in equation (A12).

### Lemma 3
Assume the informed trader follows the strategy (19). Then, $\forall t \in [nT - 1, nT]$,

$$\frac{dP \left( t, \hat{Y}_t \right)}{P \left( t, \hat{Y}_t \right)} = \lambda d\hat{Y}_t + \frac{h}{c} \sigma^2_v dt \quad (A18)$$

with $P_{nT-1} = e^{\mu_p - \frac{1}{2} (\frac{h}{c})^2 \sigma^2_v + \frac{1}{2} (\frac{c-hc}{c})^2 \sigma^2_v}$ and $\lambda = \frac{\sigma_v}{\sigma_z}$, where

$$\mu_p = c x_{nT-1} + N (nT), \quad \sigma^2_v = c^2 \left[ Q_{nT-1} - Q_{nT} \right]. \quad (A19)$$

Further, with respect to inside trader’s filtration, $P_t$ converges almost surely to $A (\hat{x}_{nT}, nT)$ at time $nT$. 

37
Proof. First, since the price rule is a function of the total adjusted order flow \( \hat{Y}(t) \), I rewrite the inside trader’s trading rule as a function of \( \hat{Y}(t) \):

\[
\theta_t = \log \left[ A(\hat{x}_{nT}, nT) \right] - \mu_P + \frac{h}{c} \sigma_v^2 \frac{Y_t}{nT - t}
\]

\[
= \log \left[ A(\hat{x}_{nT}, nT) \right] - \mu_P + \frac{h}{c} \sigma_v^2 \frac{\hat{Y}_t}{nT - t} + \frac{1}{\lambda c} \sigma_v^2 \left( t - (nT - 1) \right)
\]

\( (A20) \)

Apply Ito’s Lemma to \( V_t = P_t E \left[ H(\hat{x}_{nT}, nT) | \mathcal{F}_t^Y \right] = P_t \Lambda_t \),

\[
\frac{dV_t}{V_t} = \frac{\lambda d\hat{Y}_t + \frac{h}{c} \sigma_v^2 dt}{P_t \Lambda_t} = \frac{dP_t}{P_t} + \frac{d\Lambda_t}{\Lambda_t} + \frac{dP_t}{P_t} \frac{d\Lambda_t}{\Lambda_t} 
\]

With the law of motion of \( V_t \) in equation (A15) and the law of motion of \( \Lambda_t \) in equation (A17), I find

\[
\frac{dP_t(t, \hat{Y}_t)}{P(t, \hat{Y}_t)} = \lambda d\hat{Y}_t + \frac{h}{c} \sigma_v^2 dt \equiv dR_t \quad (A21)
\]

where \( R_t \) is the cumulative return.

As stated in Lemma 2, \( \hat{Y}_t \rightarrow (\log \left[ H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT) \right] - \mu_P) / (\frac{c-h}{c} \lambda) \) with probability 1 as \( t \rightarrow nT \) with respect to the insider’s filtration, the equation (A21) implies

\[
\log P_t = \log P_{nT-1} + \lambda \hat{Y}_t - \frac{1}{2} \sigma_v^2 - \frac{h}{c} \sigma_v^2 \left( t - (nT - 1) \right)
\]

\( (A22) \)

\[
= \log P_{nT-1} + \lambda \hat{Y}_t - \frac{1}{2} \left( \frac{c-h}{c} \right) \sigma_v^2 \left( t - (nT - 1) \right)
\]

\[
\rightarrow c\hat{x}_{nT-1} - \frac{1}{2} \left( \frac{h}{c} \right)^2 \sigma_v^2 + N(nT) + \frac{1}{2} \left( \frac{c-h}{c} \right)^2 \sigma_v^2 + c(\hat{x}_{nT} - \hat{x}_{nT-1}) - \frac{1}{2} \left( \frac{c-h}{c} \right) \sigma_v^2
\]

\[
\rightarrow c\hat{x}_{nT} + N(nT) = \log A(\hat{x}_{nT}, nT)
\]

almost surely as \( t \rightarrow nT \) from insider’s information. This is equivalent to \( P_t \rightarrow A(\hat{x}_{nT}, nT) \) with probability 1 as \( t \rightarrow nT \).
A.2 Step 2: Insider’s Optimal Strategy

Now I can prove the main result for this step. I establish that if the dynamics of price follows equation (18), then the optimal trading strategy of the insider is indeed of the form given in equation (19) through a verification proof.

Recall that the informed trader chooses the order rate to maximize her trading profit given the make makers’ pricing rule

\[ P(t) = P(t, \hat{Y}_t). \]

The principle of optimality implies the following Bellman equation

\[
J(t, y, A(\hat{x}_{nT}, nT)) = \max_{\theta_t \in A} \left\{ \int_t^{nT} \left( A(\hat{x}_{nT}, nT) - P(s, \hat{Y}_s) \right) \theta_s ds | \hat{Y}_t = y, A(\hat{x}_{nT}, nT) \right\}
\]

subject to

\[
d\hat{Y}_t = \left[ \theta_t - \frac{1}{\lambda} \frac{h}{c} \sigma_v^2 \right] dt + dZ_t. \tag{A23}
\]

The principle of optimality implies the following Bellman equation

\[
\max_{\theta_t \in A} \left\{ (A(\hat{x}_{nT}, nT) - P(t, y)) \theta_t + J_t + J_y \left[ \theta_t - \frac{1}{\lambda} \frac{h}{c} \sigma_v^2 \right] + \frac{1}{2} \sigma_z^2 J_{yy} \right\} = 0 \tag{A24}
\]

where the subscripts denote the derivatives. The necessary conditions for having an optimal solution to the Bellman equation (A24) are

\[
J_y(t, y, A(\hat{x}_{nT}, nT)) = P(t, y) - A(\hat{x}_{nT}, nT) \tag{A25}
\]

\[
J_t + \frac{1}{2} \sigma_z^2 J_{yy} - \frac{1}{\lambda} \frac{h}{c} \sigma_v^2 J_y = 0 \tag{A26}
\]

Lemma 4. Define

\[
d\omega(s) = dZ(s) - \frac{1}{\lambda} \frac{h}{c} \sigma_v^2 ds, \quad \forall nT \geq s \geq t \geq nT - 1, \tag{A27}
\]

where \( Z(s) \) is a Brownian motion with instant variance \( \sigma_z^2 \). If there exists a strictly monotone function \( g(\cdot) \) such that the pricing rule is

\[
P(t, y) = E[g(y + \omega(nT) - \omega(t))], \tag{A28}
\]
then
\[ J(t, y, A(\hat{x}_{nT}, nT)) = E[j(y + \omega(nT) - \omega(t), A(\hat{x}_{nT}, nT))] \] (A29)
is a smooth solution to the Bellman equations (A25) and (A26), where
\[ j(y, v) = \int_{\gamma^{-1}(v)} [v - g(x)] \, dx \geq 0, \quad \forall (v, y) \]

Proof. By Theorem 7.6 in Chapter 5 of Karatzas and Shreve (1991) (Feynman-Kac representation), the value function \( J \) defined in equation (A29), is a unique solution to the Bellman equation (A26) with the terminal condition \( J(nT, y, A(\hat{x}_{nT}, nT)) = j(y, A(\hat{x}_{nT}, nT)) \).

Taking the derivative under the expectation operator yields\(^{23}\)
\[ J_y(t, y, A(\hat{x}_{nT}, nT)) = E[j_y(y + \omega(nT) - \omega(t), A(\hat{x}_{nT}, nT))] \]
\[ = E[g(y + \omega(nT) - \omega(t))] - A(\hat{x}_{nT}, nT) \]
\[ = P(t, y) - A(\hat{x}_{nT}, nT), \]
which shows \( J(t, y, A(\hat{x}_{nT}, nT)) \) also satisfy equation (A25) with \( P(t, y) \) as defined by (A28).

