The Dark Side of Circuit Breakers

Hui Chen Anton Petukhov Jiang Wang Hao Xing*

February 27, 2023

Abstract

Market-wide trading halts, also called circuit breakers, have been widely adopted as part of the stock market architecture, in the hope of stabilizing the market during dramatic price declines. We develop an intertemporal equilibrium model to examine how circuit breakers impact market behavior and welfare. We show that a circuit breaker tends to lower the overall level of the stock price and significantly alters its dynamics. In particular, as the price approaches the circuit breaker, its volatility rises drastically, accelerating the chance of triggering the circuit breaker – the so-called "magnet effect"; in addition, returns exhibit increasing negative skewness and positive drift, while trading activity spikes up. Our empirical analysis finds supportive evidence for the model's predictions. Moreover, we show that a circuit breaker can affect the overall welfare either negatively or positively, depending on the relative significance of investors' trading motives for risk sharing vs. irrational speculation.

*Chen: Sloan School of Management at MIT and NBER; Petukhov: Citadel; Wang (corresponding author): Sloan School of Management at MIT, CAFR and NBER; Xing: Questrom School of Business at Boston University. We thank Daniel Andrei, Doug Diamond, Jennifer Huang, Leonid Kogan, Pete Kyle, Hong Liu, Lubos Pastor, Steve Ross, Liyan Yang, and seminar participants at Boston University, INSEAD, MIT, Peking University, PBCSF, University of Chicago, NBER Asset Pricing meeting, Minnesota Asset Pricing Conference, NYU Shanghai Volatility Institute Conference, Hanqing Summer Finance Workshop, CICF and WFA for comments.

1 Introduction

Stock market crashes in the absence of clear macroeconomic causes raise questions about the confidence in the financial market from market participants, policy makers, and the general public alike. While the mechanisms behind these sudden price drops are still not well understood, various measures have been adopted to intervene in the trading process in the hope of stabilizing prices and restoring market order. These measures, sometimes referred to as "throwing sand in the gears," range from market-wide trading halts, price limits, to restrictions on order flows, positions, margins, and even transaction taxes, just to name a few.[1] They have grown to be an important part of the overall market architecture. Yet, the merits of these measures, from either a theoretical or an empirical perspective, remain largely unclear (see, for example, Grossman, 1990).

Arguably the most prominent among these measures is the market-wide circuit breaker (MWCB), which was first introduced in the U.S. in 1988 after the 1987 Black Monday stock market crash.[2] It temporarily halts trading in all stocks and related derivatives when a designated market index drops by a pre-specified amount during a trading session. Since then, circuit breakers of various forms have been adopted around the globe.[3] Table 1 shows the adoption of market-wide circuit breakers (MWCB) and price limits among the leading stock markets in both the developed and developing economies. The U.S. MWCB was first triggered on October 27, 1997, which led to its redesign. It then stayed untouched (including during the "Flash Crash" of 2010) until March 2020, when it was triggered four times in a span of two weeks at the onset of the COVID-19 pandemic. Following the turbulent stock market declines in 2015, China introduced its MWCB in January 2016. After being triggered on the first day of its installment and again in the same week, it was immediately abolished.

These recent events have revived the debate about circuit breakers and market interventions

[1] Contingent trading halts and price limits are part of the normal trading process for individual stocks and futures contracts. But the motivations behind them vary. For example, the trading halt of an individual stock prior to major corporate announcements is motivated by the desire for fair information disclosure, while daily price limits on futures are motivated by the desire to guarantee the proper implementation of the mark-to-market mechanism as well as to deter market manipulation. In this paper, we focus on market-wide trading halts in the underlying markets such as stocks as well as their derivatives, which have very different motivations and implications.

[2] See the Online Appendix (OA.1) for a brief history of the MWCB mechanism in the U.S.

[3] According to a 2016 report, "Global Circuit Breaker Guide" by ITG, over 30 countries around the world have rules of trading halts in the form of circuit breakers, price limits and volatility auctions.

Table 1: Adoption of market-wide circuit breakers (MWCB) among leading stock markets by market capitalization (source: World Bank). The table also reports markets with price limits on individual stocks. Markets with market-wide circuit breakers as well as individual stock price limits are denoted by Y and N otherwise. Y/N for China denotes its adaptation of circuit breakers and then their abandonment.

Year 2020	Market Cap (Tn \$)	Market Cap Rank	Circuit Breaker	Price Limit
Developed markets				
United States	40.7	1	Y	Y
Japan	6.7	3	N	Y
Hong Kong	6.1	4	N	N
United Kingdom	3.6	5	Y	Y
France	2.8	6	Y	N
Canada	2.6	7	Y	N
Developing markets				
China (mainland)	12.2	2	Y/N	Y
India	2.6	8	Y	Y

in general. What are the goals of circuit breakers and other forms of trading brakes? How do they actually impact the market? How to assess their success or failure? How may their effectiveness depend on the particular market and its participants, the actual design, and the specific market conditions? More broadly, the same questions can be raised about any form of intervention in the normal trading process.

In this paper, we develop an intertemporal equilibrium model in which investors trade either to share risks or to speculate on their own beliefs. We examine how the introduction of a circuit breaker changes investors' trading and equilibrium price behavior. We show that a circuit breaker in general lowers the overall price level. More importantly, it substantially alters the price dynamics. In particular, we show that as the market drops closer to the circuit breaker, the conditional price volatility rises sharply. In addition, returns exhibit negative skewness with an increasing magnitude. Both effects reflect the fact that a circuit breaker tends to destabilize the price during large market declines. Consequently, as the market falls closer to the circuit breaker, the likelihood of a circuit breaker trigger rises at an accelerated pace. This is the so-called "magnet effect" often suspected by market participants. Our model also predicts that expected return and trading volume tend to increase as the market approaches the circuit breaker. We find supporting evidence for these predictions using transaction-level data from E-mini S&P 500 futures.

We also use the model to examine the welfare implications of a circuit breaker, which critically depends on investors' trading motives. If they are trading primarily for risk sharing, the introduction of a circuit breaker reduces overall welfare. If, however, they are trading mainly to speculate on their irrational beliefs (e.g., due to panic at times of market turmoil), a circuit breaker can improve overall welfare under the objective probability measure. When both trading motives are present, there will be a trade-off between the two effects.

In our model, the financial market consists of a stock and a bond. There are two (classes of) heterogeneous investors in the market. The heterogeneity can take two different forms: it can be in the investors' beliefs about stock fundamentals (payoffs), or in their utility function, which can be state-dependent. Under our formulation, these two forms of heterogeneity are mathematically equivalent in yielding the equilibrium market behavior despite the differences in their economic interpretations and welfare implications. For simplicity, we will adopt the heterogeneous beliefs interpretation in our exposition and return to the alternative interpretation in the welfare analysis.

In particular, we assume that the two investors have heterogeneous beliefs about the stock's future payoff. One investor's belief is set to be the same as the objective belief, while the other investor's belief is different and will thus be referred to as the irrational investor. They trade competitively in the financial market.

Without the circuit breaker, the market is dynamically complete, and both investors trade continuously to achieve the efficient allocation under their own beliefs. With the opportunity of continuous portfolio rebalancing, the relatively more optimistic investor is willing to take on much bigger stock positions relative to the more pessimistic investor.

The introduction of a circuit breaker fundamentally changes investors' trading behavior and the equilibrium price. We start by considering the market equilibrium at the triggering point of the circuit breaker, which occurs after a series of negative shocks. The trading halt forces the more optimistic investor to hold her stock position for an extended period without the opportunity to rebalance in response to new shocks. Such extreme illiquidity substantially reduces her willingness to hold the stock upon market closure. As a result, the stock price has to drop drastically at market closure in order to entice the pessimistic investor to absorb more stock shares, disproportionately raising his importance in determining the prices.

Next, consider the price behavior near the circuit breaker. No arbitrage requires that the stock price be continuous over time. Thus, the low price level at market closure and its continuity imply that the stock price will be dropping drastically as the market moves closer to the circuit breaker threshold. The sharp price decline in response to lowering fundamentals corresponds to heightened price sensitivity to fundamental shocks, which causes the conditional price volatility to rise significantly.

The mechanism above leads to the following predictions from our model. (1) The presence of a circuit breaker lowers the stock price overall relative to its level without the circuit breaker. (2) The (conditional and realized) price volatility increases *at an increasing rate* as the market moves closer the circuit breaker. (3) Due to the negative correlation between volatility and price, the realized return skewness also turns negative as as the market moves closer the circuit breaker. (4) The conditional expected return rises as the market approaches the circuit breaker and the price level drops. This is simply because the price drop is mostly due to the increasing influence of the more pessimistic investor instead of changes in fundamentals (e.g., expected future payoffs). (5) The unwinding of the more optimistic investor's stock positions when approaching the circuit breaker results in a fast increase in trading volume. These price and volume behaviors are in sharp contrast to those without the circuit breaker, which are relatively stable around the same price level.

