Online Appendix for: "How Does the Stock Market Absorb Shocks?"

Murray Z. Frank* Ali Sanati†

Abstract

This appendix contains supplementary material to the analysis in the paper. Section 1 presents market reaction to news shocks in Drugs industry. Section 2 explains the regression approach in details. Shock absorption patterns are studied using the regression approach in Section 3. Relative merits of the two methodologies are discussed in Section 4. The arbitrageurs' test using intermediary capital ratios are repeated using monthly data in Section 5. Details of the models used in the textual analysis are provided in Section 6. A stylized model to clarify the proposed perspective to understand the shocks is discussed in Section 7.

*Carlson School of Management, University of Minnesota, Minneapolis, MN 55455.
†Kogod School of Business, American University, Washington, DC 20016.

1 Response to News Shocks in Drugs Industry

In this part we replicate Table 3 of the main paper, which shows the market reaction to news shocks, for a sub-sample of firms in the Drugs industry (SIC 283). Intuitively, one would expect biotech and pharmaceutical stocks to be more news dependent because usually news about their products and patents have a large impact on their future earnings.

Table 1 presents the results. Interestingly, our results show 2 to 4 times stronger post-news reaction for this group of firms. However, post-shock return patterns are generally the same as the case of all news.

2 Alternative Methodology: the Regression Approach

A large body of the literature studying the stock market reaction to price shocks use the regression method. Some recent examples of this method are in studies by **?** and **?**. As we discussed before, we study the market reaction to both news and no-news price shocks. Empirically, we calculate the price shock at each day, for all firms in our sample. If there is a news story about a firm in a day, we flag that observation as a news shock, otherwise we call it a no-news shock. When both news and no-news shocks are present, there are two alternative ways of running regressions.

First, to run the following regression separately for news and no-news observations.

$$ACAR_{it+2,t+40} = \alpha_i + \beta_1 ARet_{it} + \beta_2 Size_{it} + \beta_3 B/M_{it} + \beta_4 MOM_{it} + \beta_5 VOL_{it} + \varepsilon_{it} \quad (1)$$

i and t are firm and time indices, respectively. The dependent variable, $ACAR_{it+2,t+40}$, is the average cumulative abnormal return on firm i's stock for a post-shock period from day $t+2$ to $t+40$. On the right-hand side, α_i is the firm fixed effect, $ARet_{it}$ is the abnormal return on day t, and the rest are the controls for firm characteristics that predict expected returns. The controls are standard, as in **?**. They include yearly measures of firm size ($Size_{it}$), book to market ratio (B/M_{it}), return momentum for the past 12 months excluding the most recent month (MOM_{it}), and average daily return volatility for the previous month (VOL_{it}).

The variable of interest in this regression is β_1. When we run this regression separately for no-news, all news, positive news or negative news observations, in each case, a positive β_1 shows drift in returns after the shock and a negative β_1 shows reversal. We interpret the drift pattern in post-shock returns as the under-reaction, and the reversal as the over-reaction at day 0. This becomes more intuitive where we study news shocks. Suppose there is a change in fundamental value of the firm with each news story. When the news becomes public (day 0), the stock market reacts to the new information and move the price in the appropriate direction. Now if the price change is exactly of a magnitude that leaves the price at the fundamental value (correct reaction), we expect to see no specific pattern for returns in the following days. Whereas, if the magnitude of the price change on day 0 is smaller than the correct reaction (under-reaction), we expect to see the price moving in the same direction (drift) in the following days. Reciprocally, if the magnitude of the price change at day 0 is larger than the correct reaction (over-reaction), market corrects itself in the following days and pushes back the price to the fundamental value, so reversal occurs.

There is ample evidence suggesting the differences between small and large firms in capital markets, regarding their information environment, limits to arbitrage, etc. Therefore, it is reasonable to investigate relation of post-news ACARs and day 0 abnormal returns for firms of different size. This can be captured by adding an interaction term of size and day 0 abnormal return to the regression. So, in addition to (1), we also estimate the following regression.

$$\begin{aligned} ACAR_{it+2,t+40} = &\alpha_i + \beta_1 ARet_{it} + \beta_2 ARet_{it} * Size_{it} + \beta_3 Size_{it} + \beta_4 B/M_{it} \\ &+ \beta_5 MOM_{it} + \beta_6 VOL_{it} + \varepsilon_{it} \end{aligned} \tag{2}$$

In this setting, β_1 shows the average effect for all stocks and β_2 shows the additional effect for large stocks.

The second approach to run regressions is to pool all news and no-news observations. By allowing a dummy variable to identify news observations, the average effect for all observations and also the marginal effect for news shocks are estimated in the following specification.

$$\begin{aligned} ACAR_{it+2,t+40} = & \alpha_i + \beta_1 ARet_{it} + \beta_2 News_{it} * ARet_{it} + \beta_3 News_{it} + \beta_4 Size_{it} \\ & + \beta_5 B/M_{it} + \beta_6 MOM_{it} + \beta_7 VOL_{it} + \varepsilon_{it} \end{aligned} \tag{3}$$

All variables are defined same as before. The only additional variable is $News_{it}$, a bi-variate dummy variable, which takes 0 if there is no news about firm i on day t and 1 otherwise. β_1 shows the drift or reversal pattern for no-news observations and β_1 and β_2 together show the patterns for news observations.

Furthermore, instead of a single dummy variable for news, we are able to separately estimate the patterns for positive and negative news by allowing separate dummies for each group.

$$\begin{aligned} ACAR_{it+2,t+40} = & \alpha_i + \beta_1 ARet_{it} + \beta_2 PosNews_{it} * ARet_{it} + \beta_3 PosNews_{it} \\ & + \beta_4 NegNews_{it} * ARet_{it} + \beta_5 NegNews_{it} + \beta_6 Size_{it} \\ & + \beta_7 B/M_{it} + \beta_8 MOM_{it} + \beta_9 VOL_{it} + \varepsilon_{it} \end{aligned} \tag{4}$$

Similar to the previous setting β_1 shows the average post-shock pattern for no news observations. Taking β_1 into account along with β_2 and β_3, one can infer the patterns for positive and negative news, respectively.

3 Shock Absorption Patterns Using the Regression Approach

Table 2 shows the results for specifications (1) and (2), i.e. studying shock absorption patterns separately for various groups of observations. The first two columns show the results if we pool all news observations, and columns 5-7 show the results for all, positive and negative no-news shocks. The negative estimated coefficient on the interaction of day 0 abnormal return ($ARet_0$) and size in column 1, and on $ARet_0$ in column 2, along with the consistently negative coefficient on $ARet_0$ in columns 5-7 for all no-news groups confirm results form the previous studies, i.e. smaller short term overreaction to news and larger short term overreaction to no-news (**?** and **?**). Results for the no-news case is extensively replicated and discussed in the literature, however we believe this is not the whole picture about the news observations.

On one hand, if we include the interaction term of $ARet_0$ and Size as in column 1 of table 2, the coefficient on $ARet_0$ changes its sign to positive and lose its significance. This means as a result of pooling all news observations together, for an average firm we estimate no link between $ARet_0$ and $ACAR_{2,40}$, and only observe a much lower overreaction for large firms, captured by the negative coefficient on the interaction term.

On the other hand, if we distinguish positive and negative news observations, columns 3 and 4 of table 2 respectively, we observe significantly different patterns. We label a news observation to be a positive (negative) news if the abnormal return on the news date is positive (negative). For negative news, for an average firm, we observe a significant drift pattern after the news date, due to the positive sign of the coefficient on $ARet_0$; however, for positive news there is reversal, although it is not statistically significant. This means for an average firm, the stock market under-reacts to bad news and over-reacts to good news. Notice that taking both β_1 and β_2 into account, results suggest that both patterns are alleviated for large firms. We elaborate this point and explain that this finding is consistent with the model predictions later when we discuss arbitrageurs' tests.

Table 3 shows the results for the pooling regressions. The results of the specification 3 and 4 are in the first two columns; column 3 and 4 show the results for running the same regressions using raw returns instead of abnormal returns. Due to the statistically significant negative coefficients shown in the first row (β_1) of the table, in all cases reversal patterns for no-news observations is identified. Running standard pooling regressions in column 1 or using raw returns in column 3, as in **?**, we get similar results as previously shown in the literature, i.e. large reversals after no-news shocks and smaller reversals following news shocks. However, in columns 2 and 4, when we distinguish positive and negative news we observe different patterns. In this setting, taking coefficient estimates on $ARet_{it}$ together with the signed interaction term $Pos/Neg\ News * ARet_{it}$ suggest post-news reversal for positive news and almost no link between post-news returns and day 0 returns for negative news. Nevertheless, an important point is the significant negative coefficients on dummy variables for news and positive/negative news.

The latter point together with inconsistent results from the other two regression settings using literally the same data set, suggest the inability of regression approach to show all the patterns, even though it is heavily used in the literature for this purpose. As we discussed in methodology section, a potential source of the problem might be due to the data characteristics. Although, there is a large variation in day 0 abnormal returns, we do not observe as much variation in the post-shock ACARs in each group of observations. In other words, post-news ACARs are fairly similar within each group of positive news and negative news, but different across groups. Hence, much of the action could be captured by the dummy variables for positive and negative news and there is not much variation left to be picked by the interaction term.

4 Comparing the Two Methodologies

Although we show the results using both methodologies, there are a number of advantages for the event studies approach. First and foremost, is that in the event studies approach we are able to test for significance of ACARs particularly in different periods and for various groups. This point becomes more prominent if we expect to see different shock absorption patterns under different situations, for instance for positive versus negative shocks. One might think that by adding proper control variables, e.g. in this case dummies for positive and negative shocks, regression coefficients could potentially capture the patterns. However, for the specific goal of our study, which is to emphasize the differences between shock absorption patterns of positive and negative news, regression setting is not able to depict a “complete" picture of the underlying story. One reason for this failure is due to some of the data characteristics. For instance, as a matter of fact the magnitude and sign of post-news ACARs for negative and positive news stories are very similar. Also, among observations of each group (positive or negative news), there is not much variation in post-news ACARs. Therefore, much of the action is captured by dummies for positive and negative news in the regressions and there is not much variety left to be captured by the interaction term of the dummy and day 0 abnormal return. Nevertheless, we know that the important phenomenon to look for is the link between day 0 abnormal returns and the post-shock ACARs in different situations and comparison of these links across situations. As we show in results, the task can be done more clearly using event studies.