Lemma 5. Any continuous trading strategy that makes \( \lim_{t \to 1^-} P(t, \hat{Y}_1(t)) = v \) is optimal, where \( P(t, y) \) is as defined by (21).

\(^{23}\)The proof that the derivative of the right-hand side of (A29) can be taken under the expectation operator is similar to Back (1992).
Proof. For any trading strategy \( \theta_t \), apply Ito’s Lemma to the value function,

\[
J\left(nT, \hat{Y}_{nT}, A(\hat{x}_{nT}, nT)\right) = J\left(nT - 1, \hat{Y}_{nT-1}, A(\hat{x}_{nT}, nT)\right) + \int_{nT-1}^{nT} \left\{ J_t dt + J_y d\hat{Y}_t + \frac{1}{2} J_{yy} \left(d\hat{Y}_t\right)^2 \right\}
\]

\[
\overset{(A23)}{=} J\left(nT - 1, \hat{Y}_{nT-1}, A(\hat{x}_{nT}, nT)\right) + \int_{nT-1}^{nT} \left\{ J_t dt + J_y \left(\frac{\theta_t - \frac{1}{\lambda} \frac{h}{c} \sigma_v^2\right) dt + dZ_t \right\} + \frac{1}{2} \sigma^2 J_{yy}
\]

\[
\overset{(A26)}{=} J\left(nT - 1, \hat{Y}_{nT-1}, A(\hat{x}_{nT}, nT)\right) + \int_{nT-1}^{nT} \left\{ J_y (\theta_t dt + dZ_t) \right\}
\]

\[
\overset{(A25)}{=} J\left(nT - 1, \hat{Y}_{nT-1}, A(\hat{x}_{nT}, nT)\right) - \int_{nT-1}^{nT} A(\hat{x}_{nT}, nT) - P\left(t, \hat{Y}_t\right) (\theta_t dt + dZ_t)
\]

I can rearrange this as

\[
\int_{nT-1}^{nT} \left( A(\hat{x}_{nT}, nT) - P\left(t, \hat{Y}_t\right) \right) \theta_t dt = J\left(nT - 1, \hat{Y}_{nT-1}, A(\hat{x}_{nT}, nT)\right) - J\left(nT, \hat{Y}_{nT}, A(\hat{x}_{nT}, nT)\right)
\]

\[
- \int_{nT-1}^{nT} \left( A(\hat{x}_{nT}, nT) - P\left(t, \hat{Y}_t\right) \right) dZ_t
\]

The left-hand side is the profit of the informed trader, and the right-hand side is bounded above by

\[
J\left(nT - 1, \hat{Y}_{nT-1}, A(\hat{x}_{nT}, nT)\right) - \int_{nT-1}^{nT} \left( A(\hat{x}_{nT}, nT) - P\left(t, \hat{Y}_t\right) \right) dZ_t \quad (A30)
\]

due to the nonnegativity of \( J\left(nT, \hat{Y}_{nT}, A(\hat{x}_{nT}, nT)\right) \) in equation (A29). The no-double-strategies condition

\[
E \int_{nT-1}^{nT} P_t^2 dt < \infty
\]

implies that the stochastic integral in (A30) has a zero expectation. Therefore,

\[
E \int_{nT-1}^{nT} \left\{ A(\hat{x}_{nT}, nT) - P_t \right\} \theta_t dt \leq J\left(nT - 1, P_{nT-1}, A(\hat{x}_{nT}, nT)\right)
\]
with equality if and only if \( \hat{Y}_{nT} = g^{-1}(A(\hat{x}_{nT}, nT)) \), which is equivalent to \( P(nT, \hat{Y}_{nT}) = A(\hat{x}_{nT}, nT) \) from equation (A28). Thus, \( J(nT - 1, \hat{Y}_{nT - 1}, A(\hat{x}_{nT}, nT)) \) is an upper bound on the informed trader’s expected profit, conditional on the termination value \( A(\hat{x}_{nT}, nT) \), and the upper bound is realized - and the corresponding strategy is consequently optimal - if and only if \( P(nT, \hat{Y}_{nT}) = A(\hat{x}_{nT}, nT) \). In Lemma 3, I have already shown that the trading strategy in (19) implies \( P(t, \hat{Y}_t) \rightarrow A(\hat{x}_{nT}, nT) \) with probability 1 as \( t \rightarrow nT \). It follows that the strategy (19) is optimal.

Since \( \hat{Y}_{nT} = g^{-1}(A(\hat{x}_{nT}, nT)) \) a.s., for any scalar \( a \), the probability, given the market makers’ information at time \( nT - 1 \), that \( \hat{Y}_{nT} \leq a \) is \( F(g(A(\hat{x}_{nT}, nT))) \) where \( F \) is the distribution function of \( A(\hat{x}_{nT}, nT) \). According to Lemma , the distribution function of \( \hat{Y}_{nT} \), given the market makers’ information at time 0, is normal distribution with mean zero and variance \( \sigma^2_z \) and I denote it as \( N \). Therefore, \( N = F \circ g \), implying \( g = F^{-1} \circ N \). When \( \log A(\hat{x}_{nT}, nT) \) is normally distributed with mean \( c\hat{x}_{nT - 1} + N(nT) \) and variance \( c^2 [Q_{nT - 1} - Q_n] = \sigma^2_v \). Set \( g(y) = F^{-1}(N(y)) \):

\[
F(g(y)) = N^*(\frac{\log g(y) - [c\hat{x}_{nT - 1} + N(nT)]}{\sigma_v}) = N^*(\frac{\hat{Y}_t}{\sigma_z}),
\]

so

\[
g(y) = \exp(c\hat{x}_{nT - 1} + N(nT) + \lambda y) \quad (A31)
\]

where \( \lambda = \frac{\sigma_v}{\sigma_z} \) and \( g(y) \) is a increasing function in \( y \) since \( c > h > 0 \). Thus,

\[
P(t, \hat{Y}_t) = E \left[ g \left( \hat{Y}_t + \omega(nT) - \omega(t) \right) \right]
\]

\[
= E \left[ \exp \left( c\hat{x}_{nT - 1} + N(nT) + \lambda \left( \hat{Y}_t + Z_{nT} - Z_t - \frac{1}{\lambda} \frac{h}{c} \sigma^2_v (nT - t) \right) \right) \right]
\]

\[
= \exp \left( c\hat{x}_{nT - 1} + N(nT) + \lambda \hat{Y}_t + \frac{1}{2} \sigma^2_v (nT - t) - \frac{h}{c} \sigma^2_v (nT - t) \right)
\]

\[
= \exp \left( \log P_{nT - 1} + \lambda \hat{Y}_t - \left[ \frac{1}{2} \sigma^2_v - \frac{h}{c} \sigma^2_v \right] (t - (nT - 1)) \right) \quad (A32)
\]
where $P_{nT-1} = e^{c\hat{x}_{nT-1} - \frac{1}{2}\left(\frac{\hat{x}_n}{\hat{v}}\right)^2 + N(nT) + \frac{1}{2}(\frac{\hat{x}_n}{\hat{v}})^2}$, which is exactly equation (A22).

After I find the policy function $g(y)$, through equation (A29), it’s very straightforward to show the maximized expected profit of the insider is

$$J(t, P(t, \hat{Y}_t), A(\hat{x}_{nT}, nT)) = \frac{1}{2} \lambda \sigma_v \sigma_z (nT - t) A(\hat{x}_{nT}, nT) + \frac{1}{2} \sigma^2_z P_{yy}(t, \hat{Y}_t) - \frac{1}{2} \frac{1}{\lambda} \left[ \log A(\hat{x}_{nT}, nT) - \log \frac{A(t, \hat{Y}_t)}{P(t, \hat{Y}_t)} \right],$$

(A33)

**Lemma 6.** With respect to the market makers’ filtration, both pricing rule $P(t, \hat{Y}_t)$ and price-response function $P_y(t, \hat{Y}_t)$ are submartingales with a deterministic growth rate $\frac{h_c \sigma^2_v}{c} \sigma^2_v$ per unit of time.