We also show that the changes in price dynamics in the presence of the circuit breaker gives rise to the so-called "magnet effect." That is, when the price approaches the circuit breaker threshold, the likelihood of hitting the circuit breaker increases substantially compared to under normal conditions. This is because the rising price volatility in the neighborhood of the circuit breaker greatly accelerates the chance of actually reaching it.

We further explore the model's predictions empirically. By their purpose and design, circuit breakers are rarely hit. Thus, it is hard to assess their impact by relying purely on the actual triggering events themselves. Instead, we take advantage of the dynamic nature of our model and examine its unique predictions on the behavior of prices (including realized volatility, skewness, and average return) and volume as the market approaches the circuit breaker (predictions (2) through (5) as laid out above), without necessarily hitting it.

We use transaction-level data for the E-mini S&P 500 futures from 2013 to 2020 to construct the volatility, skewness, return, and trading volume measures and then run piecewise linear regressions on a measure of the distance to circuit breaker (DTCB) while controlling for the leverage effect, a time trend, intraday seasonality, as well as lagged dependence of the dependent variables. The empirical results are overall consistent with our model's predictions. In particular, for return volatility, the unconditional regression coefficient on DTCB is -17.5, implying that a drop in DTCB by 1% (7% is the Level-1 trigger level) leads to an increase in volatility of 17.5 bps (normalized to daily scale). However, when DTCB is within the range of 2% and below, the regression coefficient nearly doubles to -30.0, reflecting an accelerating rise in return volatility as DTCB drops closer to 0. These results reflect the destabilizing influence of the circuit breaker on price in its neighborhood and the magnet effect.

While our model focuses on circuit breakers, it can be extended to study other forms of interventions in the trading process such price limits, trading restrictions, and other forms of market freezes and slowdowns. The underlying mechanism driving the impact of a circuit breaker also applies to these situations.

Related Literature

Prior theoretical work on circuit breakers focus on their role in restoring orderly trading and reducing "excess" volatility in a market with various microstructure imperfections, such as limited market participation, non-synchronized trading, asymmetric information. This is in part motivated by the apparent break-down in the trading process during the 1987 market crash. For example, Greenwald and Stein (1991) argue that, in a market with limited participation and the resulting execution risk, circuit breakers can help to better synchronize trading for market participants and improve the efficiency of allocations (see also Greenwald and Stein, 1988).[4] Subrahmanyam (1994) shows that in the presence of partial participation/optimization and asymmetric information, circuit breakers can increase ex ante price volatility when investors with fixed orders shift their trades to earlier periods with lower liquidity supply (see also Subrahmanyam, 1995). Using a setting similar to Greenwald and Stein (1991), Kodres and O'Brien (1994) show that circuit breakers can reduce the welfare loss from the initially imperfect trading process, at least for some of the market participants.

[4] In Greenwald and Stein (1991), limited participation takes several forms. In particular, value traders, who act as price stabilizers, enter the market at different times with uncertainty. This uncertainty in their participation, which is assumed to be exogenous, gives rise to the additional risk in execution prices. Also, these value traders can only rely on market orders or simple limit orders, rather than limit order schedules, in their trading.

A common starting point of these work is a trading process with major imperfections in a noisy rational expectations setting. Leaving noise trades outside the model, these models do capture important aspects of the market, but they are partial equilibrium in nature.

By developing a general equilibrium model in an intertemporal setting, we achieve the following: First, we properly capture investors' most basic trading needs, risk sharing and speculation, and their resulting trading behavior, with and without circuit breakers. Hence, the resulting impact of a circuit breaker is a full equilibrium outcome. Second, a general equilibrium model captures the welfare of all market participants, and hence allows us to examine the full extent of a circuit breaker's welfare impact. Third, our intertemporal setting yields unique predictions on how the circuit breaker changes price dynamics, especially how it destabilizes the market as the price approaches it and gives rise to the magnet effect. Moreover, these predictions are testable even without the circuit breaker being ever triggered.[5] In addition, our model also provides a basis to further include other forms of imperfections beyond irrational speculation, such as costs of participation, failure of coordination, asymmetric information and strategic behavior, which may be relevant to capture and quantify more fully the cost and benefit of circuit breakers. Nonetheless, focusing solely on their marginal influence may understate the fundamental merits of the market mechanism itself.[6]

The empirical work on market-wide circuit breakers is scarce due to the fact that their likelihood to be triggered is very small by design. Goldstein and Kavajecz (2004) provide a detailed analysis on the behavior of market participants in the period around October 27, 1997, the only time the U.S. circuit breaker has been triggered since its introduction until very recently. They find that leading up to the trading halt, market participants accelerated their trades. In addition, they show that sellers' behavior is less influenced when approaching circuit breaker than the buyers', who are withdrawing from the market by canceling their

[5]The predictions of our model on price volatility is different from those from Subrahmanyam (1994) in nature. Apart from the differences in modeling choices, such as general equilibrium vs. noisy rational expectations equilibrium and mild imperfections vs. more severe imperfections, Subrahmanyam (1994)'s results are about the ex ante price volatility at the trading halt point while ours are about the dynamics of volatility when the market approaches the circuit breaker. While there are many channels a market intervention can change the overall price volatility, our predictions on volatility dynamics are distinctive in the presence of circuit breakers and directly testable, which we further confirm from the data.

[6]In this spirit, our paper is closely related to Hong and Wang (2000), who study the effects of periodic market closures in the presence of asymmetric information. The liquidity effect caused by market closures as we see here is qualitatively similar to what they find. By modeling the stochastic nature of a circuit breaker, we are able to fully capture its impact on market dynamics, such as volatility and skewness.

buy limit orders. These patterns are all consistent with what our model predicts: buyers (more optimistic investors) are cutting back from the market and sellers (more pessimistic investors) are becoming the marginal traders when the circuit breaker is approaching.[7]

Our model shows that the presence of a circuit breaker will lead to unique price dynamics in its neighborhood. In particular, the model produces testable predictions about the dynamic behavior of return moments for a wide range of prices without the circuit breaker being triggered. Such an approach clearly demonstrates the power of an effective model in guiding our empirical analysis. It is also confirmed by our empirical results.

A related empirical literature is on the impact of conditional trading restrictions on individual securities including futures. For example, Bertero and Mayer (1990) and Roll (1988, 1989) study the effects of trading halts based on price limits imposed on individual stocks around the 1987 stock market crash while finding different results.[8] Although the focus of our paper is on market-wide circuit breakers, the results we obtain are broadly compatible with the empirical findings on the impact of trading halts for individual assets.

The rest of the paper is organized as follows. Section 2 describes the basic model for our analysis. Section 3 provides the solution to the model. In Section 4, we examine how a circuit breaker changes investor behavior and equilibrium price dynamics and derive testable predictions on the circuit breaker's unique impact on price and volume dynamics. Section 5 considers the circuit breakers' welfare implications. In Section 6, we empirically examine the model's predictions on price and volume dynamics. Section 7 presents additional discussion on the model's robustness and possible extensions. Section 8 concludes. The appendix contains the key steps of the proofs. An online appendix further provides additional institutional details on circuit breakers, details of the proofs, additional results from the model, and the robustness of the empirical results.

[7] Ackert, Church, and Jayaraman (2001) study the impact of market-wide circuit breakers through experiments. They find that circuit breakers do not impact prices significantly but alter market participants' trading behavior substantially by accelerating trading when the breakers are approaching.

[8] Chen (1993), Lauterbach and Ben-Zion (1993), Santoni and Liu (1993), Lee, Ready, and Seguin (1994), Kim and Rhee (1997), Corwin and Lipson (2000), Christie, Corwin, and Harris (2002), Jiang, McInish, and Upson (2009), Gomber et al. (2012), among others, study the effects of trading halts and price limits on the market behavior of individual stocks. Chen et al. (2017) examine the impact of daily price limits on trading patterns and price dynamics in the Chinese stock market. Brennan (1986), Kuserk, Locke, and Sayers (1992), Berkman, Steenbeek et al. (1998), Coursey and Dyl (1990), Ma, Rao, and Sears (1989b), Ma, Rao, and Sears (1989a), Chen and Jeng (1996) study effects of trading restrictions related to price fluctuations in futures markets.

2 The Model

In this section, we present a simple model for circuit breakers as the basis of our analysis. We will provide additional discussions on the model's assumptions later in this section.