Another benefit of the event study methodology is that we can easily sort our data on different characteristics into deciles or quintiles and compare the shock absorption patterns in different groups. This feature is in particular very useful in the last section of our study for testing hypotheses.

5 Monthly Intermediary Capital Ratio

In this section we repeat the arbitrageurs' test that uses intermediary capital ratio growth rate as a proxy for intensity of arbitrage activity in the market. **?** provide the data on intermediary capital ratio at both quarterly and monthly frequencies. To compute the equity capital ratio they need market value of equity and book value of debt for each intermediary. They get the data from CRSP-Compustat merged data set for US firms and from Datastream for non-US firms. In the quarterly version of the data, both parts are evaluated at the quarterly frequency. In the monthly version, market value of equity is evaluated at the monthly frequency but book value of debt is at the quarterly frequency.

Table 4 shows the results when we repeat the test with monthly data (comparable results using quarterly data are shown in Table 9 in the main paper). Results are qualitatively similar to what we found in the baseline version of the test. As we move from the first to the fifth quintile (low to high arbitrage activity), post-shock returns lose economic and statistic significance in response to both positive and negative shocks.

6 Textual Analysis: Model Details

In this section we present more details on the parameters and structure of the models used for categorizing the news sample.[1] Appendix A of the paper presents the four main steps

[1]Model explanations presented here are partly adopted from RapidMiner tutorials.

of the categorization process: (1) Creating a training sample; (2) Processing articles in the training sample; (3) Creating, training, and validating the categorization model using the training sample; (4) Using the model to categorize the original sample. In the third step, we compare the performance of three different models of categorization and pick Naive Bayes as the baseline model. Table A.1 of the paper shows the performance comparison.

The k-Nearest Neighbor (k-NN) algorithm is based on comparing an unknown unit of the sample–a news story, in our case–with the k training units which are the nearest neighbors of the unknown unit. The first step of the application of the k-NN algorithm on a new unit is to find the k closest training examples. "Closeness" is defined in terms of a distance in the n-dimensional space, defined by the n attributes in the training sample. In our case, the attributes are elements of the word vectors created from the training sample.

Different types of measures–such as, numerical, nominal, Bergmann Divergence, etc.– can be used. Also, different metrics, such as the Euclidean distance, can be used to calculate the distance between the unknown unit and the training examples. The structure and parameters of the k-NN models used in our study is shown in the table below.

Model:	k-NN
Structure:	k= 1 or 3 Measure types: Numerical measures Numerical measure: Cosine Similarity

The next operator that we use, learns a model by means of a feed-forward neural network trained by a back propagation algorithm. Back propagation algorithm is a supervised learning method which can be divided into two phases: propagation and weight update. The two phases are repeated until the performance of the network is good enough. In back propagation algorithms, the output values are compared with the correct answer to compute the value of some predefined error-function. By various techniques, the error is then fed back through the network. Using this information, the algorithm adjusts the weights of each connection in order to reduce the value of the error function by some small amount. After repeating this process for a sufficiently large number of training cycles, the network will usually converge to some state where the error of the calculations is small. In this case, one would say that the network has learned a certain target function.

We can define the structure of the neural network with the number of hidden layers. In our case, size of the hidden layer is calculated from the number of attributes of the input example set, which is known after creating word vectors of the training sample. The layer size will be set to (number of attributes + number of classes) / 2 + 1, in which, the attributes are elements of the word vectors created from the training sample and there are nine classes as shown in Table 4 of the paper.

Next we must specify the number of training cycles used for the neural network training. In back-propagation the output values are compared with the correct answer to compute the value of some predefined error-function. The error is then fed back through the network. Using this information, the algorithm adjusts the weights of each connection in order to reduce the value of the error function by some small amount. This process is repeated n number of times. We set n=250.

The learning rate, which determines how much we change the weights at each step, is set to 0.3. The momentum, which simply adds a fraction of the previous weight update to the current one and helps with preventing local maxima, is set to 0.2. Also, we set the program to shuffle the training sample before learning. Finally, the Neural Net operator uses a usual sigmoid function as the activation function. Therefore, the value range of the attributes are normalized to -1 and +1. The structure and parameters of the Neural Network model is summarized in the table below.

Model:	Neural Network
Structure:	Layers: 3 (input, hidden, output) Size of hidden layer = (# attributes + # classes) / 2 + 1 Training cycles = 250 Error epsilon = 1.0e-5 Learning rate = 0.3 Momentum = 0.2 Shuffle training sample Normalize the value range of attributes

7 Model

In this section we present a stylized model that helps clarify how interactions of different groups of investors in response to shocks, generates predictable post-shock returns. At the end we compare model implications to the empirical findings in previous sections.

Our model builds on the market structure of **??**. We use this framework because it is a simple model in which stock prices reflect the interactions among various types of agents. In the model, each day agents submit their demand for a firm's stock, and there is also a unit supply of the stock. At the end of each day market clears and the price is determined. Given this structure, it is easy to implement financial frictions pertaining to each agent as a direct change in their demand functions. Although the market structure in our model resembles that of Shleifer and Vishny, the market participants, their characteristics, and also the frictions in our model are different. In our model, we assume there are three types of traders: retail investors, arbitrageurs, and liquidity traders.

Retail investors have attention bias, in the sense that they do not trade unless their attention is grabbed. They trade only in response to positive news and they may not fully internalized the aggregate effects of their individual trades. So their trade may impose mispricing when there is a news shock. Empirical evidence supporting these assumptions are discussed in Section 4 of the main paper.

Arbitrageurs are rational and have limited capital in the short-run. They can buy, sell or short sell stocks to maximize their profits. We assume that short selling requires arbitrageurs to deposit the value of the trade in margin accounts, thus limited capital limits them on both buying and short selling. In the long-run they have access to sufficient capital to remove all mispricings. So the model can be viewed as reflecting limits to arbitrage (**?**) and slow moving capital (**?**).

Liquidity traders have an inelastic demand for the stock. The real world counterpart of this group are households who invest a certain part of their income in a stock portfolio, for instance in their retirement accounts. This investment happens regardless of the current market conditions. In reality, much of this type of demand might be invested through mutual funds and pension funds, but for simplicity we abstract from intermediaries and assume that the liquidity traders directly invest in the stock market.

In our model we assume that when there is a good news, all three types of traders participate in the market, but in the case of bad news retail investors are absent and only liquidity traders and arbitrageurs trade the stocks. To justify this assumption note that in the model, demand of the liquidity traders reflects liquidity demands of all types of traders. So indeed the behavior of retail investors in the model is normalized to their liquidity demands. This allows us to have retail investors who do not trade on the bad news days, because their liquidity demand is already reflected. Given the empirical evidence that short selling is not pervasive among retail investors, it is fairly reasonable to assume that they have no other incentive to trade in response to bad news.

7.1 Model Setup

Shares of a firm are being traded at the market. There are four trading days. At each date, there is a stock price that can be different from its fundamental value. We call the difference between the stock price and the fundamental value a mispricing. At date 0, we assume that there is no mispricing and both the fundamental value and the stock price of the firm are V. When there is no news, the only demand for the stocks is liquidity demand. At the beginning of date 1, there is a news story about a shock to the fundamental value of the firm (ΔV). Denote the true fundamental value at the news date by $V' \equiv V + \Delta V$. We call a positive shock to the fundamental value ($\Delta V > 0$) a good news and a negative shock ($\Delta V < 0$) a bad news.

Liquidity traders. At date 0, their demand is V, which sets the stock price given the unit supply of stocks. Their demand is not affected by ΔV and is equal to V, even after the news becomes public. This is the inelasticity of liquidity demand that we discussed before. Therefore, liquidity traders' stock holding at date t is $H_t^L = \dfrac{V}{p_t}$.

Arbitrageurs. Arbitrageurs are risk neutral rational agents and choose their trading strategy to maximize their profit. They precisely assess ΔV and decide about their initial trade accordingly. At date 1 and 2, arbitrageurs have limited aggregate capital of C_1 and C_2, respectively. But, there is unlimited arbitrage capital at date 3 that eliminates mispricings. Thus the stock price at date 3 equals V'. Without loss of generality, we assume a representative arbitrageur who owns the aggregate arbitrage capital. Let A_t be the arbitrageur's endogenous demand for the stock. Therefore her stock holding at date t is $H_t^A = \frac{A_t}{p_t}$.

Retail investors. We assume a continuum of retail investors with limited attention and they trade only if an attention grabbing event attracts them. There are two sources that attract retail investors. At date 1, it is the news that attracts them. Measure $\kappa > 1$ of retail investors trade in reaction to the news shock. These retail investors who trade at the news date hold their stocks until the terminal date. However at subsequent dates (when there is no news) it is the previous day's price change that draws their attention. [2] Meaning that of the retail investors who miss the news (because they are busy), the date 1 price jump attracts measure 1, and they trade at date 2. We assume that date 2 trade of this group is in the same direction as the price change at date 1. We also assume that news grabs more attention than a price jump because of media coverage, etc., thus the mass of retail investors who trade at the news date (measure κ) is larger than the mass who trade subsequently (measure 1). Importantly, retail investors trade only in response to good news. We discussed this point earlier in this section. Figure 1 illustrates the time line.

Figure 1: Time line

$t = 0$	$t = 1$	$t = 2$	$t = 3$
Fundamental value and stock price are V.	News becomes public. Arbitrageurs trade to maximize overall profits. In the case of good news, measure κ of retail traders receive the news and buy stocks. In the case of bad news, retail invesotrs do not trade.	In the case of good news, measure 1 of retail traders buy stock shares. Arbitrageurs trade to maximize overall profits.	There is enough aggregate arbitrage capital to eliminate mispriceings. Stock price is V'.