**Proof.** As in Back (1992), I explicitly indicate the conditional expectation at time $t$ given the market makers’ information by $E^M[\cdot]$ and the conditional expectation given the informed trader’s information by $E^I[\cdot]$.

From equation (A28),

$$P(t, \omega(t)) = E^I[g(\omega(1))]; \quad P_y(t, \omega(t)) = E^I[g_y(\omega(1))],$$

which implies both $P(t, \omega(t))$ and $P_y(t, \omega(t))$ are martingales under the inside trader’s information set. However, this can not tell us the property of $P(t, \hat{Y}_t)$ under market makers’ information since $\hat{Y}_t$ with respect to market makers’ information is not the same as $\omega(t)$ with respect to the informed trader’s information from equation (A27). Thus, I study the property of $P(t, Z_t)$ under informed trader’s information.

It’s very straightforward to check this pricing rule satisfies

$$P(t, \hat{Y}_t) - \frac{1}{\lambda} \frac{h_c \sigma^2_v}{c} P_y(t, \hat{Y}_t) + \frac{1}{2} \sigma^2_z P_{yy}(t, \hat{Y}_t) = 0,$$

which is a natural implication of Lemma 4.
The pricing rule in equation (A28) yields

\[ P(t, Z_t) = E[I \left[ g(Z_t + \omega(nT) - \omega(t)) \right] | Z_t] \]

\[ \stackrel{(A31)}{=} E^I \left[ \exp \left( c\hat{x}_{nT-1} + N(nT) + \lambda \left( Z_{nT} - \frac{1}{\lambda} h \sigma_v^2(nT - t) \right) \right) \right] | Z_t] \]

\[ = E^I \left[ P(nT, Z_{nT}) \right] \exp \left( -\frac{h}{c} \sigma_v^2(nT - t) \right) \tag{A34} \]

where the last equality comes from equation (A32):

\[ P(nT, Z_{nT}) = \exp \left( \log P_{nT-1} + \lambda Z_{nT} - \left[ \frac{1}{2} \sigma_v^2 - \frac{h}{c} \sigma_v^2 \right] \right) \]

\[ = \exp \left( c\hat{x}_{nT-1} + N(nT) + \lambda Z_{nT} \right). \]

Rearrange equation (A34), I find

\[ P(t, Z_t) \exp \left( -\frac{h}{c} \sigma_v^2(t - (nT - 1)) \right) = E^I \left[ P(nT, Z_{nT}) \right] \exp \left( -\frac{h}{c} \sigma_v^2 \right), \]

which implies \( P(t, Z_t) \exp \left( -\frac{hc}{(c-h)^2} \sigma_v^2(t - (nT - 1)) \right) \) is a martingale under inside trader’s information set. Since the distribution of \( Z_t \) with respect to the informed trader’s information is the same as the distribution of \( \hat{Y}_t \) with respect to market makers’ information,

\[ P(t, \hat{Y}_t) \exp \left( -\frac{h}{c} \sigma_v^2(t - (nT - 1)) \right) = E^M \left[ P(nT, \hat{Y}_{nT}) \right] \exp \left( -\frac{h}{c} \sigma_v^2 \right) \]

\[ = E^M \left[ P(nT, \hat{Y}_{nT}) \right] \exp \left( -\frac{h}{c} \sigma_v^2 \right) \]

where the last equality using the Markov property of a Brownian motion. This implies

\[ P(t, \hat{Y}_t) \exp \left( -\frac{h}{c} \sigma_v^2(t - (nT - 1)) \right) \]

is a martingale under inside trader’s information set. This is equivalent to say \( P(t, \hat{Y}_t) \) is a submartingale with a deterministic growth rate \( \frac{h}{c} \sigma_v^2 \) per unit of time since both \( h \) and \( c \) are
strictly positive. Similar argument applies to the price-response function $P_y(t, \hat{Y}_t)$, which is also a submartingale with a deterministic growth rate $\frac{h}{c}\sigma^2_v$ per unit of time.

From Lemma 6, it’s very straightforward to show the unconditional expected return for any $t \in [nT - 1, nT]$ is

$$\log E \left[ \frac{P_t}{P_{nT-1}} | \mathcal{F}_{nT-1}^Y \right] = \frac{h}{c}\sigma^2_v (t - (nT - 1)) = hc [Q_{nT-1} - Q_{nT}] (t - (nT - 1)), \quad (A35)$$

which implies the pre-FOMC announcement drift grows at a constant rate $\frac{h}{c}\sigma^2_v$. In the meantime, the posterior variance of $\ln P_{nT}$ at time $t \in [nT - 1, nT]$ is

$$Var \left[ \ln P_{nT} | \mathcal{F}_t^Y \right] = Var \left[ \ln A (\hat{x}_{nT}, nT) | \mathcal{F}_t^Y \right]$$

$$= Var \left\{ \ln [H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT)] - \ln [H (\hat{x}_{nT}, nT)] | \mathcal{F}_t^Y \right\}$$

$$= Var \left\{ \ln [H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT)] | \mathcal{F}_t^Y \right\} + Var \left\{ \ln [H (\hat{x}_{nT}, nT)] | \mathcal{F}_t^Y \right\}$$

$$- 2 Cov \left( \ln [H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT)], \ln [H (\hat{x}_{nT}, nT)] \right)$$

$$= \left( \frac{c - h}{c} \right)^2 \sigma^2_v (nT - t) - 2 \frac{-h (c - h)}{c^2} \sigma^2_v (nT - t) + \left( \frac{h}{c} \right)^2 \sigma^2_v (nT - t)$$

$$= \sigma^2_v (nT - t) = c^2 [Q_{nT-1} - Q_{nT}] (nT - t).$$

Therefore, the reduction of the uncertainty at time $t$ comparing to $nT - 1$ is

$$Var \left[ \ln P_{nT} | \mathcal{F}_t^Y \right] - Var \left[ \ln P_{nT} | \mathcal{F}_{nT-1}^Y \right] = c^2 [Q_{nT-1} - Q_{nT}] [ (nT - 1) - t],$$

which implies the uncertainty reduces at a constant rate $c^2 [Q_{nT-1} - Q_{nT}]$ per unit of time. This completes the proof of proposition 1.
Appendix B  Proof of Theorem 2

First, I show that given \( \theta (s, V) \) and \( \theta (s, \bar{V}) \), how the market makers estimate the probability that the inside trader has private information and the value of \( H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \) conditional on \( \delta = 1 \) through the nonlinear filtering. Second, I show the price rule is a function of the adjusted order flow \( \hat{Y}_t \). Third, I construct the equilibrium and establish the optimality through a verification proof.\(^{25}\)

B.1 Step 1: Market makers’ updating

Lemma 7. Let \( \mu (t, V) \) be the estimate of the unnormalized density function of the random variable \( \hat{V} = H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \) given the stochastic differential equation (26) when the insider is informed. Then \( \mu (t, V) \) must satisfy the following stochastic differential equation (Zakai equation):

\[
d\mu (t, V) = \frac{\theta (t, V)}{\sigma^2} \mu (t, V) dY (t), \quad \mu (0, V) = f (V),
\]

which has a unique solution

\[
\mu (t, V) = f (V) \exp \left[ \frac{1}{\sigma^2} \left( \int_0^t \theta (s, V) dY (s) - \frac{1}{2} \int_0^t \theta^2 (s, V) ds \right) \right].
\]

Hence, the value estimate \( V (t) \) of \( H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \) is given by

\[
V (t) \equiv E [H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) | \mathcal{F}_1 (t)] = \frac{\int_V V \mu (V, t) dV}{\int_V \mu (V, t) dV} \quad (A36)
\]

where \( f (v) = \frac{df (v)}{dv} \) is the prior probability density function at time 0.