We consider a continuous-time endowment economy over a finite time interval $[0,\ T]$. Uncertainty is described by a one-dimensional standard Brownian motion Z, defined on a filtered complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$, where $\{\mathcal{F}_t\}$ is the augmented filtration generated by Z.

Financial Market

There is a single share of an aggregate stock, which pays a terminal dividend of D_T at time T. The process for D is exogenous and publicly observable, given by:

$$dD_t = \mu D_t dt + \sigma D_t dZ_t, \quad D_0 = 1, \tag{1}$$

where $t \in [0,\ T]$ and μ and $\sigma > 0$ are the expected growth rate and the volatility of D_t, respectively.[9] Besides the stock, there is also a riskless bond with zero net supply. Each unit of the bond yields a terminal payoff of 1 at T.

Both the stock and the bond are traded competitively in a financial market. Since there is no intermediate payoff/consumption, we use the riskless bond as the numeraire. Thus, the price of the bond is always 1.[10] Let S_t denote the price of the stock at t (cum-dividend).

Agents

There are two agents, A and B, who are initially endowed with zero units of the bond and ω and $1-\omega$ shares of the stock, respectively, with $0 \leqslant \omega \leqslant 1$ determining the initial wealth distribution between the two agents.

Both agents can trade in the market. Let θ_t^i and ϕ_t^i denote the stock and bond holdings of agent i at t, $i = A, B$, respectively. We impose the usual restrictions on trading strategies

[9] For simplicity, throughout the paper we will refer to D_t as "dividend" and S_t/D_t as the "price-dividend ratio", even though the dividend will only be paid at T. More generally, D_t can be understood as the expectation of D_T at t.

[10] Since there is no intermediate consumption in the model, it is natural to use the bond price as the numeraire. This modeling choice is also motivate by the problem at hand, which is about trading halt during an otherwise continuous trading day. Given that a trading day is a rather short horizon, we would expect the bond price or the interest rate to stay mostly constant. In this case, using the bond price as the numeraire is a reasonable approximation. Under this numeraire, the bond price stays at 1.

to rule out arbitrage. Agent i's wealth is then given by $W_t^i = \phi_t^i + \theta_t^i S_t$, and

$$dW_t^i = \theta_t^i dS_t, \tag{2}$$

with W_T^i being her/his terminal wealth.

We assume that the agents have preferences in the form of expected utility over their terminal wealth. They choose trading strategies to maximize their own expected utilities. For tractability, we further assume that both agents' utility function takes the logarithmic form:

$$u^i(W_T^i) = \ln(W_T^i), \quad i = A,\, B. \tag{3}$$

The two agents have different beliefs about the terminal dividend D_T and they "agree to disagree" (i.e., they do not learn from each other or from prices).[11] Agent A has the objective belief in the sense that her belief measure is consistent with $\mathbb{P}$, the physical probability measure. In particular, $\mu^A = \mu$. Agent B's belief measure, denoted by $\mathbb{P}^B$, however, is different from but equivalent to $\mathbb{P}$.[12] In particular, he believes that the dividend growth rate at time t is given by:

$$\mu_t^B = \mu + \delta_t, \tag{4}$$

where δ follows an Ornstein-Uhlenbeck process:

$$d\delta_t = -\kappa \left(\delta_t - \bar{\delta}\right) dt + \nu \, dZ_t, \tag{5}$$

with $\kappa \geqslant 0$ and $\nu \geqslant 0$. (5) then describes the dynamics of the gap between agent B's belief and the physical probability measure, which is the same as agent A's belief.

Notice that δ_t is driven by the same Brownian motion as the dividend. With $\nu > 0$, agent B becomes more optimistic (pessimistic) following positive (negative) shocks to the dividend, and the impact of these shocks on his belief decays exponentially at the rate κ. Thus, the parameter ν controls how sensitive B's conditional belief is to realized dividend shocks, while κ determines the relative importance of shocks from recent past vs. distant past. The average long-run disagreement between the two agents is $\bar{\delta}$. In the special case where $\nu = 0$ and $\delta_0 = \bar{\delta}$, the disagreement between the two agents stays constant over time. In another special

[11] The formulation here follows the earlier work of Detemple and Murthy (1994) and Zapatero (1998), among others.

[12] More precisely, $\mathbb{P}$ and $\mathbb{P}^B$ are equivalent when restricted to any σ-field $\mathcal{F}_T = \sigma(\{D_t\}_{0 \leqslant t \leqslant T})$. Two probability measures are equivalent if they agree on zero probability events. Agents beliefs should be equivalent to prevent seemingly arbitrage opportunities under any agents' beliefs.

case where $\kappa = 0$, δ_t follows a random walk.

Given the two agents' beliefs, $\mathbb{P}$ and $\mathbb{P}^B$, respectively, let η be the Radon-Nikodym derivative of $\mathbb{P}^B$ with respect to $\mathbb{P}$. From Girsanov's theorem, we then have:

$$\eta_t = \exp\left(\frac{1}{\sigma}\int_0^t \delta_s dZ_s - \frac{1}{2}\frac{1}{\sigma^2}\int_0^t \delta_s^2 ds\right). \tag{6}$$

Intuitively, agent B will be more pessimistic than A when $\delta_t < 0$; in that case, those paths with high realized values for $\int_0^t \delta_s dZ_s$, which appears after a sequence of negative shocks to Z, will be assigned higher probabilities under $\mathbb{P}^B$ than under $\mathbb{P}$. Similarly, paths with positive shocks to Z will be assigned higher probabilities under $\mathbb{P}^B$ when $\delta_t > 0$.

Difference in beliefs is a simple way to introduce heterogeneity among agents, which generates trading. The heterogeneity in beliefs can also be interpreted as heterogeneity in utility, with state dependence. In particular, we have:

$$\mathbb{E}^B[u(W_T^B)] = \mathbb{E}\left[\eta_T u(W_T^B)\right] = \mathbb{E}\left[\tilde{u}^B(W_T^B, \eta_T)\right], \quad \tilde{u}^B(W_T^B, \eta_T) \equiv \eta_T \ln(W_T^B). \tag{7}$$

Thus, we can re-interpret the two heterogeneous agents as having the same objective belief but different utility functions. While agent A has the simple logarithmic utility function over her terminal wealth as given in (3), agent B has a state-dependent utility function, $\tilde{u}^B(\cdot)$, as given in (7). Later in this section, we will provide more discussion on the economic meaning behind such a state-dependent utility. Although these two different interpretations give rise to the same market behavior, they can lead to different welfare and policy implications for circuit breaker, which we discuss in Section 5.

Circuit Breaker

To capture the essence of a market-wide circuit breaker, we assume that the market will be closed whenever the stock price S_t hits a threshold $(1-\alpha)S_0$, where S_0 is the endogenous initial price of the stock, and $\alpha \in [0, 1]$ is a constant parameter determining the floor of downside price fluctuations during the interval $[0, T]$. For simplicity, we assume that the market will remain closed until T after the circuit breaker is triggered.

In practice, the circuit breaker threshold is typically based on the closing price of the previous trading day instead of the opening price of the current trading session. However, the distinction between today's opening price and the prior day's closing price is not crucial for our analysis. The circuit breaker not only depends on but also endogenously affects the initial stock price, just like it does for prior day's closing price in practice. We provide more details on the actual operation of circuit breakers towards the end of this section.

Market Equilibrium

We now define the market equilibrium in the presence of a circuit breaker. No arbitrage requires that the stock price process is continuous. Let τ denote the time when the circuit breaker is triggered, i.e., when S_t first hits $(1-\alpha)S_0$. Hence, τ is given by:

$$\tau = \inf\{\, t \geqslant 0 : S_t \leqslant (1-\alpha)S_0 \,\}, \quad \alpha \in [0, 1]. \tag{8}$$

Let $\tau \wedge T$ to denote $\min\{\tau, T\}$. We then have:

Definition 1 (Equilibrium with Circuit Breaker)**.** *In the presence of a circuit breaker, the market equilibrium is defined by an $\mathcal{F}_t$-stopping time τ, trading strategies $\{\theta_t^i, \phi_t^i\}$, $i = A, B$, and a continuous stock price process S_t, all defined on $[0, \tau \wedge T]$, such that:*

1. *Taking stock price process S_t as given, the two agents' trading strategies maximize their expected utilities under their respective beliefs and budget constraints.*
2. *For all $t \in [0, T]$, both the stock and bond markets clear:*
$$\theta_t^A + \theta_t^B = 1, \quad \phi_t^A + \phi_t^B = 0. \tag{9}$$
3. *The stopping time τ is consistent with the circuit breaker rule in (8).*

Discussion

The model described above takes a parsimonious form, mainly for clarity and tractability. Here, we elaborate a bit more on the model's specifications.