Retail investors are not able to assess ΔV precisely, but their reaction to the news is in

[2] **?** document this behavior for retail investors.

the same direction as ΔV. This is consistent with retail investors being atomic agents who submit their demands individually and underestimate their aggregate price impact. This effect is captured by a demand shock that increases the retail demand above the impact of the change in the fundamental value. Dollar value of retail investors' demand at each date is $\Delta V + S_t$, where S_t is the demand shock at date t, and is defined as

$$S_t = \beta\ \kappa\ \Delta V_t + \beta\ \Delta p_{t-1} \tag{5}$$

Where β determines the magnitude of the retail demand's deviation from fundamental values, ΔV_t is the shock to the fundamental value, and $\Delta p_{t-1} = p_{t-1} - p_{t-2}$ is the previous day's price change. The first expression on the right-hand side of (5) captures the demand shock to measure κ of retail investors whose attention is grabbed and trade at the news date. Notice that the first expression is only non-zero on date 1, because there is no shock to the fundamental value on other dates. The second term is zero on the news date because there is no price change before the news. However, it captures the demand shock to measure 1 of retail investors who miss the news at date 1, but the date 1 price jump attracts them and they trade at date 2.

Note that a positive β implies overreaction and a negative β implies underreaction to shocks by retail investors. Empirical evidence seem to support the idea that high retail trading leads to overreaction (**??**), so we consider $\beta \in [0, 1]$. A larger β represents a larger retail investors' attention bias, which leads to larger demand shocks. Also, a larger shock to the fundamental value or a larger price jump results in larger demand shocks. Using equation (5), the retail investors' demand shock at dates 1 and 2 are $S_1 = \kappa\ \beta \times (V' - V)$ and $S_2 = \beta \times (p_1 - V)$. Having their demand, the quantity of retail investors' asset holding at date t is $H_t^R = \dfrac{\Delta V + S_t}{p_t}$.

In our model, liquidity traders and retail investors are non-optimizing agents. That is their aggregate demand is either inelastic (in the case of liquidity traders) or is a deterministic function of the model variables (in the case of retail investors). The only agent who optimizes trading strategies is the arbitrageur. Therefore, an equilibrium in this model is defined as the arbitrageur's optimal policy (stock demand) and a stock price on each day.

Market Clearing. The price is determined by the market clearing condition that equates supply and demand of shares. At each date, we assume that there is a constant unit supply of shares in the market. The market clearing condition at date 1 is

$$H_1^L + H_1^R + H_1^A = 1 \tag{6}$$

The market clearing condition at date 2 depends on the arbitrageur's strategy at date 1 because she can trade her stock holdings from date 1. If the arbitrageur buys at date 1, i.e. $A_{1B} \in (0, C_1]$, she could supply her date 1 stock holdings at date 2. But this does not happen if she sells at date 1, i.e. $A_{1S} \in [-C_1, 0]$. Hence, date 2 market clearing conditions are different under the buying strategy, and the selling strategy.

The arbitrageur chooses her trades at each date to maximize net profits from day 1 to 3, subject to capital constraints and a participation constraint. To save notation and without loss of generality, we assume equal arbitrage capital at times 1 and 2: $C_1 = C_2 = C$. So the arbitrageur's problem is

$$\begin{aligned}
\underset{A_1, A_2}{\text{Max}} \quad & \Pi(A_1, A_2; p_1, p_2) & \\
\text{s.t.} \quad & -C \leqslant A_1 \leqslant C & \text{(date 1 capital const.)} \\
& \underline{\phi}(A_1; C) \leqslant A_2 \leqslant \bar{\phi}(A_1; C) & \text{(date 2 capital const.)} \\
& \Pi(A_1^*, A_2^*) \geqslant 0 & \text{(participation const.)}
\end{aligned}$$

$\Pi(.)$ is the arbitrageur's net profit from trades, which obviously depends on her demand for stocks and prices. Note that the arbitrageur closes all trading positions at date 3, so her trade at the terminal date is determined by her trades at the first two periods. The date 2 capital constraint depends on date 1 trade; for instance if the arbitrageur simply does not trade on the first day, she has $2C$ available capital at date 2. Finally, the participation constraint ensures that there is an overall gain to the trades, otherwise the arbitrageur just avoids trading.

Solving for the arbitrageur's optimal strategy and the equilibrium prices precisely requires the distinction between positive shocks and negative shocks due to the time 2 market clearing, and the presence of retail investors who are only active in response to positive shocks.

7.2 Analysis of Positive News Shocks

All three groups of investors trade in response to a positive news. The arbitrageur faces a trade-off in choosing her trading strategy at date 1. She has no control on retail investors' date 1 trade in response to news. However, because her date 1 trade has a price impact, her choice affects date 2 trades of the retail investors. In the sense that if she buys stocks at date 1, this increases date 1 price. The date 1 price jump is the source of attracting retail investors to trade at date 2. Therefore, although under buying strategy she buys at a higher price at date 1, it will increase retail investors' demand at date 2. This in turn helps the arbitrageur to sell at a higher date 2 price. This strategy resembles bubble riding. On the other hand, selling strategy at date 1 helps the arbitrageur to sell at medium prices at both dates. So it is not immediately clear which strategy dominates the other one. The following proposition sketches the arbitrageur's optimal trading strategy.

Proposition 7.1 *In response to a positive news, the arbitrageur's optimal strategy is to trade against the mispricing by adopting a selling strategy with* $A_1^* = -C$*, and* $A_2^* = -2C + C\frac{p_2^*}{p_1^*}$*, if the arbitrage capital is limited* ($C \in [0, C^*]$) *and retail investors have attention bias* ($\beta \in [\beta^*, \hat{\beta}]$)*, where* C^**,* β^**, and* $\hat{\beta}$ *are defined in Equations (21), (14), and (19), respectively.*

Proof See Appendix A.1.

The intuition is that if the attention bias is too large ($\beta > \hat{\beta}$), it makes buying strategies relatively more attractive. The reason is that larger values of β incentivize the bubble riding behavior, that is to buy at date 1 in order to attract more retail investors tomorrow, which allows the arbitrageur to sell at a higher price at date 2. But for moderate attention bias the arbitrageur's optimal strategy is to trade against the mispricing, not riding a bubble. So the arbitrageur sells the stocks on date 1 and 2, but because of the capital constraints she cannot remove all the mispricing at these dates.

Having the optimal strategies, the equilibrium prices at date 1 and 2 in response to a positive shock are determined:

$$p_1^* = V' + \beta\kappa\,\Delta V - C \tag{7}$$

$$p_2^* = \frac{\Big(V' - \beta V - 2C + \beta(V' + \beta\kappa\,\Delta V - C)\Big)(V' + \beta\kappa\,\Delta V - C)}{V' + \beta\kappa\,\Delta V - 2C} \tag{8}$$

Hence, we can look at the post-shock price patterns in equilibrium:

Corollary 7.2 *In response to a positive news at date 1:*

1. *The stock market overreacts at the news day, that is* $p_1^* > V'$.
2. *The mispricing declines over time (i.e. no bubble riding), that is* $p_1^* > p_2^* > V'$.
3. *The price decline is steeper at date 2 compared to date 3, that is* $p_1^* - p_2^* > p_2^* - V'$.

Proof See Appendix A.3.

Corollary 7.2 shows the predicted equilibrium price patterns under the model assumptions. The first part shows that the model predicts that the stock market overreacts to good news on the news day. The second part shows that the largest mispricing occurs at the news date. The price declines at date 2 and finally at date 3 it is equal to the fundamental value. The third part predicts that most of the price correction occurs at date 2 and a smaller part is left for date 3. That is the price graph has a steeper slope between date 1 and 2 than between date 2 and 3. All of these predictions are consistent with our empirical findings in Tables **??** and **??**. There we documented overreaction to good news on the news day, followed by gradual removal of mispricings. Also, comparing $ACARs$ for the first 10 days after the positive shocks, to $ACARs$ over days 11 to 40 confirms the steeper price corrections in the earlier post-shock days.

Corollary 7.3 *Larger attention bias in individual investors results in larger overreaction to good news, that is larger mispricings at dates 1 and 2:* $\dfrac{\partial p_1^*}{\partial \beta} > 0$ *and* $\dfrac{\partial p_2^*}{\partial \beta} > 0$.

Proof See Appendix A.4.

Corollary 7.3 shows the impact of β on price patterns. We defined β as determining the extent that the aggregate demand of attention biased retail investors exceeds the rational demand in response to positive news. So the more retail investors trade in response to a shock the higher the effective β will be. The model prediction with respect to β is also in line with our findings in the data. In the retail investors' tests we found that more retail trading activity leads to more overreaction to good news.

Corollary 7.4 *Larger arbitrage capital results in smaller overreaction to good news at date 1, that is* $\frac{\partial p_1^*}{\partial C} < 0$.

Proof Obvious.

Corollary 7.4 predicts the impact of arbitrage capital on price patterns. This prediction is also supported by our empirical findings in the arbitrageur's tests. We found that in situations where the effective arbitrage capital is small, overreaction to positive news is significantly stronger.

7.3 Analysis of Negative News Shocks

There are two main differences between the analysis of positive and negative news shocks. The obvious one is that in the case of bad news, by definition, the shock to the fundamental value is negative, i.e. $\Delta V < 0$. But the more important difference is that retail investors do not actively trade in response to the negative shocks.

The arbitrageur is free to buy or sell stocks to maximize overall profits from the trade. The following proposition determines the optimal strategy for the arbitrageur in response to bad news.

Proposition 7.5 *In response to a negative news, the arbitrageur's optimal strategy is to trade against the mispricing by adopting a selling strategy with* $A_1^* = A_2^* = -C$*, if the arbitrage capital is limited:* $C \in \left[0, \min\{\frac{V}{2}, -\Delta V\}\right]$.

Proof See Appendix A.5.