Proof. See Lemma 1 in Li (2013).

By Ito’s Lemma, I can derive the estimation of \( H (\hat{x}_{nT}, nT) A (\hat{x}_{nT}, nT) \) with respect to market makers’ perceptive conditional on \( \delta = 1 \).

\(^{25}\)The method of proof is largely based on Li (2013), who applies the “sequential detection” in the filtering literature.
Lemma 8. The value estimate given by (A36) satisfies the stochastic differential equation
\[ dV(t) = \lambda(t) \left( dY(t) - \hat{\theta}(t) dt \right), \quad (A37) \]
where
\[ \hat{\theta}(t) = E \left[ \theta(t, \tilde{V}) | \mathcal{F}_1(t) \right] = \frac{\int_V \theta(t, V) \mu(V, t) \, dv}{\int_V \mu(V, t) \, dv} \quad (A38) \]
and
\[ \lambda(t) \equiv \frac{E[\theta(t, \tilde{v}) | \mathcal{F}_1(t)] - V(t) \hat{\theta}(t)}{\sigma^2 z}. \quad (A39) \]

In addition,
\[ \hat{Y}_1(t) \equiv Y(t) - \int_0^t \hat{\theta}(s) \, ds \quad (A40) \]
is a Brownian Motion with instant variance $\sigma^2 z$ under $\mathcal{F}_1(t)$.

Therefore, the market makers estimate the probability that the cumulative order flow is generated by the inside trader has private information or not. This updating problem can be solved as to calculate the likelihood ratio between the two hypotheses, $\delta = 1$ versus $\delta = 0$.

Lemma 9. The logarithm of the likelihood ratio between hypotheses in equation (26) is given by
\[ \phi(t) \equiv \frac{1}{\sigma^2 z} \left( \int_0^t \left[ \hat{\theta}(s) - \theta(s, \tilde{V}) \right] dY(s) - \frac{1}{2} \int_0^t \left[ \hat{\theta}^2(s) - \theta^2(s, \tilde{V}) \right] ds \right) \]
where $\hat{\theta}$ is as defined by (A38).

Lemma 10. The market makers’ estimate of the probability that the strategic trader has private information
\[ \pi(t) = E[\delta | \mathcal{F}(t)] = \frac{\pi_0 \exp[\phi(t)]}{1 - \pi_0 + \pi_0 \exp[\phi(t)]} \quad (A41) \]
satisfies the following stochastic differential equation:
\[ d\pi(t) = \frac{\pi(t) [1 - \pi(t)]}{\sigma^2 z} \left( \hat{\theta}(t) - \theta(t, \tilde{V}) \right) d\hat{Y}(t), \quad \pi(0) = \pi_0 \quad (A42) \]
where
\[
\dot{Y}(t) = Y(t) - \int_0^t \left( \pi(s) \dot{\theta}(s) + [1 - \pi(s)] \theta(s, \bar{V}) \right) ds
\]  
(A43)

is the information process, which is a Brownian motion with instantaneous variance \( \sigma_z^2 \) under the filtration \( F(t) \).

Proof. The definition of \( \pi(t) \) in equation (A41) is obtained by the Bayes’ rule. By Ito’s Lemma,
\[
d\pi(t) = \pi(1 - \pi) d\phi + \frac{1}{2} \pi(1 - 2\pi)(d\phi)^2
\]
\[
= \frac{\pi(t)}{\sigma_z^2} \left( \dot{\theta}(t) - \theta(t, \bar{V}) \right) d\dot{Y}(t)
\]
where
\[
\dot{Y}(t) = Y(t) - \int_0^t \left( \pi(s) \dot{\theta}(s) + [1 - \pi(s)] \theta(s, \bar{V}) \right) ds.
\]

Let \( \Pi(t,y) \) be an arbitrary function in \( C^{1,2} \) on \([0,1] \times \mathbb{R} \) with a close range \([0,1] \). Since \( \ln[H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT)] \) is normally distributed with mean \( \mu_v \) and variance \( \left( \frac{c-h}{c} \right)^2 \sigma_v^2 \) at time \( nT-1 \), I define \( h(y) = \exp(\mu_v + \frac{c-h}{c} \lambda y) \) and \( \bar{V} = e^{\mu_v + \frac{1}{2} \left( \frac{c-h}{c} \right)^2 \sigma_v^2} \). This implies \( h^{-1}(\bar{V}) = \frac{\bar{V} - \sigma_v^2}{2\lambda} \).

I guess the insider’s trading strategy follows
\[
\theta(t, y, V) = \frac{h^{-1}(V) - h^{-1}(\bar{V}) - \Pi(t, y) \left[ y - h^{-1}(\bar{V}) \right]}{nT - t} + \bar{\theta}(t, y), \quad (A44)
\]
and
\[
\Theta(t, y) = \frac{1 - \Pi(t, y) \left[ y - h^{-1}(\bar{V}) \right]}{nT - t} + \bar{\theta}(t, y). \quad (A45)
\]

\( \bar{\theta}(t, y) \) is the expected trading volume of the inside trader under market makers’ perspective, which is not a function of \( V \) and will be defined later.\(^{26}\) Before I show that \( \theta(t, y, V) \) forms an

\(^{26}\)Li (2013) is a special case of this economy where \( \bar{\theta}(t, y) = 0 \) under the assumption that the market maker is risk neutral.
equilibrium, let me first characterize the market makers’ conditional value estimate, given the insider’s order rate $\theta(t, y, V)$.

**Lemma 11.** Let $\hat{Y}_1(t)$ be a Brownian bridge that satisfies

$$d\hat{Y}_1(t) = \left[ \theta(t, \hat{Y}_1(t), V) - \Theta(t, \hat{Y}_1(t)) \right] dt + dZ(t)$$

(A46)

with $\hat{Y}_1(nT - 1) = 0$. If the informed trader’s order rate is $\theta(t, \hat{Y}_1(t), V)$, as defined by (A44), $\Theta(t, \hat{Y}_1(t))$ as defined by (A45), is then the market makers’ expected order rate from the insider, conditional on the insider having private information. That is,

$$\hat{\theta}(t) = \mathbb{E} \left[ \theta(t, \hat{Y}_1(t), V) \mid \mathcal{F}_1(t) \right] = \Theta(t, \hat{Y}_1(t)).$$

Furthermore, the expected value of $H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT)$ under $\mathcal{F}_1(t)$ is

$$V(t) = H(t, \hat{Y}_1(t)),$$

where

$$H(t, y) = \mathbb{E} \left[ h(y + Z(1) - Z(t)) \right].$$

The insider’s trading strategy given in the above Lemma implies that when she is not better informed,

$$\theta(t, y, \bar{V}) = -\Pi(t, y) \frac{y - h^{-1}(\bar{V})}{nT - t} + \hat{\theta}(t, y),$$

which implies

$$\hat{\theta}(t) - \theta(t, y, \bar{V}) = \Theta(t, y) - \theta(t, y, \bar{V}) = \frac{y - h^{-1}(\bar{V})}{nT - t},$$

(A47)

and

$$\Pi(t, y) \Theta(t, y) + [1 - \Pi(t, y)] \theta(t, y, \bar{V}) = \hat{\theta}(t, y).$$