In the model, the heterogeneity between agents, which gives rise to their trading needs, is introduced in the form different beliefs. However, as discussed earlier, this specification can have other interpretations such as different utility functions. In this case, agent B's utility function becomes state-dependent and takes the form in (7). If η_t is positively (negatively) related to dividend shocks, which is the case in (5) with a positive (negative) ν, agent B's utility increases (decreases) with aggregate dividend/consumption, holding constant his wealth/consumption. The former case (positive relationship between η and D) is reminiscent of behavior biases like "representativeness", and the latter case (negative relationship between

η and D) is reminiscent of behavior like "catching up with the Joneses" (see, e.g., Abel, 1990). We adopt the interpretation of different beliefs in most of the paper, for convenience in exposition, but will return to both interpretations in Section 5 in the welfare analysis.

One can also introduce heterogeneity in other forms to generate trading such as different endowment shocks (see, e.g., Wang, 1994). Our formulation using different beliefs/preferences is not only simple to specify and interpret, but also helpful in our welfare analysis.

In the model, we consider the possibility of only one circuit breaker, and once triggered, the market will be closed until T, the end of the trading session.[13] In reality, there can be multiple circuit breakers during a trading session, with different triggers. For example, currently in the U.S., a market-wide trading halt can be triggered at three levels, namely, 7% (Level 1), 13% (Level 2), and 20% (Level 3) in price drops from prior day's closing price of the S&P 500 Index, respectively. In addition, for Level 1 and 2 circuit breakers, the market is closed only for 15 minutes and then reopens until the end of the day, unless Level 3 is triggered.[14] Although our analysis focuses on the single circuit-breaker case, for sake of clarity, our setup can be extended to allow multiple circuit breakers and market reopening.

3 Solution to Equilibrium

We now present the solution to the market equilibrium. We first consider the case without the circuit breaker and then the case with the circuit breaker. To distinguish the two cases, we use the symbol "^" to denote variables in the case without the circuit breaker.

3.1 Without Circuit Breaker: The Benchmark Case

In the absence of the circuit breaker, the market is dynamically complete. The equilibrium allocation in this case can be characterized as the solution to the following planner's problem:

$$\max_{\widehat{W}_T^A, \widehat{W}_T^B} \mathbb{E}_0 \left[\lambda \ln \left(\widehat{W}_T^A \right) + (1-\lambda)\, \eta_T \ln \left(\widehat{W}_T^B \right) \right] \tag{10}$$
$$\text{s.t.} \quad \widehat{W}_T^A + \widehat{W}_T^B = D_T.$$

[13]The fact that the price of the stock reverts back to the fundamental value D_T at T resembles the rationale of circuit breaker to "restore order" in the market. In this sense, instead of viewing T as the end of economy, it can be viewed as the reopening of the market.

[14]More details on the history of market-wide circuit breakers in the U.S. and their current form are contained in Online Appendix OA.1.

From the agents' first-order conditions and budget constraints, $\lambda = \omega$. The following proposition summarizes the market equilibrium including the stock price and individual agents' portfolio holdings.

Proposition 1. *Without the circuit breaker, the equilibrium stock price is:*

$$\widehat{S}_t = \frac{\omega + (1-\omega)\,\eta_t}{\omega + (1-\omega)\,\eta_t\, e^{a(t,T)+b(t,T)\delta_t}} e^{\left(\mu-\sigma^2\right)(T-t)} D_t, \quad t \in [0,T], \tag{11}$$

where

$$\begin{aligned} a(t,T) &= \left[\frac{\kappa\bar{\delta}-\sigma\nu}{\nu/\sigma-\kappa} + \frac{1}{2}\frac{\nu^2}{(\nu/\sigma-\kappa)^2}\right](T-t) - \frac{1}{4}\frac{\nu^2}{(\nu/\sigma-\kappa)^3}\left[1-e^{2(\nu/\sigma-\kappa)(T-t)}\right] \\ &\quad + \left[\frac{\kappa\bar{\delta}-\sigma\nu}{(\nu/\sigma-\kappa)^2} + \frac{\nu^2}{(\nu/\sigma-\kappa)^3}\right]\left[1-e^{(\nu/\sigma-\kappa)(T-t)}\right], \qquad (12a) \\ b(t,T) &= \frac{1}{\nu/\sigma-\kappa}\left[1-e^{(\nu/\sigma-\kappa)(T-t)}\right]. \qquad (12b) \end{aligned}$$

The two agents' shares of total wealth at time t are:

$$\widehat{\omega}_t^A = \frac{\omega}{\omega + (1-\omega)\eta_t}, \quad \widehat{\omega}_t^B = 1 - \widehat{\omega}_t^A, \tag{13}$$

and their stock and bond holdings are:

$$\widehat{\theta}_t^A = \widehat{\omega}_t^A\left(1 - \widehat{\omega}_t^B \frac{\delta_t}{\sigma\sigma_{\widehat{S},t}}\right), \quad \widehat{\theta}_t^B = 1 - \widehat{\theta}_t^A; \quad \widehat{\phi}_t^A = \widehat{\omega}_t^A\,\widehat{\omega}_t^B \frac{\delta_t}{\sigma\sigma_{\widehat{S},t}}\widehat{S}_t, \quad \widehat{\phi}_t^B = -\widehat{\phi}_t^A. \tag{14}$$

The conditional volatility of $\widehat{S}_t$, $\widehat{\sigma}_{S,t}$, can be computed in closed-form from Equation (11), which is given in the appendix.

As (14) shows, there are several forces affecting the agents' portfolio positions. First, all else equal, agent A owns more shares of the stock when B has more pessimistic beliefs (smaller $\delta_t < 0$). This effect becomes weaker when the volatility of stock return $\widehat{\sigma}_{S,t}$ is high. Second, the wealth distribution, as given in (13), also affects the agents' portfolio holdings, as the richer agent tends to hold more shares of the stock.

We can gain more intuition on the equilibrium stock price given in (11). Let $\widehat{S}_t^A$ and $\widehat{S}_t^B$ be the stock prices in the two single-agent economies with only agent A and B, respectively, as the representative agent, which are given by:

$$\widehat{S}_t^A = \frac{1}{\mathbb{E}_t[D_T^{-1}]} = e^{\left(\mu-\sigma^2\right)(T-t)} D_t, \quad \widehat{S}_t^B = \frac{1}{\mathbb{E}_t^B[D_T^{-1}]} = e^{\left(\mu-\sigma^2\right)(T-t)-a(t,T)-b(t,T)\delta_t} D_t. \tag{15}$$

We can then rewrite the equilibrium stock price as follows:

$$\widehat{S}_t = \left(\frac{\omega}{\omega + (1-\omega)\,\eta_t} \mathbb{E}_t[D_T^{-1}] + \frac{(1-\omega)\,\eta_t}{\omega + (1-\omega)\,\eta_t} \mathbb{E}_t^B[D_T^{-1}] \right)^{-1} = \left(\widehat{\omega}_t^A \frac{1}{\widehat{S}_t^A} + \widehat{\omega}_t^B \frac{1}{\widehat{S}_t^B} \right)^{-1}. \quad (16)$$

That is, the stock price with both agents is simply a harmonic average of the stock prices in the two single-agent economies, with the weights, $\widehat{\omega}_t^A$ and $\widehat{\omega}_t^B$, being their shares of total wealth. Therefore, controlling for wealth distribution, the equilibrium stock price is higher (lower) when agent B's belief is more (less) optimistic, i.e., when δ_t is larger (smaller).

We consider two special cases to gain more intuition about the model. One special case is when the disagreement between the two agents is the zero, i.e., $\delta_t = 0$ for all $t \in [0, T]$. The stock price then becomes:

$$\widehat{S}_t = \widehat{S}_t^A = e^{\left(\mu - \sigma^2\right)(T-t)} D_t. \quad (17)$$

This is a version of the Gordon growth formula, with σ^2 being the risk premium for the stock. The instantaneous volatility of stock returns becomes the same as the volatility of dividend growth, $\widehat{\sigma}_{S,t} = \sigma$. The shares of the stock held by the two agents will remain constant and be equal to their initial endowments, $\widehat{\theta}_t^A = \omega$ and $\widehat{\theta}_t^B = 1 - \omega$.