The intuition of negative news cases is relatively simpler, mainly because retail investors do not participate in the market. The liquidity demand is inelastic with respect to the news. So the arbitrageur sells the stock that moves the price closer to the new fundamental value. However, limited arbitrage capital prevents the arbitrageur to remove all the mispricing at dates 1 and 2.

Having the optimal strategies, the equilibrium prices in response to a negative shock can be computed as $p_1^* = p_2^* = V - C$. It is easy to observe the price patterns and the impact of arbitrage capital on prices:

Corollary 7.6 *In response to a negative news at date 1:*

1. *The stock market underreacts, that is* $p_1^* = p_2^* > V'$.
2. *Larger arbitrage capital results in smaller under-reaction to bad news, that is a smaller mispricing at dates 1 and 2:* $\dfrac{\partial p_1^*}{\partial C} = \dfrac{\partial p_2^*}{\partial C} < 0$.

Proof Obvious.

The predicted results in Corollary 7.6 are again consistent with our empirical findings. In Tables **??** and **??** we found that the stock market underreacts to negative news shocks. That is an initial negative price shock, which is followed by even more negative price drifts. Also, we showed in the arbitrageur's tests that when there is abundant arbitrage capital there is virtually no underreaction to negative news, and as the arbitrage capital shrinks the underreaction patterns become stronger.

Note that corollaries 7.4 and 7.6 underscore the importance of arbitrage capital in financial markets. Part of what took place during the financial crisis of 2008-2009 appears to have been a scarcity of arbitrage capital to deal with a great deal of news.[3] This scarcity was driven partly by the arbitrageurs own financial difficulties, and partly by regulatory pressure for these firms to reduce leverage. From the perspective of the model this ought

[3] We thank Gerry Hoberg for suggesting that our model can provide an interpretation of what took place during the financial crisis. There was a great deal of news to absorb, so the retail traders were active. At the same time arbitrage capital was being reduced, so the pricing was less tightly tied to true values than usual.

to have produced more predictable stock price patterns, and hence average prices that were further from the true asset values.

7.4 Analysis of No-News shocks

So far we have discussed the price shocks caused by news providing information about a change in the fundamental value of the firm. However, often price shocks are not accompanied by news, the case that we call no-news shocks. In these situations the price shock is due to random fluctuations in the supply or demand of the shares. We use the same market structure to study market reaction to these shocks.

To be consistent with the previous analysis, we assume a unit supply of shares at all dates, and that the no-news price shocks are caused by an exogenous demand shock to liquidity traders at date 1. Assume that the demand shock changes the liquidity traders' demand from V to V' at date 1, but their demand is equal to V on all other days. Since this demand shock is exogenous, the arbitrageur does not know about that until after day 1 settlement when the price jump is observed. Thus, in this case the arbitrageur does not trade at date 1. Retail investors are also not present at date 1 due to the lack of news or any other source of attraction. So date 1 price is $p_1^* = V'$, which is different from the fundamental value V.

At date 2, the previous day's price jump attracts retail investor's attention, but induce retail trade only if the initial price jump is positive. The arbitrageur chooses date 2 trade to maximize the overall profit, knowing that there will be no mispricing at the terminal date.

Given this simple setting, the arbitrageur's optimal choice at date 2 is straightforward. In case of negative no-news shocks, the original source of mispricing (the liquidity shock) goes away, so the price will be back to fundamental value without arbitrageur's action. In the case of positive no-news shocks, attracted retail trades are the new source of mispricing. So the arbitrageur trades to make profit out of the mispricing. If the arbitrage capital is limited and the attention bias is not excessively large, the arbitrageur trades against the mispring, i.e. sells the stock and avoids bubble riding. Arbitrageur's optimal strategy and price implications of no-news shocks are discussed more formally in the following Corollary.

Corollary 7.7 *Both positive and negative no-news price shocks are reversed in subsequent days. In response to positive no-news shocks, the arbitrageur's optimal strategy is to trade against the mispricing, if arbitrage capital is limited:* $C \in \left[0, \beta \frac{\Delta V}{2}\right]$.

Proof See Appendix A.7.

Model predictions of the stock market response to no-news shocks are consistent with our empirical findings in Table **??**, where we documented the well-known reversal after no-news shocks. Here we showed that in addition to market responses to news shocks, this can also be naturally explained by considering interactions of different groups of investors.

Table 1: Drugs industry (SIC 283) analysis

This table shows the results of the event study analysis for the sub-sample of the drugs industry (SIC 283). Panel A shows the results for the comprehensive analysis of all news, and the separated analyses of the negative and positive sub-samples. Columns show the abnormal return on the news date ($ARet_0$), and the average cumulative abnormal returns over the period from $t1$ to $t2$ days after the news date ($ACAR_{[t1,t2]}$). To calculate abnormal returns, the benchmark is the average return of all matched CRSP firms on a triple-sort into deciles of size, B/M and momentum. All returns are calculated in basis points. In panel B, negative and positive observations are distinguished according to the sign of the $ARet_0$. Observations of each sign are separately sorted into quintiles based on the $ARet_0$. t-statistics are shown in parentheses. The *, ** and *** symbols denote statistical significance at the 10%, 5% and 1% levels.

Panel A:				ACAR			
	Shock	Obs.	$ARet_0$	[1,21]	[1,63]	[2,21]	[2,63]
All news		1437	-12.17**	-3.54***	-2.95***	-3.48***	-2.89***
			(-1.98)	(-3.91)	(-4.69)	(-3.81)	(-4.58)
Negative news	-	703	-169.21***	-3.16**	-1.96**	-3.12**	-1.92**
			(-21.83)	(-2.29)	(-2.08)	(-2.30)	(-2.04)
Positive news	+	734	138.24***	-3.90***	-3.85***	-3.82***	-3.78***
			(26.62)	(-3.24)	(-4.55)	(-3.10)	(-4.44)
Panel B: sort on $ARet_0$							
Q1 (Largest Negative)	-	141	-458.38***	-1.08	-0.06	-0.87	-0.17
			(-18.21)	(-0.28)	(-0.03)	(-0.24)	(-0.07)
Q2	-	140	-188.53***	-5.16*	-2.91	-5.08*	-2.70
			(-82.39)	(-1.72)	(-1.37)	(-1.67)	(-1.25)
Q3	-	141	-115.17***	1.38	1.29	0.20	0.87
			(-82.11)	(0.51)	(0.65)	(0.08)	(0.45)
Q4	-	140	-62.73***	-1.06	-0.95	-0.74	-0.88
			(-60.75)	(-0.39)	(-0.52)	(-0.29)	(-0.49)
Q5 (Smallest Negative)	-	141	-20.61***	-9.13***	-6.46***	-8.53***	-6.08***
			(-21.08)	(-3.08)	(-3.11)	(-2.84)	(-2.92)
Q1 (Smallest Positive)	+	147	15.09***	-5.13**	-3.65**	-5.61**	-3.72**
			(19.57)	(-2.09)	(-1.98)	(-2.30)	(-2.04)
Q2	+	147	51.57***	-4.80**	-4.93***	-3.67	-4.55***
			(53.31)	(-2.22)	(-2.96)	(-1.59)	(-2.71)
Q3	+	146	97.65***	-3.22	-3.41**	-3.49	-3.67**
			(81.24)	(-1.13)	(-2.02)	(-1.22)	(-2.17)
Q4	+	147	166.13***	-1.08	-2.17	-0.93	-1.86
			(76.45)	(-0.39)	(-0.95)	(-0.32)	(-0.81)
Q5 (Largest Positive)	+	147	360.45***	-5.18	-5.16**	-5.38	-5.16**
			(28.64)	(-1.55)	(-2.49)	(-1.55)	(-2.46)

Table 2: Regressions run separately for news and no-news observations using ACAR

Columns 1, 3 and 4 of this table report the coefficient estimates of equation (2), and columns 2 and 5-7 report those of equation (1). Each column denotes a sub-sample used to estimate the regression. The dependent variable, $ACAR_{2,40}$, is the average cumulative abnormal return over day 2 to 40 after the shock date. To calculate abnormal returns, the benchmark is the average return of all matched CRSP firms on a triple-sort into deciles of size, B/M and momentum. The independent variables include the shock date abnormal return ($ARet_{it}$), firm size ($Size_{it}$), interaction between abnormal returns and size ($ARet_{it} * Size_{it}$), book to market (B/M_{it}), annual return momentum (MOM_{it}), and monthly return volatility (VOL_{it}). All the returns are calculated in percentage points. Heteroskadisticity-robust standard errors are shown in parentheses. The *, ** and *** symbols denote statistical significance at the 10%, 5% and 1% levels.

	All News		Pos News	Neg News	All No-News	Pos No-News	Neg No-News
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$ACAR_{2,40}$	$ACAR_{2,40}$	$ACAR_{2,40}$	$ACAR_{2,40}$	$ACAR_{2,40}$	$ACAR_{2,40}$	$ACAR_{2,40}$
$ARet_{it}$	0.011	-0.003*	-0.011	0.050**	-0.003***	-0.001	-0.004***
	(0.008)	(0.002)	(0.011)	(0.021)	(0.000)	(0.000)	(0.001)
$ARet_{it} * Size_{it}$	-0.001*		0.001	-0.005**			
	(0.001)		(0.001)	(0.002)			
$Size_{it}$	0.002	0.002	0.000	-0.005	-0.008***	-0.008***	-0.008**
	(0.005)	(0.005)	(0.005)	(0.004)	(0.002)	(0.002)	(0.002)
B/M_{it}	0.002	0.002	0.003	0.003	-0.016***	-0.016***	-0.016***
	(0.004)	(0.004)	(0.007)	(0.004)	(0.004)	(0.004)	(0.004)
MOM_{it}	1.288***	1.274***	1.325***	1.492***	-1.487	-1.369	-1.562
	(0.424)	(0.419)	(0.406)	(0.496)	(1.142)	(1.077)	(1.224)
VOL_{it}	0.585	0.560	0.189	0.338	1.104***	0.821**	1.245***
	(0.469)	(0.467)	(0.545)	(0.521)	(0.209)	(0.197)	(0.211)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	25,875	25,875	12,593	13,282	1,620,078	795,571	824,507

Table 3: Pooling regressions

Column 1 and 3 of this table report the coefficient estimates of equation (3), and column 2 and 4 report those of equation (4). The dependent variable in the first two columns is the average cumulative abnormal return over day 2 to 40 after the shock date ($ACAR_{2,40}$). To calculate abnormal returns, the benchmark is the average return of all matched CRSP firms on a triple-sort into deciles of size, B/M and momentum. In column 3 and 4, the dependent variable is the average raw return over day 2 to 40 after the shock date ($AvgRet_{2,40}$). The independent variables include the shock date abnormal return ($ARet_{it}$), (positive/negative) news indicator ($(Pos/Neg)News_{it}$), interaction between returns and (positive/negative) news indicators ($(Pos/Neg)News_{it} * (ARet_{it}$ or $Ret_{it})$) , firm size ($Size_{it}$), book to market (B/M_{it}), annual return momentum (MOM_{it}), and monthly return volatility (VOL_{it}). All the returns are calculated in percentage points. Heteroskadisticity-robust standard errors are shown in parentheses. The *, ** and *** symbols denote statistical significance at the 10%, 5% and 1% levels.