(A48)
From equation (A42), the market makers’ estimate of the probability that the strategic trader has private information satisfies

\[
d\Pi\left(t, \hat{Y}_1(t)\right) = \frac{\Pi\left(t, \hat{Y}_1(t)\right) \left[1 - \Pi\left(t, \hat{Y}_1(t)\right)\right] \hat{Y}_1(t) - h^{-1}(\hat{V})}{\sigma^2_z} \times \left(\frac{1}{nT - t} \left[\hat{Y}_1(t) - h^{-1}(\hat{V})\right] dt\right)_{A49}
\]

with \(\Pi\left(nT - 1, \hat{Y}_1(nT - 1)\right) = \pi_0\). This is true because

\[
dY(t)_{A43, A48} = d\hat{Y}(t) + \theta(t, y)
\]

\[
= d\hat{Y}_1(t) + \Theta(t, y)_{A40}
\]

\[
d\hat{Y}_1(t) + \frac{1 - \Pi\left(t, \hat{Y}_1(t)\right)}{nT - t} \left[\hat{Y}_1(t) - h^{-1}(\hat{V})\right] + \theta(t, y)_{A45}.
\]

As in Li (2013), given the trading strategy described above, I show that solution to the stochastic differential equation (A49) is:

\[
\Pi(t, y) = \frac{\pi_0 \exp\left(\frac{1}{2\sigma^2} \left[\frac{y - h^{-1}(\hat{V})}{nT - t}\right]^2 + \frac{1}{2} \ln(nT - t) - \frac{[h^{-1}(\hat{V})]^2}{2\sigma^2}\right)}{1 - \pi_0 + \pi_0 \exp\left(\frac{1}{2\sigma^2} \left[\frac{y - h^{-1}(\hat{V})}{nT - t}\right]^2 + \frac{1}{2} \ln(nT - t) - \frac{[h^{-1}(\hat{V})]^2}{2\sigma^2}\right)},
\]

which is the optimal probability estimate of market makers.

When the insider does not have private information, from equation (A47), I can rewrite the dynamics of the probability estimate as

\[
d\Pi\left(t, \hat{Y}_1(t)\right) = \frac{\Pi\left(t, \hat{Y}_1(t)\right) \left[1 - \Pi\left(t, \hat{Y}_1(t)\right)\right] \hat{Y}_1(t) - h^{-1}(\hat{V})}{\sigma^2_z} \times \left(\frac{1}{nT - t} \left[\hat{Y}_1(t) - h^{-1}(\hat{V})\right] dt\right)_{A50}
\]

50
Thus, there are two different dynamics of the probability estimate conditional on whether the insider is informed. As shown in Li (2013), the probability estimate always resides in $(0, 1)$ for all $t < nT$ and it converges to 1 or 0 upon announcements depending on whether the insider has private information or not.

Since $\ln \left[ H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT) \right]$ is normally distributed with mean $\mu_v$ and variance $\left( \frac{c-h}{c} \right)^2 \sigma_v^2$ at time $nT - 1$, $h(y) = \exp \left( \mu_v + \frac{c-h}{c} \lambda y \right)$ and $V = e^{\mu_v + \frac{1}{2} \left( \frac{c-h}{c} \right)^2 \sigma_v^2}$. From Lemma 11, the estimation of $H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT)$ conditional on $\delta = 1$ follows

$$V(t) = E[H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT) | F_1(t)]$$

$$= H(t, \hat{Y}_1(t)) = E \left[ h(\hat{Y}_1(t) + Z(nT) - Z(t)) \right]$$

$$= \exp \left( \mu_v + \frac{c-h}{c} \lambda \hat{Y}_1(t) + \frac{1}{2} \left( \frac{c-h}{c} \right)^2 \sigma_v^2 (nT - t) \right),$$

while the estimation of $H(\hat{x}_{nT}, nT) A(\hat{x}_{nT}, nT)$ conditional on $\delta = 0$ is $\bar{V} = e^{\mu_v + \frac{1}{2} \left( \frac{c-h}{c} \right)^2 \sigma_v^2}$, where $\mu_v = (c-h) x_{nT-1} + N(nT)$.

Similarly, the estimation of SDF $H(\hat{x}_{nT}, nT)$ conditional on $\delta = 1$ follows

$$\Lambda(t) = E[H(\hat{x}_{nT}, nT) | F_1(t)]$$

$$= \exp \left( \mu_\Lambda - \frac{h}{c} \lambda \hat{Y}_1(t) + \frac{1}{2} \left( \frac{h}{c} \right)^2 \sigma_v^2 (nT - t) \right),$$

while the estimation of SDF $H(\hat{x}_{nT}, nT)$ conditional on $\delta = 0$ is $\bar{\Lambda} = e^{\mu_\Lambda + \frac{1}{2} \left( \frac{h}{c} \right)^2 \sigma_v^2}$, where $\mu_\Lambda = -hx_{nT-1} + \mathcal{H}(nT)$.

Therefore, this implies the price defined in equation (24) depends only on the current ad-
justed trading flow \( \hat{Y}_1(t) \), which follows

\[
P\left(t, \hat{Y}_1(t)\right) = \frac{\Pi\left(t, \hat{Y}_1(t)\right) V\left(t, \hat{Y}_1(t)\right) + \left(1 - \Pi\left(t, \hat{Y}_1(t)\right)\right) \bar{V}}{\Pi\left(t, \hat{Y}_1(t)\right) \Lambda\left(t, \hat{Y}_1(t)\right) + \left(1 - \Pi\left(t, \hat{Y}_1(t)\right)\right) \bar{\Lambda}}
\]

\[
= \frac{\Pi\left(t, \hat{Y}_1(t)\right) e^{\mu v + \frac{\sigma^2}{2} \lambda \hat{Y}_1(t)} + \frac{1}{2} \left(\frac{\sigma^2}{\epsilon}\right)^2 (\epsilon - c_2) (nT - t) + \left(1 - \Pi\left(t, \hat{Y}_1(t)\right)\right) e^{\mu v + \frac{1}{2} \left(\frac{\sigma^2}{\epsilon}\right)^2 c_2}}{\Pi\left(t, \hat{Y}_1(t)\right) e^{\mu v + \frac{1}{2} \left(\frac{\sigma^2}{\epsilon}\right)^2 c_2}}
\]

\[
= P_{nT-1}\left(t, \hat{Y}_1(t)\right) e^{-\frac{\sigma^2}{2} \lambda \hat{Y}_1(t)} - \frac{1}{2} \left(\frac{\sigma^2}{\epsilon}\right)^2 (\epsilon - c_2) (nT - t) + 1 - \Pi\left(t, \hat{Y}_1(t)\right)
\]

\[
\cdot \frac{1}{1 - \Pi\left(t, \hat{Y}_1(t)\right)} \left(1 - \Pi\left(t, \hat{Y}_1(t)\right)\right), \quad (A51)
\]

where \( P_{nT-1} = e^{c_{x_nT-1} - \frac{1}{2} \left(\frac{\sigma^2}{\epsilon}\right)^2 \sigma_2^2 + N(nT) + \frac{1}{2} \left(\frac{\sigma^2}{\epsilon}\right)^2 \sigma_2^2} \).

**B.2 Step 2: Insider’s Optimal Strategy**

In this section, I show that if the dynamics of price follows equation (A51) where \( \), then the optimal trading strategy of the insider is indeed of the form given in equation (A44) through verification proof.