Another special case is when the disagreement is constant over time ($\delta_t = \delta$ for all t). The results for this case are obtained by setting $\nu = 0$ and $\delta_0 = \bar{\delta} = \delta$ in Proposition 1. (11) then simplifies to:

$$\widehat{S}_t = \frac{\omega + (1-\omega)\,\eta_t}{\omega + (1-\omega)\,\eta_t\, e^{-\delta(T-t)}} e^{\left(\mu - \sigma^2\right)(T-t)} D_t. \quad (18)$$

As expected, $\widehat{S}_t$ increases with δ, which reflects agent B's optimism on dividend growth. And its volatility now becomes stochastic.

3.2 With Circuit Breaker

We now solve for the market equilibrium in the presence of the circuit breaker. In this case, the market is no longer complete over the entire time span $[0, T]$, especially if the circuit breaker is triggered. However, a crucial feature of the model is that the market remains dynamically complete until the circuit breaker is triggered. At the trigger time, the market enters a closing equilibrium in which market participants trade one last time before the market closes and then hold on to their positions until T. We first characterize the equilibrium at and before market closure under an arbitrary trigger rule for the circuit breaker. Using these characterizations, we then construct a market equilibrium under the circuit breaker trigger rule we are interested in, which is (8), hence proving the existence of market equilibrium. Finally, we establish sufficient conditions for the uniqueness of market equilibrium.

Equilibrium characterization

Consider an arbitrary trigger rule for the circuit breaker, which is not necessarily the threshold hitting rule in (8). Suppose that the circuit breaker is triggered before the end of the trading session, i.e., $\tau < T$, here τ is the trigger time which could depend on the full history of the economy. We start by characterizing the closing equilibrium at τ. Since the two agents behave competitively, they take the stock price S_τ as given and choose their stock and bond holdings to maximize their expected utility over terminal wealth, subject to the budget constraint:

$$\begin{aligned} V^i(W^i_\tau, \tau) \ &= \max_{\theta^i_\tau, \phi^i_\tau} \mathbb{E}^i_\tau \left[\ln(\theta^i_\tau D_T + \phi^i_\tau) \right], \quad i = A, B \\ &\text{s.t.} \quad \theta^i_\tau S_\tau + \phi^i_\tau = W^i_\tau. \end{aligned} \tag{19}$$

Market clearing at time τ requires:

$$\theta^A_\tau + \theta^B_\tau = 1, \quad \phi^A_\tau + \phi^B_\tau = 0, \tag{20}$$

which determines the closing price S_τ as a function of W^i_τ, $i = A, B$, and the other state variables at τ.

For any $\tau < T$, the Inada condition implies that terminal wealth for both agents needs to remain non-negative, which implies that $\theta^i_\tau \geqslant 0$ and $\phi^i_\tau \geqslant 0$. That is, neither agent will take short or levered positions in the stock at τ. This is a direct result of the inability to rebalance one's portfolio after market closure, which is an extreme version of illiquidity.

Next, before the trigger time, both agents trade continuously to maximize their expected indirect utilities at $\tau \wedge T$, i.e., $V^i(W^i_{\tau\wedge T}, \tau \wedge T)$. Given that the market is dynamically complete before $\tau \wedge T$, the equilibrium is equivalent to the following planner problem:

$$\max_{W^A_{\tau\wedge T}, W^B_{\tau\wedge T}} \mathbb{E}_0 \left[\lambda\, V^A(W^A_{\tau\wedge T}, \tau \wedge T) + (1-\lambda)\, \eta_{\tau\wedge T} V^B(W^B_{\tau\wedge T}, \tau \wedge T) \right], \tag{21}$$

subject to the resource constraint:

$$W^A_{\tau\wedge T} + W^B_{\tau\wedge T} = S_{\tau\wedge T} = \begin{cases} D_T, & \text{if} \quad \tau \geqslant T \\ S_\tau, & \text{if} \quad \tau < T. \end{cases} \tag{22}$$

The resulting wealth allocation $W^i_{\tau\wedge T}$ is used in the closing equilibrium (19). Meanwhile, it also helps to specify the market clearing price before τ like the complete market counterpart.

The following result characterizes the equilibrium price before and at the circuit breaker trigger time.

Proposition 2. *Let*

$$\underline{\delta}(t) = -\frac{a(t,T)}{b(t,T)}, \tag{23}$$

where $a(t,T)$, $b(t,T)$ are given by (12) in Proposition 1. Suppose that the market closes at time $\tau < T$. Then,

(i) At τ both agents will hold all of their wealth in the stock, $\theta^i_\tau = \frac{W^i_\tau}{S_\tau}$, and nothing in the bond, $\phi^i_\tau = 0$, $i = A, B$. The market clearing price is:

$$S_\tau = \min\left\{\widehat{S}^A_\tau, \widehat{S}^B_\tau\right\} = \begin{cases} e^{\left(\mu-\sigma^2\right)(T-\tau)} D_\tau, & \text{if} \quad \delta_\tau > \underline{\delta}(\tau) \\ e^{\left(\mu-\sigma^2\right)(T-\tau)-a(\tau,T)-b(\tau,T)\delta_\tau} D_\tau, & \text{if} \quad \delta_\tau \leqslant \underline{\delta}(\tau) \end{cases} \tag{24}$$

where $\widehat{S}^i_\tau$ denotes the stock price in a single-agent economy populated by agent i, as given in (15), and

(ii) For $t \leqslant \tau \wedge T$, the market clearing price is:

$$S_t = \left(\omega^A_t \mathbb{E}_t\left[S^{-1}_{\tau\wedge T}\right] + \omega^B_t \mathbb{E}^B_t\left[S^{-1}_{\tau\wedge T}\right]\right)^{-1}, \tag{25}$$

where ω^i_t is the share of total wealth owned by agent i before market closure, and is identical to $\widehat{\omega}^i_t$ in (13). Moreover, the market clearing price S is continuous on $[0, \tau \wedge T]$.

Because neither agent will take levered or short positions during market closure, with zero net bond supply, there will not be any lending or borrowing over this period. Therefore, the two agents will invest all their respective wealth in the stock at the market closure. This implies that at τ, the relatively optimistic investor faces the leverage constraint, and the relatively pessimistic investor becomes the marginal investor. Consequently, the stock price at τ, S_τ, cannot be lower than the pessimistic investor's evaluation. Otherwise, continuity requires that the stock price would also be lower than the pessimistic investor's evaluation right before τ. In this case, both investors would want to take on leverage to invest in the under-valued stock, which contradicts with the unconstrained equilibrium before τ. Thus, the market clearing price at the trigger time must be the pessimistic investor's evaluation.

The result that S_τ only depends on the belief of the relatively pessimistic agent is qualitatively different from the complete market case, where the stock price is a wealth-weighted average of the prices under the two agents' beliefs. This is a crucial result: the lower stock valuation upon market closure affects both the stock price level and dynamics before market closure, which we analyze in Section 4. Notice that having the lower expectation of the growth rate at the current instant is not sufficient to make the agent marginal. One also needs to take into account the agents' future beliefs and the risk premium associated with future fluctuations in the beliefs, which are summarized by $\underline{\delta}(t)$.

Before the circuit breaker trigger time, the market clearing price in (25) is reminiscent of its complete market counterpart (16). Unlike in the case of complete markets, the expectations in (25) are no longer the inverse of the stock prices from the respective representative agent economies.

Equilibrium existence and uniqueness

Proposition 2 characterizes the market equilibrium under an arbitrary triggering rule for the circuit breaker. It thus allows us to obtain the equilibrium under the specific triggering rule given in (8). We achieve this in two steps. First, we consider the equilibrium under a triggering rule based on a given constant lower bound for the equilibrium stock price, denoted by $\underline{S}$. From Proposition 2, this fully specifies the stock price on $[0,\, \tau \wedge T]$, including at $t = 0$. Next, we require the triggering price to satisfy $\underline{S} = (1-\alpha)S_0$ and show that an equilibrium exists. This will then give an equilibrium under circuit breaker trigger rule (8).

Consider a triggering rule defined by a constant lower bound for the equilibrium stock price $\underline{S}$. That is,

$$\tau = \inf\{\, t \geqslant 0 : S_t \,\leqslant\, \underline{S}\,\}. \tag{26}$$

In this case, (24) implies that

$$D_\tau = \underline{D}(\tau, \delta_\tau; \underline{S}) \equiv \begin{cases} \underline{S}e^{-(\mu-\sigma^2)(T-\tau)}, & \text{if} \quad \delta_\tau > \underline{\delta}(\tau) \\ \underline{S}e^{-(\mu-\sigma^2)(T-\tau)+a(\tau,T)+b(\tau,T)\delta_\tau}, & \text{if} \quad \delta_\tau \leqslant \underline{\delta}(\tau). \end{cases} \tag{27}$$

The above property of the dividend process at τ for triggering rule (26) allows us to construct an equilibrium under triggering rule (8) and hence establish its existence. To this end, we separate our construction into three steps:

(i) For a given $\underline{S}$, define a stopping time:

$$\tau' = \inf\{t \geqslant 0 : D_t \leqslant \underline{D}(t, \delta_t; \underline{S})\}, \tag{28}$$

where $\underline{D}(t, \delta_t; \underline{S})$ is given in (27).