	ACAR		AvgRet	
	(1) $ACAR_{2,40}$	(2) $ACAR_{2,40}$	(3) $AvgRet_{2,40}$	(4) $AvgRet_{2,40}$
$ARet_{it}$ or Ret_{it}	-0.003***	-0.003***	-0.006***	-0.006***
	(0.000)	(0.001)	(0.001)	(0.001)
$News_{it} * (ARet_{it}$ or $Ret_{it})$	-0.001		0.003**	
	(0.002)		(0.002)	
$News_{it}$	-0.017***		-0.03***	
	(0.008)		(0.010)	
$PosNews_{it} * (ARet_{it}$ or $Ret_{it})$		-0.001		0.001
		(0.002)		(0.002)
$NegNews_{it} * (ARet_{it}$ or $Ret_{it})$		-0.001		0.005*
		(0.006)		(0.003)
$PosNews_{it}$		-0.016**		-0.022**
		(0.006)		(0.009)
$NegNews_{it}$		-0.020**		-0.026***
		(0.009)		(0.009)
$Size_{it}$	-0.008***	-0.008***	-0.019***	-0.019***
	(0.002)	(0.002)	(0.002)	(0.002)
B/M_{it}	-0.016***	-0.016***	-0.028***	-0.028***
	(0.004)	(0.004)	(0.007)	(0.007)
MOM_{it}	-1.008	-1.008	-0.412	-0.421
	(1.083)	(1.084)	(1.364)	(1.381)
VOL_{it}	1.114***	1.113***	2.447***	2.481***
	(0.205)	(0.206)	(0.513)	(0.523)
Firm FE	Yes	Yes	Yes	Yes
N	1,623,203	1,623,203	1,623,203	1,623,203

Table 4: Availability of arbitrage capital based on intermediary capital ratio (monthly frequency)

News observations are sorted into quintiles based on aggregate intermediary capital ratio growth rate. Columns show the abnormal return on the news date ($ARet_0$), and the average cumulative abnormal returns over the period from t1 to t2 days after the news date ($ACAR_{[t1,t2]}$). To calculate abnormal returns, the benchmark is the average return of all matched CRSP firms on a triple-sort into deciles of size, B/M and momentum. All returns are calculated in basis points. Negative and positive observations are distinguished according to the sign of the $ARet_0$, and reported in the top and the bottom panel, respectively. t-statistics are shown in parentheses. The *, ** and *** symbols denote statistical significance at the 10%, 5% and 1% levels.

Sort:				ACAR			
Cap. ratio growth rate (η_t^Δ)	Shock	Obs.	$ARet_0$	[1,21]	[1,63]	[2,21]	[2,63]
Q1 (Smallest η_t^Δ)	-	3173	-211.12***	-1.94**	-2.19***	-1.77**	-2.08***
			(-40.47)	(-2.26)	(-3.43)	(-2.05)	(-3.23)
Q2	-	3234	-167.41***	-1.01*	-0.21	-0.98*	-0.18
			(-44.82)	(-1.74)	(-0.45)	(-1.67)	(-0.39)
Q3	-	3197	-164.99***	-0.45	-2.43***	-0.54	-2.49***
			(-37.99)	(-0.66)	(-5.13)	(-0.79)	(-5.25)
Q4	-	3138	-158.44***	-0.77	-1.23***	-0.77	-1.22***
			(-47.32)	(-1.16)	(-2.66)	(-1.13)	(-2.61)
Q5 (Largest η_t^Δ)	-	3158	-186.66***	1.20	0.08	1.18	0.06
			(-44.19)	(1.46)	(0.14)	(1.41)	(0.10)
Q5 - Q1	-		24.46***	3.14***	2.27***	2.95**	2.14**
			(3.64)	(2.64)	(2.60)	(2.45)	(2.44)
Q1 (Smallest η_t^Δ)	+	3052	208.89***	-2.79***	-3.22***	-3.10***	-3.29***
			(38.03)	(-3.20)	(-5.05)	(-3.56)	(-5.17)
Q2	+	3003	166.38***	-0.82	-0.72*	-1.06	-0.79*
			(44.71)	(-1.22)	(-1.72)	(-1.54)	(-1.65)
Q3	+	3028	166.34***	-2.16***	-2.41***	-2.46***	-2.53***
			(39.23)	(-3.36)	(-5.14)	(-3.73)	(-5.33)
Q4	+	3035	167.13***	-1.87***	-0.81*	-2.04***	-0.85*
			(45.92)	(-2.89)	(-1.70)	(-3.12)	(-1.77)
Q5 (Largest η_t^Δ)	+	3030	200.15***	-1.34*	-1.92***	-1.60*	-2.08***
			(40.05)	(-1.68)	(-3.26)	(-1.96)	(-3.52)
Q5 - Q1	+		-8.74	1.45*	1.30**	1.51*	1.2*
			(1.17)	(1.72)	(2.00)	(1.76)	(1.89)

A Proofs

A.1 Proof of Proposition 7.1

Proof The arbitrageur knows that at date 3 there is no mispricing in the stock market, so she sells as much as it is feasible for her so long as the price at date 2 is above the new fundamental value, V'. Therefore, arbitrageur's demand at date 2 is equal to the negative of her available capital at date 2, that is

$$A_2 = -\Big[C_2 + \big(C_1 - |A_1|\big) + H_1^A \times (p_2 - p_1)\Big]$$

where C_2 is the new capital available to the arbitrageur at date 2, $\big(C_1 - |A_1|\big)$ is the amount of capital left from date 1, and $H_1^A \times (p_2 - p_1)$ is profit (loss) from her stock holdings at date 1. The arbitrageur's demand at date 1, A_1, has opposite signs if the arbitrageur buys or sells stocks. Therefore, A_{2B} and A_{2S}, the arbitrageur's date 2 demand under buying and selling strategies, respectively, are

$$\begin{cases} A_{2B} = -[C_1 + C_2 + H_{1B}^A p_{2B} - 2H_{1B}^A p_{1B}] \\ A_{2S} = -[C_1 + C_2 + H_{1S}^A p_{2S}] \end{cases} \tag{9}$$

Date 2 market clearing conditions in the two cases are

$$\begin{cases} H_2^L + H_2^R + H_2^A = 1 + H_{1B}^A & \text{if buying strategy at } t = 1 \\ H_2^L + H_2^R + H_2^A = 1 & \text{if selling strategy at } t = 1 \end{cases} \tag{10}$$

Substituting the stock holdings (H_t^i) for all agents into date 1 market clearing condition (Equation 6), date 1 stock price is

$$p_1 = V' + S_1 + A_1$$

To find the arbitrageur's optimal strategy we examine the selling and buying strategies separately. This is important due to the difference in the arbitrageur's profit function and also the market clearing conditions as in equation (10).

First consider the case of selling strategy in which the arbitrageur short sells stocks at date 1 and her demand is $A_{1S} \in [-C, 0]$. The arbitrageur cash out after the final date at the price V', so her profit at the terminal date is

$$\begin{aligned} \Pi_S &= (H_{1S}^A + H_{2S}^A)(V' - p_{2S}) \\ &= 2C(1 - \frac{V'}{p_{2S}}) \end{aligned} \tag{11}$$

Where p_{2S} is date 2 price. Using the market clearing condition under the selling strategy from equation (10), date 2 price is

$$p_{2S} = \frac{\Big(V' - \beta V - 2C + \beta(V' + S_1) + \beta A_{1S}\Big)(V' + S_1 + A_{1S})}{V' + S_1 + 2A_{1S}} \tag{12}$$

Having the profit function, the arbitrageur's optimal selling strategy solves the following problem:

$$\begin{aligned} \underset{A_{1S}}{\text{Max}} \quad & 2C(1 - \frac{V'}{p_{2S}}) \\ \text{s.t.} \quad & -C \leqslant A_{1S} \leqslant 0 \qquad \text{(capital constraint)} \\ & \Pi_S(A^*_{1S}) \geqslant 0 \qquad \text{(IR constraint)} \end{aligned} \tag{13}$$

The following lemma sketches the optimal selling strategy for the arbitrageur.

Lemma A.1 *In response to a positive shock, the optimal selling strategy for arbitrageurs at date 1 is $A^*_{1S} = -C$, i.e. the capital constraint is binding, if the arbitrage capital is limited and retail investors' attention bias is large enough, that is $C \in \big[0, \frac{1}{2}(V' - \beta V)\big]$ and $\beta > \beta^*$, where β^* is defined in the solution to*

$$f(\beta^*) = 0 \tag{14}$$

and $f(\beta) \equiv (\kappa^2\,\Delta V^2)\beta^3 + 2\kappa\,\Delta V(V' - \frac{V}{2} - C)\beta^2 + \Big((V' - C)^2 - V(V' - C) - 2\kappa C\,\Delta V\Big)\beta + 2C^2 - CV' \geqslant 0.$

Proof See Appendix A.2.