Given the market makers’ pricing rule, \( P(t) = P\left(t, \hat{Y}_1(t)\right) \), the informed trader chooses the order rate to maximize her trading profit. That is,

Recall that the informed trader chooses the order rate to maximize her trading profit given the make makers’ pricing rule \( P(t) = P\left(t, \hat{Y}_1(t)\right) \):

\[
J(t, y, A(\hat{x}_nT, nT), \pi_0) = \max_{\theta_t \in A} \left[ \int_t^{nT} \left( A(\hat{x}_nT, nT) - P(s, \hat{Y}_s) \right) \theta_s ds | \hat{Y}_t = y, A(\hat{x}_nT, nT) \right]
\]

subject to

\[
d\hat{Y}_t = \left[ \theta(t) - \hat{\theta}(t) \right] dt + dZ_t. \quad (A52)
\]

The principle of optimality implies the following Bellman equation

\[
\max_{\theta_t \in A} \left\{ (A(\hat{x}_nT, nT) - P(t, y)) \theta_t + J_t + J_y \left[ \theta_t - \hat{\theta}(t) \right] + \frac{1}{2} \sigma^2 J_{yy} \right\} = 0 \quad (A53)
\]

52
where the subscripts denote the derivatives. The necessary conditions for having an optimal solution to the Bellman equation (A53) are

\[ J_y(t, y, A(\hat{x}_{nT}, nT)) = P(t, y) - A(\hat{x}_{nT}, nT) \] (A54)

\[ J_t + \frac{1}{2} \sigma^2 J_{yy} - \hat{\theta}(t) J_y = 0. \] (A55)

These necessary conditions lead to the following results:

**Lemma 12.** Suppose the expected order rate \( \hat{\theta}(t) = \Theta(t, \hat{Y}_1(t)) \), where \( \hat{Y}_1(t) \) is the adjusted order at \( t \). Let \( \omega(t) = y \) and suppose that the stochastic differential equation

\[ d\omega(s) = dZ(s) - \Theta(s, \omega(s)) ds, \quad \forall 1 \geq s \geq t \]

has a unique solution, where \( Z(s) \) is a Brownian motion with instant variance \( \sigma^2 \). If there exists a strictly monotone function \( g(\cdot) \) such that the pricing rule is

\[ P(t, y) = E[g(\omega(1)) | \omega(t) = y], \] (A56)

then

\[ J(t, y; v, \pi_0) = E[j(v, \omega(1)) | \omega(t) = y] \]

is a smooth solution to the Bellman equations (18) and (19), where

\[ j(v, y) = \int_y^{g^{-1}(v)} [v - g(x)] dx \geq 0, \quad \forall (v, y) \]

Before I prove the above Lemma, I would like to determine the expected trading volume of the insider \( \bar{\theta}(t, y) \) in equation (A44). The equation (A56) implies that \( P(t, \omega(t)) \) is a martingale under the filtration generated by \( \omega \). This implies

\[ P_t - \hat{\theta}(t) P_y + \frac{1}{2} \sigma^2 P_{yy} = 0. \]

So I want to find the expected trading volume \( \bar{\theta} \), which is a function of \( (t, \hat{Y}_1(t)) \) and makes

\[ P_t - \Theta(t, \hat{Y}_1(t)) P_y + \frac{1}{2} \sigma^2 P_{yy} = 0, \] (A57)

53
for any \( t \) and \( \hat{Y}_1(t) \).

As \( \Pi \left( t, \hat{Y}_1(t) \right) \) is a martingale under the market makers’ information set, \(^{27}\)

\[
\begin{align*}
d\hat{Y}_1(t) &= dY(t) - \Theta \left( t, \hat{Y}_1(t) \right) dt \\
&= \left[ d\hat{Y}(t) + \bar{\theta} \left( t, \hat{Y}_1(t) \right) dt \right] - \Theta \left( t, \hat{Y}_1(t) \right) dt \\
&= d\hat{Y}(t) - \left[ \Theta \left( t, \hat{Y}_1(t) \right) - \bar{\theta} \left( t, \hat{Y}_1(t) \right) \right] dt
\end{align*}
\]

and the fact that \( \hat{Y}(t) \) is a Brownian motion, I have

\[
\Pi_t - \left[ \Theta \left( t, \hat{Y}_1(t) \right) - \bar{\theta} \left( t, \hat{Y}_1(t) \right) \right] \Pi_y + \frac{1}{2} \sigma^2 \Pi_{yy} = 0 \tag{A58}
\]

Moreover, as

\[
\Pi_y = \Pi \left( 1 - \Pi \right) \frac{\hat{Y}_1(t) - h^{-1}(\bar{v})}{\sigma^2 (nT - t)},
\]

I have

\[
\sigma^2 \Pi_y - \left[ \Theta \left( t, \hat{Y}_1(t) \right) - \bar{\theta} \left( t, \hat{Y}_1(t) \right) \right] \Pi = 0 \tag{A59}
\]

\(^{27}\)To show this

\[
\begin{align*}
d\Pi \left( t, \hat{Y}_1(t) \right) &= \frac{\Pi \left( t, \hat{Y}_1(t) \right) \left[ 1 - \Pi \left( t, \hat{Y}_1(t) \right) \right] \hat{Y}_1(t) - h^{-1}(\bar{v})}{nT - t} d\hat{Y}(t) \\
&= \frac{\Pi \left( t, \hat{Y}_1(t) \right) \left[ 1 - \Pi \left( t, \hat{Y}_1(t) \right) \right] \hat{Y}_1(t) - h^{-1}(\bar{v})}{nT - t} \left[ dY(t) - \bar{\theta} \left( t, \hat{Y}_1(t) \right) dt \right] \\
&= \frac{\Pi \left( t, \hat{Y}_1(t) \right) \left[ 1 - \Pi \left( t, \hat{Y}_1(t) \right) \right] \hat{Y}_1(t) - h^{-1}(\bar{v})}{nT - t} \left( d\hat{Y}(t) + \frac{1 - \Pi \left( t, \hat{Y}_1(t) \right)}{nT - t} \left[ \hat{Y}_1(t) - h^{-1}(\bar{v}) \right] dt \right)
\end{align*}
\]
\[ P\left(t, \hat{Y}_1 (t)\right) = \frac{\Pi \left(t, \hat{Y}_1 (t)\right) e^{\frac{c-h}{c} \lambda \hat{Y}_1 (t) - \frac{1}{2} \left(\frac{c-h}{c}\right)^2 \sigma^2 (t-(nT-1))} + 1 - \Pi \left(t, \hat{Y}_1 (t)\right)}{\Pi \left(t, \hat{Y}_1 (t)\right) e^{-\frac{h}{c} \lambda \hat{Y}_1 (t) - \frac{1}{2} \left(\frac{h}{c}\right)^2 \sigma^2 (t-(nT-1))} + 1 - \Pi \left(t, \hat{Y}_1 (t)\right)} = \frac{A \left(t, \hat{Y}_1 (t)\right)}{B \left(t, \hat{Y}_1 (t)\right)} \]

where \(A \left(t, \hat{Y}_1 (t)\right) = \Pi \left(t, \hat{Y}_1 (t)\right) e^{\frac{c-h}{c} \lambda \hat{Y}_1 (t) - \frac{1}{2} \left(\frac{c-h}{c}\right)^2 \sigma^2 (t-(nT-1))} + 1 - \Pi \left(t, \hat{Y}_1 (t)\right)\), and \(B \left(t, \hat{Y}_1 (t)\right) = \Pi \left(t, \hat{Y}_1 (t)\right) e^{-\frac{h}{c} \lambda \hat{Y}_1 (t) - \frac{1}{2} \left(\frac{h}{c}\right)^2 \sigma^2 (t-(nT-1))} + 1 - \Pi \left(t, \hat{Y}_1 (t)\right)\). To save notations, let \(D \left(t, \hat{Y}_1 (t)\right) = e^{\frac{c-h}{c} \lambda \hat{Y}_1 (t) - \frac{1}{2} \left(\frac{c-h}{c}\right)^2 \sigma^2 (t-(nT-1))}\) and \(E \left(t, \hat{Y}_1 (t)\right) = e^{-\frac{h}{c} \lambda \hat{Y}_1 (t) - \frac{1}{2} \left(\frac{h}{c}\right)^2 \sigma^2 (t-(nT-1))}\).