(ii) Define

$$S'_t = \left(\omega^A_t \, \mathbb{E}_t\left[(\widehat{S}^{min}_{\tau'\wedge T})^{-1}\right] + \omega^B_t \, \mathbb{E}^B_t\left[(\widehat{S}^{min}_{\tau'\wedge T})^{-1}\right]\right)^{-1}, \quad t \leqslant \tau', \tag{29}$$

where $\widehat{S}^{min}_t = \min\{\widehat{S}^A_t, \widehat{S}^B_t\}$.

(iii) Evaluating S'_t at $t = 0$, we have the map F: $F(\underline{S}/(1-\alpha)) = S'_0$. We then look for a fixed point for the map F: $F(S_0) = S_0$.

In the construction described above, the definition of τ' is motivated by (27) and is exogenously specified. It is given by a triggering rule specified by the dividend process with a parameter $\underline{S}$. We will prove in the next result that τ' is exactly the circuit breaker trigger time τ in (8) for the equilibrium constructed. The process S' in (29) is motivated by (25), and parametrized by $\underline{S}$ in the construction. Once we identify a fixed point S_0 for the map F, then S' becomes the market clearing price starting from the initial price S_0 and τ' is the circuit breaker trigger time in (8), hence a corresponding equilibrium is obtained. The following result formally presents the existence of equilibrium.

Proposition 3. *Under circuit breaker rule (8), there exists a market equilibrium. In particular, τ' defined in (28) is the circuit breaker trigger time, i.e., $\tau' = \tau$ in (8).*

For the uniqueness of equilibrium, we consider cases of constant and stochastic δ separately. In the case of stochastic δ, given by (5), we are interested in the situation with $\bar{\delta} = 0$, that is, when there is no persistent disagreements between the two agents. The case with persistent disagreement is captured by the constant disagreement case. In the former case, we also set the initial disagreement to be zero for simplicity, i.e., $\delta_0 = 0$. The following result identifies the marginal agent at the circuit breaker trigger time in the case of stochastic δ.

Lemma 1. *Suppose that $\bar{\delta} = \delta_0 = 0$ in (5). Then agent B is the marginal agent when the circuit breaker is triggered at τ for $0 \leqslant \tau < T$.*

It is intuitive that agent B would be the marginal agent at τ. Since the circuit breaker will be triggered when the stock price has fallen sufficiently from the initial value, it will be on paths with more negative dividend shocks. Thus, starting with $\delta_0 = 0$, these shocks will make δ_t fall below zero at τ, making agent B more pessimistic relative to agent A. Lemma 1 states that this is indeed the case for general parameters.

With the results above, we have the following sufficient conditions for the uniqueness of equilibrium.

Proposition 4. *The market equilibrium is unique when:*

(i) δ *is constant, or*

(ii) $\bar{\delta} = \delta_0 = 0$ *and* $\omega < 1 - \alpha$.

In the case of stochastic δ, $\omega < 1 - \alpha$ implies that the weight of agent B's initial wealth is at least α. This is a technical condition and is not necessary. Our extensive numerical analysis shows that the equilibrium remains unique even when this condition is violated.

4 Impact of Circuit Breaker on Market Behavior

We now examine how the circuit breaker changes market dynamics, including the behavior of stock price and agents' stock holdings. We first consider the case of constant disagreement. This case allows us to obtain analytical results about how the circuit breaker changes price dynamics and to illustrate the mechanism behind. We next consider the general case with time-varying disagreements, which allows us to gain more general and quantitative results on the impact of the circuit breaker. Throughout this section, we consider $\alpha \in (0, 1]$ to allow more trading and price evolution since the case of $\alpha = 0$ corresponds to trading only at $t = 0$.

4.1 Constant Disagreement

For constant disagreement, we have $\delta_t \equiv \delta$. To fix ideas, we work with the case of $\delta < 0$, in which B is the more pessimistic agent. The results remain the same for the case of $\delta > 0$, in which A becomes the more pessimistic agent, by relabeling.

Price level, volatility, and magnet effect

By comparing the stock prices and their volatility in the equilibria with and without the circuit breaker, respectively, we can analyze the impact of the circuit breaker.

Price level We first consider the equilibrium price level. Because agent B is relatively more pessimistic, Proposition 2 implies that the stock price at τ, when $\tau < T$ and the circuit breaker is triggered, equals agent B's evaluation. That is, $S_{\tau\wedge T} = \widehat{S}^B_{\tau\wedge T} = \left(\mathbb{E}^B_{\tau\wedge T}[D_T^{-1}]\right)^{-1}$. The stock price prior to τ in (25) can be transformed into the following equivalent form:

$$S_t = \left(\omega_t^A\,\mathbb{E}_t\big[\mathbb{E}^B_{\tau\wedge T}[D_T^{-1}]\big] + (1-\omega_t^A)\,\mathbb{E}_t^B[D_T^{-1}]\right)^{-1}, \tag{30}$$

where we have used the property of conditional expectations: $\mathbb{E}_t[S^{-1}_{\tau\wedge T}] = \mathbb{E}_t\big[\mathbb{E}^B_{\tau\wedge T}[D_T^{-1}]\big]$ and $\mathbb{E}^B_t[S^{-1}_{\tau\wedge T}] = \mathbb{E}^B_t\big[\mathbb{E}^B_{\tau\wedge T}[D_T^{-1}]\big] = \mathbb{E}^B_t[D_T^{-1}]$. Here, the stock price is still the harmonic average of their respective prices in the two single-agent economies, except that the terminal value is no longer the dividend at T, as in the case without the circuit breaker, but the price at trigger time $\tau\wedge T$, $S_{\tau\wedge T}$, which is determined by the valuation of agent B, who is the marginal investor at that time. Equation (30) helps us obtain the following comparison between stock prices with or without circuit breaker.[15]

Proposition 5. *The stock price is always lower with the circuit breaker than without, i.e., for all $t \in [0,\, T]$ and and $D_t \geqslant \underline{D}_t = \underline{D}(t,\delta)$,*

$$S_t < \widehat{S}_t. \tag{31}$$

This proposition simply states that, given the same fundamental (D_t), the equilibrium stock price is always lower with the circuit breaker than without. The reason is that, with or without the circuit breaker, the inverse stock price is in general a weighted average between optimistic and pessimistic evaluations of the two agents (cf. (16)). However, at the circuit breaker trigger time, the stock price is determined by the pessimistic evaluation, which then leads to a lower price than that without the circuit breaker, even before the trigger time.

Price dynamics and volatility Next, we examine the impact of the circuit breaker on price dynamics and volatility, especially in its the neighborhood. The following proposition summarizes the main results.

Proposition 6. *Comparing stock price with and without the circuit breaker, we have:*

(i) The price without the circuit breaker, $\widehat{S}_t$, and the price with circuit breaker, S_t, satisfy:

$$\lim_{D_t\downarrow\underline{D}_t}\widehat{S}_t > \lim_{D_t\downarrow\underline{D}_t} S_t = (1-\alpha)\,S_0, \quad \textit{for all} \quad t<T.$$

[15] Our proof in the appendix further shows that the statement of Proposition 5 also hold for stochastic δ with $\bar{\delta} = \delta_0 = 0$.

(ii) With the circuit breaker and $\delta \leqslant -\sigma^2$, the stock price volatility σ_S satisfies:

$$\lim_{D_t \downarrow \underline{D}_t} \sigma_{S,t} > \lim_{D_t \downarrow \underline{D}_t} \sigma_{\widehat{S},t} > \sigma \quad \textit{for all} \quad t < T.$$

(iii) In addition, when ω is close to 1 and $\delta < -\sigma^2$, there exists a neighborhood around $\underline{D}_t$ such that in this neighborhood, as $D_t \downarrow \underline{D}_t$, $\widehat{S}_t/S_t$ increases and converges to a constant greater than 1, and $\sigma_t^S - \sigma_t^{\widehat{S}}$ increases and converges to a constant greater than 0.

The first result simply states that when the fundamental approaches the threshold level $\underline{D}_t = \underline{D}(t, \delta)$, the probability of triggering the circuit breaker approaches 1 so that the stock price with the circuit breaker converges to the trigger level $(1 - \alpha)S_0$, while the price without the circuit breaker approaches a level strictly higher than $(1 - \alpha)S_0$ as Proposition 5 states.