Now consider the case of buying strategy in which the arbitrageur buys stocks at date 1 and her demand is $A_{1B} \in (0, C_1]$. In this case, the arbitrageur's profit is

$$\begin{aligned} \Pi_B &= (H^A_{2B})(V' - p_{2B}) \\ &= \Big(2C + A_{1B}(\frac{p_{2B}}{p_{1B}} - 2)\Big)(1 - \frac{V'}{p_{2B}}) \end{aligned} \tag{15}$$

Where, using the date 2 market clearing condition from equation (10), date 2 price is

$$p_{2B} = \frac{\Big(V' - \beta V - 2C + \beta(V' + S_1) + (2 + \beta)A_{1B}\Big)(V' + S_1 + A_{1B})}{V' + S_1 + 3A_{1B}} \tag{16}$$

Notice the difference between the first line of equations (11) and (15). When the arbitrageur sells stocks at date 1, she adjusts her margin at date 2 but holds her stock holding from date 1 to the final date, thus we see $(H^A_{1S} + H^A_{2S})$ as the quantity she holds at the final date. However, under the buying strategy she closes her long position at date 2 and turn it to a new short sell position. Therefore, the quantity she holds until the final date is only H^A_{2B}.

Lemma A.1 sketches the optimal selling strategy, $A^*_{1S} = -C$. This proposition shows that under limited capital and attention bias, the optimal selling strategy dominates all other strategies. We have already shown in lemma A.1 that A^*_{1S} dominates all other selling strategies. So to complete the proof, it is sufficient to show that the optimal selling

strategy dominates all buying strategies, too. From lemma A.1 we have $\Pi_S^* = \Pi_S(A_{1S}^*) = 2C(1 - \frac{V'}{p_{2S}^*})$. Also, from the analysis of the buying strategy in equation (15) we know $\Pi_B = \Big(2C - A_{1B}(\frac{p_{2B}}{p_{1B}} - 2)\Big)(1 - \frac{V'}{p_{2B}})$.
The ideal way to prove this proposition is to solve for the optimal buying strategy, and to compare the payoff of the optimal strategies in the two cases and find the dominant strategy. However, analytically solving for the optimal buying strategy is almost impossible due to the cumbersome algebra. Instead, we find an upper bound for Π_B, $UB(\Pi_B)$, and compare it to Π_S^*. If Π_S^* is larger than $UB(\Pi_B)$, we conclude that the optimal selling strategy dominates all buying strategies and indeed is the dominant strategy.

Since $A_{1B}(\frac{p_{2B}}{p_{1B}} - 2) < 0$, $2C$ is an upper bound for the first bracket in Π_B. Taking derivative of the second bracket to find its maximum, we have

$$\begin{aligned}\frac{\partial}{\partial A_{1B}}\Big(1 - \frac{V'}{p_{2B}}\Big) &= \frac{\partial}{\partial p_{2B}}\Big(1 - \frac{V'}{p_{2B}}\Big) \times \frac{\partial p_{2B}}{\partial A_{1B}} \\ &= \frac{\partial}{\partial p_{2B}}\Big(1 - \frac{V'}{p_{2B}}\Big) \times \frac{(2+\beta)\Big(3A_{1B}^2 + 2QA_{1B} + Q(Q - \frac{2P}{2+\beta})\Big)}{(Q + 3A_{1B})^2} > 0\end{aligned} \tag{17}$$

Where P and Q are defined in equation (23). Thus the second bracket in Π_B achieves its maximum when $A_{1B} = C$. Therefore, we can define an upper bound for the profit under the buying strategy as $UB(\Pi_B) \equiv 2C(1 - \frac{V}{p_{2B}(C)})$. Therefore, we have

$$\begin{aligned} & \Pi_S^* > UB(\Pi_B) \\ \Leftrightarrow \quad & 2C(1 - \frac{V'}{p_{2S}^*}) > 2C(1 - \frac{V'}{p_{2B}(C)}) \\ \Leftrightarrow \quad & P_{2S}^* > p_{2B}(C) \\ \Leftrightarrow \quad & \frac{(P - \beta C)(Q - C)}{(Q - 2C)} > \frac{(P + (2+\beta)C)(Q + C)}{(Q + 3C)} \\ \Leftrightarrow \quad & (\frac{Q + 3C}{Q + C}) \times (\frac{Q - C}{Q - 2C}) > \frac{P + (2+\beta)C}{P - \beta C} \\ \Leftrightarrow \quad & (1 + \frac{2C}{Q + C}) \times (1 + \frac{C}{Q - 2C}) > 1 + \frac{2C(1+\beta)}{P - \beta C} \\ \Leftrightarrow \quad & \frac{3(Q - C/3)}{(Q - 2C) \times (Q + C)} > \frac{2(1+\beta)}{P - \beta C} \\ \Leftrightarrow \quad & \frac{3(Q - C/3)}{(Q - 2C)} > \frac{2(1+\beta) \times (Q + C)}{P - \beta C} \end{aligned} \tag{18}$$

Since $\frac{3(Q-C/3)}{(Q-2C)} > 3$, it is sufficient to show $\frac{2(1+\beta)\times(Q+C)}{P-\beta C} < 3$. Substituting for P, Q, and $S_1 = \beta\kappa\,\Delta V$, the sufficient condition is equivalent to $(\kappa\,\Delta V)\beta^2 + (V' - 3V - 5C - 2\kappa\,\Delta V)\beta + (V' - 8C) > 0$.

This inequality holds if and only if $C < \dfrac{V'}{8}$ and $\beta < \hat{\beta}$, where

$$\hat{\beta} = \frac{2\kappa\Delta V + 3V - V' + 5C - \sqrt{(2\kappa\Delta V + 3V - V' + 5C)^2 - 4\kappa\,\Delta V(V' - 8C)}}{2\kappa\,\Delta V} \tag{19}$$

Notice that the other root for the quadratic expression of the sufficient condition is larger than one, thus it is not part of the solution since $\beta \in [0, 1]$. From lemma A.1 for the optimal selling strategy to be individual rational it must be the case that $\beta > \beta^*$ and $C < \frac{1}{2}(V' - \beta V)$. If $\beta^* > \hat{\beta}$ then A^*_{1S} never dominates any A_{1B}. However, β^* is increasing in C, and $\hat{\beta}$ is decreasing in C. Evaluating β^* and $\hat{\beta}$ on the two extremes of the range for C, we get

$$\begin{aligned} &\text{At } C = 0, \hat{\beta} > 0 \text{ and } \beta^* = 0 \Rightarrow \hat{\beta} - \beta^* > 0 \\ &\text{At } C = \frac{V'}{8}, \hat{\beta} = 0 \text{ and } \beta^* > 0 \Rightarrow \hat{\beta} - \beta^* < 0 \end{aligned} \tag{20}$$

Because $\hat{\beta} - \beta^*$ is a continuous function, Kakutani's fixed point theorem implies that $\exists C^* \in [0, \frac{V'}{8}]$ such that $\hat{\beta} > \beta^*, \forall C \in [0, C^*]$, and C^* is the solution to

$$\beta^*(C) = \hat{\beta}(C) \tag{21}$$

Hence, with limited arbitrage capital and large enough attention bias, the optimal selling strategy dominates every buying strategy; that is
If $C \in [0, C^*]$ and $\beta \in [\beta^*, \hat{\beta}]$, then $\Pi^*_S > \Pi_B(A_{1B}), \forall A_{1B} \in [0, C]$. ∎

A.2 Proof of Lemma A.1

Proof The optimal selling strategy of the arbitrageur solves the following problem:

$$\begin{aligned} \max_{A_{1S}} \quad & 2C\left(1 - \frac{V'}{p_{2S}}\right) \\ \text{s.t.} \quad & -C \leqslant A_{1S} \leqslant 0 \qquad \text{(capital constraint)} \\ & \Pi_S(A^*_{1S}) \geqslant 0 \qquad \text{IR constraint} \end{aligned} \tag{22}$$

First, we solve the problem without the IR constraint and then check whether IR is satisfied. Let $P \equiv V' - \beta V - 2C + \beta(V' + S_1)$ and $Q \equiv V' + S_1$; date 2 price can be written as

$$P_{2S} = \frac{(P + \beta A_{1S})(Q + A_{1S})}{Q + 2A_{1S}} = \frac{\beta A^2_{1S} + (\beta Q + P)A_{1S} + PQ}{2A_{1S} + Q} \tag{23}$$

Therefore, the profit is calculated as $\Pi_S = 2C(1-\frac{V'}{p_{2S}}) = 2C-2CV'\dfrac{2A_{1S}+Q}{\beta A_{1S}^2+(\beta Q+P)A_{1S}+PQ}$. Taking derivative with respect to A_{1S}, we have

$$\frac{\partial \Pi_S}{\partial A_{1S}} = 4\beta CV' \frac{A_{1S}^2 + (V'+S_1)A_{1S} - \frac{1}{2}(V'+S_1)(V'-\beta V-2C)}{\Big(\beta A_{1S}^2 + (\beta Q+P)A_{1S}+PQ\Big)^2} < 0\ ,\ \forall A_{1S} \in (A_0^-, A_0^+) \tag{24}$$

Where A_0^- and A_0^+ are roots of the quadratic in the numerator. Notice that $[-C,0] \subset (A_0^-, A_0^+)$ if $C < \dfrac{1}{2}(V'-\beta V)$. Thus we have the following result $A_{1S}^* = -C, \quad \text{if} \quad C \in \big[0, \dfrac{1}{2}(V'-\beta V)\big]$.