I find

\[ P_t = P_{nT-1} B^{-2} \left\{ \left( \Pi_t - \frac{1}{2} \left(\frac{c-h}{c}\right)^2 \sigma^2 \Pi \right) D - \Pi_t \right\} B - A \left\{ \left( \Pi_t - \frac{1}{2} \left(\frac{h}{c}\right)^2 \sigma^2 \Pi \right) E - \Pi_t \right\} \]

\[ P_y = P_{nT-1} B^{-2} \left\{ \left( \Pi_y + \frac{c-h}{c} \lambda \Pi \right) D - \Pi_y \right\} B - A \left\{ \left( \Pi_y - \frac{h}{c} \lambda \Pi \right) E - \Pi_y \right\} \]

\[ P_{gy} = P_{nT-1} B^{-2} \left\{ \left( \Pi_{gy} + 2 \frac{c-h}{c} \lambda \Pi_y + \left(\frac{c-h}{c}\right)^2 \lambda^2 \Pi \right) D - \Pi_{gy} \right\} B - A \left\{ \left( \Pi_{gy} - 2 \frac{h}{c} \lambda \Pi_y + \left(\frac{h}{c}\right)^2 \lambda^2 \Pi \right) E - \Pi_{gy} \right\} \]

\[ -P_{nT-1} B^{-2} \cdot 2B^{-1} \left\{ \left( \Pi_y + \frac{c-h}{c} \lambda \Pi \right) D - \Pi_y \right\} B - A \left\{ \left( \Pi_y - \frac{h}{c} \lambda \Pi \right) E - \Pi_y \right\} \left\{ \left( \Pi_y - \frac{h}{c} \lambda \Pi \right) E - \Pi_y \right\} \]
Put these derivatives into the following equation:

\[
P_t - (\Theta - \bar{\theta} + \bar{\theta}) P_y + \frac{1}{2} \sigma_z^2 P_{yy} = \left[ \left( \Pi_t - \frac{1}{2} \left( \frac{c - h}{c} \right)^2 \sigma_z^2 \right) D - \Pi_t \right] B - A \left[ \left( \Pi_t - \frac{1}{2} \left( \frac{h}{c} \right)^2 \sigma_z^2 \right) E - \Pi_t \right] - \\
(\Theta - \bar{\theta} + \bar{\theta}) \left[ \left( \Pi_y + \frac{c - h}{c} \lambda \Pi \right) D - \Pi_y \right] B - A \left[ \left( \Pi_y - \frac{h}{c} \lambda \Pi \right) E - \Pi_y \right] + \\
\frac{1}{2} \sigma_z^2 \left\{ \left[ \left( \Pi_{yy} + 2 \frac{c - h}{c} \lambda \Pi y + \left( \frac{c - h}{c} \right)^2 \lambda^2 \Pi \right) D - \Pi_{yy} \right] B - A \left[ \left( \Pi_{yy} - 2 \frac{h}{c} \lambda \Pi y + \left( \frac{h}{c} \right)^2 \lambda^2 \Pi \right) E - \Pi_{yy} \right] \right\} - B^{-1} \sigma_z^2 \left\{ \left[ \left( \Pi_y + \frac{c - h}{c} \lambda \Pi \right) D - \Pi_y \right] B - A \left[ \left( \Pi_y - \frac{h}{c} \lambda \Pi \right) E - \Pi_y \right] \right\} \left[ \left( \Pi_y - \frac{h}{c} \lambda \Pi \right) E - \Pi_y \right]
\]

where the first two terms are zero from equation (A58) and (A59). Therefore, equation (A57) holds if and only if

\[
\bar{\theta} \left( t, \hat{Y}_1 (t) \right) = B^{-1} \sigma_z^2 \left( \left( \Pi_y - \frac{h}{c} \lambda \Pi \right) E - \Pi_y \right)
\]

\[
= \frac{\frac{h}{c} \lambda \sigma_z^2 \Pi E - \Pi \left( 1 - \Pi \right) \hat{Y}_1 (t) - h^{-1} (i) \left[ nT - t \right]}{\Pi \cdot E + 1 - \Pi}
\]

\[
= \frac{\frac{1}{2} \frac{h}{c} \sigma_z^2 \Pi E - \Pi \left( 1 - \Pi \right) \hat{Y}_1 (t) - h^{-1} (i) \left[ nT - t \right]}{\Pi \cdot E + 1 - \Pi},
\]

where \( E \left( t, \hat{Y}_1 (t) \right) = e^{-\frac{\lambda}{2} Y_1(t) - \frac{1}{2} \left( \frac{h}{c} \right)^2 \sigma_z^2 (t - (nT - 1))} \) and

\[
\Pi \left( t, \hat{Y}_1 (t) \right) = \frac{\pi_0 \exp \left( \frac{1}{2 \sigma^2} \frac{Y_1 (t) - h^{-1} (i) \left[ nT - t \right]}{nT - t} + \frac{1}{2} \ln (nT - t) - \frac{\left[ h^{-1} (i) \right]^2}{2 \sigma^2} \right)}{1 - \pi_0 + \pi_0 \exp \left( \frac{1}{2 \sigma^2} \frac{Y_1 (t) - h^{-1} (i) \left[ nT - t \right]}{nT - t} + \frac{1}{2} \ln (nT - t) - \frac{\left[ h^{-1} (i) \right]^2}{2 \sigma^2} \right)}.
\]

56
It shows that the expected trading volume $\overline{\theta}$ is a function of $\left( t, \hat{Y}_1 (t) \right)$ since all the terms on the right side of the equation are functions of $\left( t, \hat{Y}_1 (t) \right)$.

From the above equation, it’s easy to see that when $\Pi \equiv 1$, it reduces to the baseline model where $\overline{\theta} \left( t, \hat{Y} (t) \right) = \frac{1}{\lambda c} \sigma^2$. When the market maker is risk neutral to future news, the expected trading volume $\overline{\theta} \left( t, \hat{Y}_1 (t) \right)$ is 0, which is consistent with Li (2013).
The model is calibrated at annually frequency. I assume the prescheduled announcements happen at the monthly frequency, that is, $T = \frac{1}{12}$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>symbol</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>aggregate output</td>
<td></td>
<td></td>
</tr>
<tr>
<td>long run output growth rate</td>
<td>$\bar{\theta}$</td>
<td>1.50%</td>
</tr>
<tr>
<td>volatility of output</td>
<td>$\sigma_y$</td>
<td>3%</td>
</tr>
<tr>
<td>persistence of the AR(1) process</td>
<td>$\rho$</td>
<td>4.5%</td>
</tr>
<tr>
<td>volatility of the AR(1) process</td>
<td>$\sigma_\theta$</td>
<td>0.25%</td>
</tr>
<tr>
<td>uncertainty and asset value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>the transparency of announcements</td>
<td>$\sigma_\mathcal{S}^2$</td>
<td>$8.5 \times 10^{-7}$</td>
</tr>
<tr>
<td>the asset value</td>
<td>$c$</td>
<td>3</td>
</tr>
<tr>
<td>preference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>risk aversion</td>
<td>$\gamma$</td>
<td>10</td>
</tr>
<tr>
<td>elasticity of intertemporal substitution</td>
<td>$\psi$</td>
<td>2</td>
</tr>
<tr>
<td>subjective discount factor</td>
<td>$\beta$</td>
<td>0.005</td>
</tr>
</tbody>
</table>
Table 2: **Summary Statistics on S&P500 Index Excess Returns and Changes in VIX.**

Note: This table reports summary statistics for the pre-announcement day 2 p.m (−1) – announcement (ann) and announcement – close changes in VIX (ΔVIX) and cumulative excess returns on the S&P500 (Cum. Return). The close time is 3:55 p.m. The samples are: (1) All FOMC announcements, (2 and 3) FOMC announcements sorted on uncertainty, which is first and third tertiles of changes in VIX (ΔVIX) between open and 2 p.m on pre-announcement dates, and (4 and 5) FOMC announcements with and without FOMC Press Conference. “Sharpe ratio” is the annualized Sharpe ratio on FOMC announcement returns. The sample period is from 1996:01 to 2019:11, and from 2011:04 for press conference sample. “No. of FOMC” is the number of FOMC in each subset. t-statistics for the mean are reported in parentheses. ***, **, and * represent 1%, 5%, and 10%.