The second result shows that when the market approaches the circuit breaker threshold, the price volatility is higher than the price volatility without the circuit breaker, which is higher than the fundamental volatility. The intuition is as follows. Following negative fundamental shocks, agent B, being the more pessimistic agent, is exerting more influence on the stock price. This is true even without the circuit breaker because the more pessimistic agent will gain wealth share following negative fundamental shocks. However, since the presence of a circuit breaker reduces the willingness of the optimistic agent A to hold stocks, agent B becomes even more important (more so than implied by his wealth share) in driving the prices and eventually becomes the sole marginal investor upon market closure. As a result, any negative shock to the stock price gets amplified endogenously as it adds to agent B's pricing influence. This is the reason behind the higher price volatility as the market approaches the circuit breaker threshold.

We further illustrate the mechanism driving the price dynamics around the circuit breaker by considering the case when ω is close to 1 (but not at 1). In this case, agent A dominates the market in terms of wealth share. Nonetheless, B, the pessimistic agent, even though infinitesimal in wealth share, is still the marginal investor at the circuit breaker triggering point. Without the circuit breaker, the equilibrium stock price is close to:

$$\widehat{S}_t = \Big(\mathbb{E}_t[D_T^{-1}]\Big)^{-1},$$

which is simply the valuation of the agent A, who dominates the market in wealth share, based on her view of the fundamental, which is D_T. With the circuit breaker, however, the equilibrium stock price becomes:

$$S_t = \left(\mathbb{E}_t\big[\mathbb{E}^B_{\tau\wedge T}[D_T^{-1}]\big]\right)^{-1}.$$

Now, agent A evaluates the stock based on its future valuation at the circuit breaker trigger time, which is determined by agent B. Part (iii) in the proposition shows that in the neighborhood of $\underline{D}_t$, as the circuit breaker approaches, the ratio $\widehat{S}_t/S_t$ increases to a limit greater than 1 and the gap between σ^S and $\sigma^{\widehat{S}}$ widens to a limit greater than 0.

Magnet effect. The "magnet effect" is a popular term among practitioners referring to changes in price dynamics that accelerate the process of reaching the trading halt trigger as the price moves closer to it. While there is no established definition of this effect, we provide a more formal notion here.

Let $p_h(S_t) \equiv \mathbb{P}(\tau \leqslant t + h|\mathcal{F}_t)$ denote the conditional probability that the stock price, currently at S_t, will hit the circuit breaker threshold $(1-\alpha)S_0$ within a given time interval h, which we refer to as conditional hitting probability, and $\widehat{p}_h(S_t)$ denote the conditional hitting probability in the absence of the circuit breaker. The difference between these two hitting probabilities, given by $\Delta p_h(S_t) \equiv p_h(S_t) - \widehat{p}_h(S_t)$, then gives the impact of the circuit breaker on the conditional hitting probability. We define the "magnet effect" as the situation when $\Delta p_h(S_t)$ becomes increasingly positive when S_t is getting closer to the threshold.

Some studies of circuit breakers associate the magnet effect with the unconditional probability of reaching the threshold. We find this association unsatisfactory. In general, the introduction of a circuit breaker changes the overall equilibrium, including the price distribution, conditional or unconditional. If, for example, the unconditional volatility of price distribution increases, the probability of it reaching any threshold also increases. The magnet effect, on the other hand, is really about the dynamic behavior of price, especially when it approaches the circuit breaker. The conditional hitting probability defined above is one way to capture how the circuit breaker changes the price dynamics near its vicinity. In other words, the magnet effect, by its very nature, is a dynamic effect, not a static one.

By our version of the magnet effect, the very presence of a circuit breaker raises the probability of the stock price hitting the threshold as the stock price moves closer to the threshold. The next proposition compares $p_h(S_t)$ and $\hat{p}_h(S_t)$ in our model. To generate nontrivial comparison, we reduce the horizon h as S_t approaches $\underline{S}$.[16]

Proposition 7. *Suppose $S_t > \underline{S}$ and $h = (S_t - \underline{S})^2$. There exists a neighborhood around $\underline{S}$ such that when S_t is in this neighborhood:*

(i) $p_h(S_t)$ increases as $S_t \downarrow \underline{S}$.

(ii) For $\delta < -\sigma^2$, $p_h(S_t) > \hat{p}_h(S_t)$. Moreover, if ω is also close to 1, then as $S_t \downarrow \underline{S}$, $p_h(S_t) - \hat{p}_h(S_t)$ increases and converges to a constant greater than 0.

Here, the magnet effect is mainly caused by the significant increase in conditional return volatility in the presence of a circuit breaker. As the stock price becomes increasingly more volatile when approaching the circuit breaker, the likelihood of hitting the circuit breaker increases as well.

Quantitative analysis of market behavior

In addition to the analytical results presented above about the stock price level and its dynamics near the circuit breaker, we can also explore quantitatively the impact of the circuit breaker on price and trading behavior over the whole state space. Given that the market behavior in this case is qualitatively similar to that in the general case with time-varying disagreement, we will discuss it next in the context of the general case. The specific results about the constant disagreement case are provided in the Online Appendix (OA.3) for comparison.

4.2 Time-varying Disagreement

We now turn to the general case with time-varying disagreement, which better captures the possibility of large intraday market swings.[17] To fix ideas, we focus on the case where the difference in beliefs δ_t follows a random walk. We do so by setting $\kappa = 0$, $\nu = \sigma$, and

[16] If h is fixed, S_t approaches $\underline{S}$, $\lim_{S_t \downarrow \underline{S}} p_h(S_t) = \lim_{S_t \downarrow \underline{S}} \hat{p}_h(S_t) = 1$, which gives the trivial result: with or without the circuit breaker, the price will hit the threshold surely.

[17] An important rationale for policy makers to introduce a circuit breaker is to curb disorderly trading, in particular "panic selling," under extreme market conditions. This is what motivates our modeling of the subjective belief deviation δ_t as time-varying, which is initially zero and only becomes significant when the market moves substantially within a day. Under such an interpretation, a paternalistic planner/regulator may view the trading motive of agent B as speculation driven by irrational beliefs, and see potential benefits from introducing a circuit breaker to curb such trading activities (see Section 5 for more discussion on this). Moreover, while it might be difficult for the planner to discern relatively mild deviations from objective probabilities under normal market conditions (Blume et al., 2018), it is more reasonable to assume that the planner can identify irrational panics.

$\delta_0 = \bar{\delta} = 0$. Thus, there is neither initial nor long-term bias in agent B's belief. The more general case with $\kappa > 0$ yields qualitatively similar results.

Under this specification, agent B extrapolates his belief about future dividend growth from realized dividends. In particular, $\delta_t = \sigma Z_t$. As a result, he becomes overly optimistic following large positive dividend shocks and overly pessimistic following large negative dividend shocks. More importantly, the circuit breaker is approached after substantial drops in D_t (from initial value 1) or Z_t (from initial value 0), which corresponds to substantially negative δ_t. As a result, in this case, agent B is always the more pessimistic agent near the circuit breaker.[18]

For calibration, we normalize $T = 1$ to denote one trading day. We set the expected value of the dividend growth $\mu = 10\%/250 = 0.04\%$ (implying an annual dividend growth rate of 10%) and its (daily) volatility $\sigma = 3\%$. The downside circuit breaker threshold is set at $\alpha = 5\%$. For the initial wealth distribution, we assume agent A (with objective belief) owns 90% of total wealth ($\omega = 0.9$) at $t = 0$.

Given the two agents' beliefs, we now examine how the circuit breaker changes the market behavior. In particular, we focus on the following aspects: price-to-dividend ratio S_t/D_t, conditional return volatility $\sigma_{S,t}$, expected return under agent A's belief, which is also the objective probability, $\mu^A_{S,t}$, and agent A's stock holding θ^A_t. Agent B's stock holding is given by $\theta^B_t = 1 - \theta^A_t$. Since the impact of the circuit breaker changes during the day, we choose the time to be $t = 0.1$ to fix ideas. We will discuss time variation in the circuit breaker's impact in the Online Appendix (OA.4).