Next we have to check under what conditions the IR constraint is satisfied.

$$\begin{aligned}
&\Pi_S^* = \Pi_S(A_{1S}^*) = 2C(1-\frac{V'}{p_{2S}(A_{1S}^*)}) \geqslant 0\\
\Leftrightarrow\quad & p_{2S}(A_{1S}^*) \geqslant V'\\
\Leftrightarrow\quad & \frac{\Big(V'-\beta V-2C+\beta(V'+S_1)-\beta C\Big)(V'+S_1-C)}{V'+S_1-2C} \geqslant V'\\
\Leftrightarrow\quad & \beta(V'+S_1-C)^2 - (\beta V+2C)(V'+S_1-C) + CV' \geqslant 0\\
\Leftrightarrow\quad & f(\beta) \equiv (\kappa^2\,\Delta V^2)\beta^3 + 2\kappa\,\Delta V(V'-\tfrac{V}{2}-C)\beta^2 + \Big((V'-C)^2 - V(V'-C) - 2\kappa C\,\Delta V\Big)\beta + 2C^2 - CV' \geqslant 0
\end{aligned} \tag{25}$$

In the last line, we substitute for S_1 using $S_1 = \beta\kappa\,\Delta V$. If $C < \frac{V'}{2}$ then we have $f(\beta)|_{\beta=0} < 0$, and we always have $f(\beta)|_{\beta=1} > 0$. Since $f(\beta)$ is continuous in β, Kakutani's fixed point theorem implies that $\exists\ \beta^*$ such that

$$f(\beta^*) = 0 \tag{26}$$

And $f(\beta) \geqslant 0$, $\forall\ \beta > \beta^*$, that is the IR constraint is satisfied.

Remark $\beta^* > \dfrac{C}{\kappa\,\Delta V}$, since $f(\dfrac{C}{\kappa\,\Delta V}) < 0$.

Proof

$$\begin{aligned}
f(\frac{C}{\kappa\,\Delta V}) &= (\kappa^2\,\Delta V^2)(\frac{C}{\kappa\,\Delta V})^3 + 2\kappa\,\Delta V(V'-\tfrac{V}{2}-C)(\frac{C}{\kappa\,\Delta V})^2 + \Big((V'-C)^2 - V(V'-C) - 2\kappa C\,\Delta V\Big)(\frac{C}{\kappa\,\Delta V}) + 2C^2 - CV'\\
&= \frac{C}{\kappa}V' - CV' < 0 \qquad \text{since } \kappa > 1. \qquad \square
\end{aligned} \tag{27}$$

If the attention bias is large enough and the arbitrage capital is limited, that is $\beta > \beta^*$ and $C < \frac{V'}{2}$, then the IR constraint is satisfied. Considering these conditions along with the previous bound on C, we conclude $A_{1S}^* = -C$, if $C \in \big[0, \dfrac{1}{2}(V'-\beta V)\big]$ and $\beta > \beta^*$. ∎

A.3 Proof of Corollary 7.2

Proof Part 1) From the remark A.2 in the proof of proposition 7.1, we know $\beta^* > \frac{C}{\kappa\,\Delta V}$. So if $\beta > \beta^*$, we have $\beta > \frac{C}{\kappa\,\Delta V} \Rightarrow \beta\kappa\,\Delta V - C > 0 \Rightarrow p_1^* > V$. ∎

Part 2) $p_1^* > p_2^* \Rightarrow \frac{p_2^*}{p_1^*} < 1$. Substituting for p_1^* and p_2^* using equations (7) and (8), this inequality is equivalent to

$$\begin{aligned} &\frac{V' - \beta V - 2C + \beta(V' + \beta\kappa\,\Delta V - C)}{V' + \beta\kappa\,\Delta V - 2C} < 1 \\ \Leftrightarrow\quad &\beta < \frac{(\kappa-1)\Delta V + C}{\kappa\,\Delta V} \end{aligned} \tag{28}$$

Let $\bar{\beta} \equiv \frac{(\kappa-1)\Delta V + C}{\kappa\,\Delta V}$. According to proposition 7.1, $\beta < \hat{\beta}$ for the selling strategy to be the dominant strategy. To complete the proof, it is sufficient to show $\bar{\beta} > \hat{\beta}$. Using equations (**??**) and (28), and simplifying, the sufficient condition is equivalent to

$$\begin{aligned} &\frac{(\kappa-1)\Delta V + C}{\kappa\,\Delta V} > \frac{2\kappa\Delta V + 3V - V' + 5C - \sqrt{(2\kappa\Delta V + 3V - V' + 5C)^2 - 4\kappa\,\Delta V(V' - 8C)}}{2\kappa\,\Delta V} \\ \Leftrightarrow\quad & 16C^2 + \Big((52\kappa - 16)\Delta V + 8V\Big)C + \Big((2\kappa-1)\Delta V + 2V\Big)^2 + \Big(\Delta V + 2V\Big)^2 > 0 \end{aligned} \tag{29}$$

Since $C > 0$, the last inequality in (29) always holds. ∎

Part 3) Substitute for p_1^* and p_2^* using equations (7) and (8), respectively.

$$\begin{aligned} & p_1^* - p_2^* > p_2^* - V' \\ \Leftrightarrow\quad & 2V' + \beta\kappa\,\Delta V - C - 2 \times \frac{\Big(V' - \beta V - 2C + \beta(V' + \beta\kappa\,\Delta V - C)\Big)(V' + \beta\kappa\,\Delta V - C)}{V' + \beta\kappa\,\Delta V - 2C} > 0 \\ \Leftrightarrow\quad & (\beta^2\,\Delta V^2 - 2\beta^3\,\Delta V^2)\kappa^2 + \Big(\beta\,\Delta V \times (V' + 2C + 2\beta V + 2\beta C)\Big)\kappa + 2(V' - c)(C + \beta C - \beta\,\Delta V) > 0 \end{aligned} \tag{30}$$

Given the ranges for β and C, the coefficients of κ^2 and κ in the last inequality in (30) are always positive. Therefore, it is sufficient to show that the last expression is positive.

$$C + \beta C - \beta\,\Delta V > 0 \Leftrightarrow \beta < \frac{C}{\Delta V - C} \tag{31}$$

Notice that $\frac{C}{\Delta V - C} > \hat{\beta}$ according to equation (19), Hence inequality (31) always holds. ∎

A.4 Proof of Corollary 7.3

Proof The first part, about the date 1 price, is straightforward by taking derivative of p_1^* with respect to β.

To show the second part, by simplifying equation (8) we get

$$p_2^* = \frac{(\kappa^2\,\Delta V^2)\beta^3 + 2\kappa\,\Delta V(V' - \frac{V}{2} - C)\beta^2 + (\kappa\,\Delta V(V' - 2C) + (V' - C)(V' - \nu - C))\beta + (V' - C)(V' - 2C)}{(\kappa\,\Delta V)\beta + (V' - 2C)} \tag{32}$$

Taking derivative with respect to β and simplifying leaves us with

$$\frac{\partial p_2^*}{\partial \beta} = \frac{(2\kappa^3\,\Delta V^3)\beta^3 + \left(2\kappa^2\,\Delta V^2(V' - \frac{V}{2} - C) + 3\kappa^2\,\Delta V^2(V' - 2C)\right)\beta^2 + \left(4\kappa\,\Delta V(V' - \frac{V}{2} - C)(V - 2C)\right)\beta + (V' - 2C)\left(C^2 - (V' + (\kappa+1)\Delta V)C + V'\,\Delta V\right)}{\left((\kappa\,\Delta V)\beta + (V' - 2C)\right)^2} \tag{33}$$

Notice that we want to show that if $C \in [0, C^*]$ and $\beta \in [\beta^*, \hat{\beta}]$, then $\frac{\partial p_2^*}{\partial \beta} > 0$. Notice that the numerator in equation (33) is increasing in β. Therefore it is sufficient to show $\frac{\partial p_2^*}{\partial \beta}|_{\beta=\beta^*} > 0$. Substituting for β^* from equation (26), the numerator in equation (33) is equal to

$$\kappa^2\,\Delta V^2(V' + V - 2C)(\beta^*)^2 + \left(2\kappa\,\Delta V((V' - C)^2 - CV')\right)\beta^* + (V' - 2C)\left(C^2 - (V' + (\kappa - 1)\Delta V)C + V'\Delta V\right) > 0 \tag{34}$$

Which is always positive for $C \in [0, C^*]$. $\blacksquare$

A.5 Proof of Proposition 7.5

Proof Because retail traders do not sell short, there are only two demand forces in this case, the demand from liquidity traders and from arbitrageurs. Therefore, adjusting equation (6) to the case of negative news shocks, the date 1 market clearing condition is

$$H_1^L + H_1^A = 1 \tag{35}$$

Also, adjusting equation (10) to this case, date 2 market clearing conditions in the two cases are

$$\begin{cases} H_2^L + H_2^A = 1 + H_{1B}^A & \text{if buying strategy at } t = 1 \\ H_2^L + H_2^A = 1 & \text{if selling strategy at } t = 1 \end{cases} \tag{36}$$

Substituting the passive demand and the arbitrageur's demand into equation (35) and solving for date 1 price, we have $p_1 = V + A_1$.

The arbitrageur's profit if she adopts the selling strategy, is calculated as in equation (11). Using the market clearing condition under the selling strategy and substituting for demands and date 1 price, date 2 price this case is $p_{2S} = \frac{(V - 2C)(V + A_{1S})}{(V + 2A_{1S})}$.

The arbitrageur's optimal selling strategy solves the same problem as in (13). Lemma A.2 shows the arbitrageur's optimal selling strategy in response to bad news.

Lemma A.2 *In response to a negative shock, the optimal selling strategy for the arbitrageurs at date 1 is* $A^*_{1S} = -C$, *i.e. the capital constraint is binding, if the arbitrage capital is limited, that is* $C \in \left[0, \min\{\frac{V}{2}, -\Delta V\}\right]$.

Proof See Appendix A.6.

However, the arbitrager is able to adopt buying strategy too. Her profit, if she buys at day 1, is the same as (15). By substituting the demands in the market clearing condition under the buying strategy in (36), date 2 price is calculated as $p_{2B} = \dfrac{(V - 2C + 2A_{1B})(V + A_{1B})}{(V + 3A_{1B})}$.