<table>
<thead>
<tr>
<th></th>
<th>(1) All</th>
<th>Sort on Uncertainty</th>
<th>Press Conference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 p.m (−1)-ann</td>
<td>ann-close</td>
<td>No. of FOMC</td>
</tr>
<tr>
<td>DM (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 p.m (−1)-ann</td>
<td>-0.134***</td>
<td>-0.161***</td>
<td>187</td>
</tr>
<tr>
<td>(-2.672)</td>
<td></td>
<td>(-4.008)</td>
<td></td>
</tr>
<tr>
<td>2 p.m (−1)-ann</td>
<td>-0.734***</td>
<td>-0.135*</td>
<td>61</td>
</tr>
<tr>
<td>(-10.204)</td>
<td></td>
<td>(-1.833)</td>
<td></td>
</tr>
<tr>
<td>2 p.m (−1)-ann</td>
<td>0.385***</td>
<td>-0.283***</td>
<td>63</td>
</tr>
<tr>
<td>(4.470)</td>
<td></td>
<td>(-3.134)</td>
<td></td>
</tr>
<tr>
<td>ann-close</td>
<td>-0.269***</td>
<td>-0.060</td>
<td>34</td>
</tr>
<tr>
<td>(-5.067)</td>
<td></td>
<td>(-1.540)</td>
<td></td>
</tr>
<tr>
<td>ann-close</td>
<td>-0.011</td>
<td>-0.049</td>
<td>35</td>
</tr>
<tr>
<td>(-0.135)</td>
<td></td>
<td>(-0.400)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2 p.m (−1)-ann</th>
<th>ann-close</th>
<th>No. of FOMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cum. Return (%)</td>
<td>0.332***</td>
<td>-0.030</td>
<td>187</td>
</tr>
<tr>
<td>(5.636)</td>
<td>(-0.446)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 p.m (−1)-ann</td>
<td>0.925***</td>
<td>-0.076</td>
<td>61</td>
</tr>
<tr>
<td>(8.549)</td>
<td>(-0.571)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 p.m (−1)-ann</td>
<td>-0.145</td>
<td>0.063</td>
<td>63</td>
</tr>
<tr>
<td>(-1.556)</td>
<td>(0.603)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ann-close</td>
<td>0.254***</td>
<td>0.099</td>
<td>34</td>
</tr>
<tr>
<td>(3.314)</td>
<td>(1.090)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ann-close</td>
<td>0.014</td>
<td>-0.157</td>
<td>35</td>
</tr>
<tr>
<td>(0.158)</td>
<td>(-1.328)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

No. of FOMC: 187, 61, 63, 34, 35.
Table 3: **Returns on the S&P500 Index.**

Note: This table shows results for regressing the changes in VIX ($\Delta VIX$) on the cumulative excess returns on the S&P500 ($Cum.\ Return$), $Cum.\ Return_t = \alpha + \beta \Delta VIX_t + \varepsilon_t$ where both $\Delta VIX_t$ and $Cum.\ Return_t$ are calculated from 2 p.m on pre-announcement date to 2 p.m on announcement date windows, and $t$ represents each FOMC announcement. The samples are: (1) All FOMC announcements, (2 and 3) FOMC announcements sorted on uncertainty, which is first and third tertiles of changes in VIX ($\Delta VIX_{t-1}$) between open and 2 p.m on pre-announcement dates, and (4 and 5) FOMC announcements with and without FOMC Press Conference. The sample period is from 1996:01 to 2019:11, and from 2011:04 for press conference sample. “Obs.” and “No. of FOMC” are the number of observations and amount of FOMC in each subset, respectively. t-statistics are reported in parentheses. ***, **, and * represent 1%, 5%, and 10%.

<table>
<thead>
<tr>
<th></th>
<th>(1) All</th>
<th>Sort on Uncertainty</th>
<th>Press Conference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(2) High</td>
<td>(3) Low</td>
</tr>
<tr>
<td>$\Delta VIX$</td>
<td>-0.911***</td>
<td>-1.026***</td>
<td>-0.744***</td>
</tr>
<tr>
<td></td>
<td>(-16.847)</td>
<td>(-7.201)</td>
<td>(-7.463)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.201***</td>
<td>0.162</td>
<td>0.133</td>
</tr>
<tr>
<td></td>
<td>(5.337)</td>
<td>(1.232)</td>
<td>(1.719)</td>
</tr>
<tr>
<td>Obs.</td>
<td>187</td>
<td>61</td>
<td>63</td>
</tr>
<tr>
<td>No. of FOMC</td>
<td>187</td>
<td>61</td>
<td>63</td>
</tr>
</tbody>
</table>
Table 4: S&P500 Index Return Time-Series Regressions.

Note: This table reports results for regressions of the time-series of pre-FOMC announcement returns on various explanatory variables for the sample period 1996:01 to 2019:11. The dependent variable is a time-series of cumulative excess returns on the S&P500 from 2 p.m on days before announcement to 2 p.m on days of scheduled FOMC announcements. The first independent variable in Column (1) and (2) is Pre-FOMC dummy ($D_{\text{FOMC}}$), which is equal to one when a scheduled FOMC announcement has been released in the following 24-hour interval and zero otherwise. The second independent variable in Column (2) is the interaction of changes in VIX and Pre-FOMC dummy ($\Delta VIX \times D_{\text{FOMC}}$). “Sharpe ratio” is the annualized Sharpe ratio on FOMC announcement returns. “Obs.” and “No. of FOMC” are the number of observations and amount of FOMC in each subset, respectively. t-statistics are reported in parentheses. ***, **, and * represent 1%, 5%, and 10%.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{\text{FOMC}}$</td>
<td>0.314***</td>
<td>0.191**</td>
</tr>
<tr>
<td></td>
<td>(3.737)</td>
<td>(2.250)</td>
</tr>
<tr>
<td>$\Delta VIX \times D_{\text{FOMC}}$</td>
<td>-0.911***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-7.592)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.010</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>(0.671)</td>
<td>(0.674)</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>1.14</td>
<td>1.14</td>
</tr>
<tr>
<td>Obs.</td>
<td>5899</td>
<td>5899</td>
</tr>
<tr>
<td>No. of FOMC</td>
<td>187</td>
<td>187</td>
</tr>
</tbody>
</table>
References


Bilyi, M., 2018, “Pre-FOMC announcement drift and speculative trading.” .


Cocoma, P., 2018, “Explaining the pre-announcement drift,” Available at SSRN 3014299.


Dai, L., Y. Wang, and M. Yang, 2019, “Insider trading when there may not be an insider,” Available at SSRN 2720736.


Drechsler, I., A. Moreira, and A. Savov, 2018, “Liquidity creation as volatility risk,” Available at SSRN 3133291.


Laarits, T., 2019, “Pre-Announcement Risk,” Available at SSRN 3443886.

Li, T., 2013, “Insider trading with uncertain informed trading,” Available at SSRN 946324.


Tourre, F., 2017, “A macro-finance approach to sovereign debt spreads and returns,” *manuscript, University of Chicago*.


Ying, C., 2020, “Heterogeneous Beliefs and FOMC Announcements,” *Available at SSRN 3286682*. 