Our goal is to find the optimal strategy for the arbitrageur at date 1. Sketch of the proof here is similar to section A.1 where we find the optimal strategy in response to positive shocks. We use the optimal selling strategy from lemma A.2. We must compare it to the optimal buying strategy to determine the arbitrageur's dominant strategy. However, to avoid cumbersome algebra, we find an upper bound for the profits of the buying strategies and compare it to the profit from the optimal selling strategy. Remind from equation (15) that a buying strategy's profit is $\Pi_B = \Big(2C + A_{1B}(\frac{p_{2B}}{p_{1B}} - 2)\Big)(1 - \frac{V'}{p_{2B}})$, where p_{1B} and p_{2B} are as stated in the text.

The upper bound for the expression in the first bracket is $2C$. Moreover, this expression is decreasing in A_{1B}, since $\frac{\partial}{\partial A_{1B}}\big(\frac{p_{2B}}{p_{1B}}\big) < 0$ and $\big(\frac{p_{2B}}{p_{1B}} - 2\big) < 0$. However, the expression in the second bracket is increasing in A_{1B} since

$$\frac{\partial}{\partial A_{1B}}\Big(1 - \frac{V'}{p_{2B}}\Big) = \frac{\partial}{\partial p_{2B}}\Big(1 - \frac{V'}{p_{2B}}\Big) \times \frac{\partial p_{2B}}{\partial A_{1B}} = \frac{\partial}{\partial p_{2B}}\Big(1 - \frac{V'}{p_{2B}}\Big) \times \Big(\frac{2VA_{1B}}{(V + 3A_{1B})^2}\Big) > 0 \quad (37)$$

Now consider the profit from the optimal selling strategy. Using equation (11), we can calculate it as

$$\Pi^*_S = \Pi_S(A^*_{1S}) = 2C\Big(1 - \frac{V'}{p^*_{2S}}\Big) \quad (38)$$

Comparing Π_B and Π^*_S, It is clear that $\Big(2C + A_{1B}(\frac{p_{2B}}{p_{1B}} - 2)\Big) \leqslant 2C$. Therefore, if, for any buying strategy $A_{1B} \in [0, C]$, Π_B were to be larger than Π^*_S, then the second bracket in Π_B should be larger than the second bracket in Π^*_S, that is

$$\begin{aligned}
& 1 - \frac{V'}{p_{2B}} > 1 - \frac{V'}{p^*_{2S}} \\
\Leftrightarrow\quad & p_{2B} > p^*_{2S} \\
\Leftrightarrow\quad & \frac{2(A_{1B})^2 + (3V - 2C)A_{1B} + V(V - 2C)}{V + 3A_{1B}} > V - C \\
\Leftrightarrow\quad & A_{1B} > \frac{-C + \sqrt{C^2 + 8CV}}{4}
\end{aligned} \quad (39)$$

From the bounds on C in lemma A.2 we know $C < \frac{V}{2}$. So we can replace V by 2C in the last line of (39) and still have the inequality preserved. Therefore, we have

$$A_{1B} > \frac{-C+\sqrt{C^2+8CV}}{4} > \frac{-C+\sqrt{C^2+8C\times(2C)}}{4} = \frac{-C+\sqrt{17C^2}}{4} > \frac{C}{2} \tag{40}$$

This means all buying strategies that dominate the optimal selling strategy, if any, must lie in $\left[\frac{C}{2}, C\right]$. Having this more restricted domain for A_{1B}, now we are able to calculate an upper bound for Π_B. Since the first bracket in Π_B is decreasing in A_{1B}, it achieves it maximum at $A_{1B} = \frac{C}{2}$. On the contrary, the second bracket in Π_B is increasing in A_{1B}, it achieves it maximum at $A_{1B} = C$. Plugging these values of A_{1B} in Π_B, the upper bound is

$$\begin{aligned} UB(\Pi_B) &\equiv \Big(2C + \frac{C}{2}\big(\frac{V-2C+C}{V+\frac{3}{2}C} - 2\big)\Big)\Big(1 - \frac{V'\times(V+\frac{3}{2}C)}{(V-2C+C)(V+\frac{C}{2})}\Big) \\ &= \Big(C\times\frac{3V+2C}{2V+3C}\Big)\Big(1-\frac{V'\times(2V+3C)}{(V-C)(2V+C)}\Big) \end{aligned} \tag{41}$$

Using equations (38) and (41), we have

$$\begin{aligned} & \Pi_S^* > UB(\Pi_B) \\ \Leftrightarrow \quad & 2C(1-\frac{V'}{p_{2S}^*}) > \Big(C\times\frac{3V+2C}{2V+3C}\Big)\Big(1-\frac{V'\times(2V+3C)}{(V-C)(2V+C)}\Big) \\ \Leftrightarrow \quad & 2(1-\frac{V'}{p_{2S}^*}) > \underbrace{(\frac{3V+2C}{2V+3C})}_{>1}\Big(1-\frac{V'\times(2V+3C)}{(V-C)(2V+C)}\Big) \\ \Leftrightarrow \quad & 2(1-\frac{V'}{p_{2S}^*}) > \Big(1-\frac{V'\times(2V+3C)}{(V-C)(2V+C)}\Big) \\ \Leftrightarrow \quad & C^2 + (V-V')C - 2V(V-V') < 0 \\ \Leftrightarrow \quad & C < \frac{-(V-V')+\sqrt{(V-V')^2+8V\times(V-V')}}{2} < -\Delta V \end{aligned} \tag{42}$$

The last inequality always hold so long as we have the limited arbitrage capital, that is $C \in \left[0, \min\{\frac{V}{2}, -\Delta V\}\right]$ as stated in lemma A.2. Hence, with this restriction on the arbitrage capital the optimal selling strategy, $A_{1S}^* = -C$, is always the dominant strategy for the arbitrageur. $\blacksquare$

A.6 Proof of Lemma A.2

Proof Sketch of the proof is similar to what we did before. The arbitrageur solves the program in (13). First, we solve the problem without the IR constraint and then add the necessary conditions to satisfy the IR constraint to the solution.

$$\frac{\partial \Pi_S}{\partial A_{1S}} = \frac{\partial \Pi_S}{\partial p_{2S}} \times \frac{\partial p_{2S}}{\partial A_{1S}} = \frac{2CV'}{(p_{2S})^2} \times \frac{-V\times(V-2C)}{(V+2A_{1S})^2} \tag{43}$$

Therefore, $\dfrac{\partial \Pi_S}{\partial A_{1S}} < 0, \quad \text{if } C < \dfrac{V}{2}.$
The last step is to find conditions under which the IR constraint is satisfied and consider them in sketching the final solution. The IR constraint is satisfied if

$$\begin{aligned} & \Pi_S^* = \Pi_S(A_{1S}^*) = 2C(1 - \frac{V'}{p_{2S}(A_{1S}^*)}) \geqslant 0 \\ \Leftrightarrow \quad & p_{2S}(A_{1S}^*) \geqslant V' \\ \Leftrightarrow \quad & \frac{(V-2C)(V-C)}{(V-2c)} \geqslant V' \\ \Leftrightarrow \quad & C \leqslant V - V' = -\Delta V \end{aligned} \tag{44}$$

Therefore, we conclude that if the arbitrage capital is limited, that is $C < \min\{\frac{V}{2}, -\Delta V\}$, then the optimal selling strategy is $A_{1S}^* = -C$. ∎

A.7 Proof of Corollary 7.7

Proof Given the assumptions about the liquidity shocks, the liquidity traders' stock holding is

$$H_t^L = \begin{cases} \frac{V'}{p_t} & t = 1 \\ \frac{V}{p_t} & t > 1 \end{cases} \tag{45}$$

As we discussed, the arbitrageur and retail investors do not trade on date 1. Therefore, using the market clearing condition in equation (6), the date 1 price is $p_1^* = V'$.

At date 2, however, the previous day's price jump attracts retail investors. Same as before, they only trade in response to positive price jumps and their demand will be $\Delta V + S_t$. Retail investors' date 2 demand shock (S_2) depends on the sign of the price shock at date 1: $\Delta p_1 = p_1^* - p_0 = V' - V$. Retail investors' stock holding at date 2 can be summarized as in equation (46).

$$H_2^R = \begin{cases} 0 & V' < V \\ \beta \Delta p_1 & V' > V \end{cases} \tag{46}$$

As expected, their stock holding is zero when the price shock is negative. In the case of positive shocks, since there is no change in the fundamental value with the no-news shocks ($\Delta V = 0$), retail investors' demand at date 2 is only due to S_t.

The date 3 price will be equal to the fundamental value (V) due to unlimited arbitrage capital. The arbitrageur chooses an optimal strategy to maximize overall profits. From the market clearing condition, date 2 price can be computed as

$$p_2 = V + A_2 + \mathbb{1}\{V' > V\} \times \beta \Delta p_1 \tag{47}$$

Basically date 2 price is determined by demands of liquidity traders (V), the arbitrageur (A_2), and retail investors ($\mathbb{1}\{V' > V\} \times \beta \Delta p_1$).

If the date 1 price shocks is negative ($V' < V$), the arbitrageur's optimal strategy is not to trade at date 2. Because any trade by the arbitrageur forces the price to deviate from the fundamental value and the arbitrageur loses money at date 3 when closing her trades at the fundamental value. However if the date 1 price shock is positive ($V' > V$) and the arbitrage capital is limited ($C \in \left[0, \beta\frac{\Delta V}{2}\right]$), the arbitrageur's optimal strategy is to sell. The upper bound in the capital constraint guaranties that the price remains above fundamental value at date 2. So the arbitrageur's optimal strategy at date 2 is summarized as:

$$A_2^* = \begin{cases} 0 & V' < V \\ -2C & V' > V \end{cases} \tag{48}$$

Substituting this optimal strategy into equation (47), the equilibrium date 2 price is

$$p_2^* = \begin{cases} V & V' < V \\ V - 2C + \beta \times (V' - V) & V' > V \end{cases} \tag{49}$$

As equation (49) suggests, in the case of date 1 negative shock, the price is back to fundamentals at date 2. In the case of date 1 positive shock, under the limited arbitrage capital, the price declines after the shock, and it reaches the fundamental value at date 3. ∎