Figure 1 plots these four quantities, S_t/D_t, $\sigma_{S,t}$, $\mu^A_{S,t}$ and θ^A_t, as a function of the fundamental D_t at $t = 0.1$. Unlike the constant disagreement case, in general, dividend D_t and time of the day t are no longer sufficient to fully determine the state of the economy. Thus, we plot the average values of the variables conditional on t and D_t here.[19]

Price-dividend ratio. We first consider the circuit breaker's impact on the price-dividend ratio, shown in panel A of Figure 1. Since agent A's belief about the dividend growth rate is constant over time, the price-dividend ratio under her belief is constant over different values of D_t, shown by the horizontal gray dash line ($\widehat{S}^A_t/D_t = 1$). Due to the variation in δ_t, which

[18] Our earlier result Lemma 1 shows that this statement still holds in the general case with $\kappa > 0$.

[19] In the case considered here, δ_t follows a random walk. The additional state variable besides t and D_t is the Radon-Nikodym derivative η_t given in (6). For simplicity, we plot the variables of interest conditional on t and D_t, averaging over the conditional distribution for η_t.

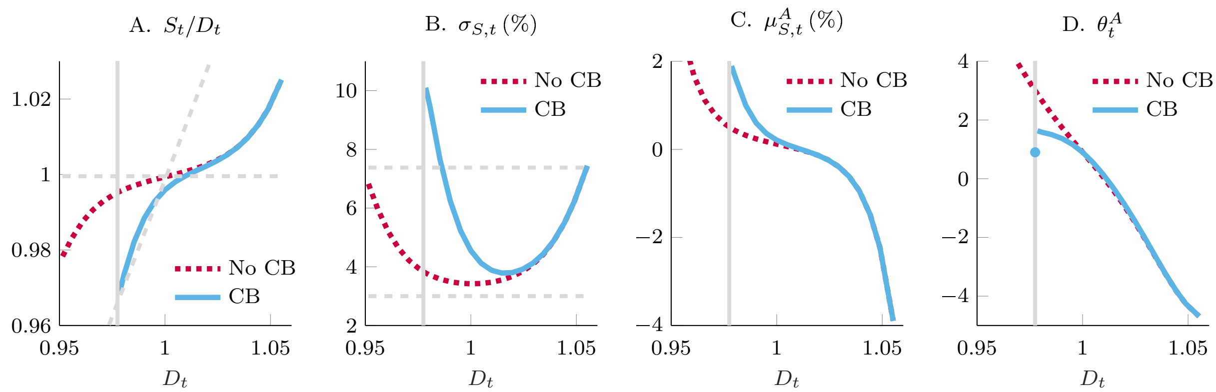

Figure 1: Price-dividend ratio, return volatility, instantaneous expected return and agent A's portfolio at $t = 0.1$ in the case of time-varying disagreement. In each panel, the gray vertical lines denotes the circuit breaker threshold $\underline{D}_t$, the gray dash lines are for the two cases with only one of the agents present in the market. The variables of interest are conditional on D_t. The parameters are $T = 1$, $\mu = 0.04\%$, $\sigma = \nu = 3\%$, $\kappa = 0$, $\delta_0 = \bar{\delta} = 0$, $\omega = 0.9$, and $\alpha = 5\%$.

is perfectly correlated with D_t, the price-dividend ratio under agent B's belief now increases with D_t, shown by the upward-sloping gray dash line ($\widehat{S}^B_t/D_t$).

As discussed in Section 3.1, when both agents are present, the price ratio without the circuit breaker is a wealth-weighted average of the prices from the two representative-agent economies populated by agent A and B, respectively. This is shown by the red dotted line in the figure, which indeed lies between the two gray dash lines. As D_t decreases, agent A becomes relatively more optimistic, and holds more stock shares. Consequently, the stock price reflects more of her belief. As D_t increases, the opposite is true. Agent B becomes relatively more optimistic, holds more stock shares, and have a larger impact on the price.

In the presence of the circuit breaker, for any given level of dividend D_t above the circuit breaker threshold, the price-dividend ratio, shown by the solid blue line, is always lower than the value without the circuit breaker. Moreover, the gap between the two price-dividend ratios becomes negligible when D_t is sufficiently large and the market is far away from the circuit breaker. However, as D_t approaches the threshold $\underline{D}_t$ and the market moves closer to the circuit breaker, the difference becomes more pronounced.[20] When the circuit breaker is hit, the difference exceeds 2%.

We also note that in the presence of the circuit breaker, the stock price declines more rapidly as the dividend approaches the trigger threshold. The reason behind this behavior can

[20]In general cases, the threshold $\underline{D}(t, \delta_t)$ depends on both t and δ_t. Since our calibration of the δ process implies a one-to-one mapping between δ_t and D_t, the threshold becomes unique for any t.

be traced to how the stock price is determined upon market closure. As explained in Section 3.2, at the instant when the circuit breaker is triggered, neither agent will be willing to take on levered positions in the stock due to the inability to rebalance the portfolio thereafter. With bonds in zero net supply, the leverage constraint always binds for the relatively optimistic agent (agent A), and the market clearing stock price has to be such that agent B is willing to hold all of his wealth in the stock, *regardless of his share of total wealth*. Indeed, we see the price-dividend ratio with circuit breaker converging to $\widehat{S}_t^B/D_t$ when D_t approaches $\underline{D}_t$, instead of the wealth-weighted average of $\widehat{S}_t^A/D_t$ and $\widehat{S}_t^B/D_t$, given by the red dotted line. The lower stock price at the circuit breaker threshold also drives the stock price lower before market closure, with the effect becoming stronger as D_t moves closer to the threshold $\underline{D}_t$. This gives rise to the accelerated decline in stock price as D_t drops, which also implies higher price sensitivity to dividend shocks.[21]

Conditional return volatility. The higher sensitivity of the price-dividend ratio to dividend shocks due to the circuit breaker manifests itself in elevated conditional return volatility, as shown in panel B of Figure 1. Quantitatively, the impact of the circuit breaker on the stock return volatility can be quite sizable. Without the circuit breaker, the conditional volatility of returns (red dotted line) peaks at about 7.3%. With the circuit breaker, the conditional volatility (blue solid line) becomes substantially higher as D_t approaches $\underline{D}_t$. In particular, the conditional volatility reaches 10% at the circuit breaker threshold.

Conditional expected return. Panel C of Figure 1 plots the conditional expected stock return under the objective probability measure, which is the same as agent A's belief. Even when there is no circuit breaker, the conditional expected return rises as dividend falls. This is because the irrational agent B is both gaining wealth share and becoming more pessimistic as D_t falls, driving prices lower. Given that this price drop is not related to changes in expected *future* payoffs, the expected return rises. The presence of the circuit breaker accelerates the increase in the conditional expected return as D_t approaches the threshold $\underline{D}_t$.

[21]The circuit breaker does rule out extreme low values for the price-dividend ratio during the trading session, which could occur at extremely low dividend values had trading continued. This could be one of the benefits of circuit breakers in the presence of market frictions. For example, when there are intra-day mark-to-market requirements for some market participants, a narrower range for the price-dividend ratio can help to reduce the chances of inefficient liquidations that could further destabilize the market. Formally modeling such frictions will be an interesting direction for future research but beyond the scope of this paper.

Agents' stock holdings. We can also analyze the impact of the circuit breaker on the equilibrium stock price by connecting it to how the circuit breaker influences the equilibrium stock holdings of the two agents. Let us again start with the case without the circuit breaker, shown by the red dotted line in panel D of Figure 1. The stock holding of agent A, $\widehat{\theta}_t^A$, continues to rise as D_t falls to $\underline{D}_t$ and beyond. This is the result of two effects: (i) with lower D_t, the stock price is lower, implying higher expected return under agent A's belief, who is taking levered positions in the stock; and (ii) lower D_t also makes agent B, who is shorting the stock, wealthier and thus more capable of lending to agent A, who then takes on more levered positions.

With the circuit breaker, while the stock holding θ_t^A takes on similar values as $\widehat{\theta}_t^A$, its counterpart in the case without the circuit breaker, for large values of D_t, it becomes visibly lower than $\widehat{\theta}_t^A$ as D_t approaches the circuit breaker threshold, as shown by the blue solid blue line. This is because agent A becomes increasingly concerned with the rising return volatility at lower D_t, which offsets the effect of higher expected return. Finally, θ_t^A takes a discrete drop when $D_t = \underline{D}_t$. With the leverage constraint binding, agent A will hold all of his wealth in the stock, which means θ_t^A will be equal to his wealth share ω_t^A. The stock price in equilibrium has to fall enough such that agent A has no incentive to sell more of his stock holding. This is indeed the case as shown in panel A.

Magnet effect – Accelerating likelihood of hitting the circuit breaker. Given our definition of the magnet effect, in Figure 2, we plot the two conditional hitting probabilities, $p_h(S_t)$ and $\widehat{p}_h(S_t)$, respectively. To avoid trivial limits of these conditional hitting probabilities, as for Proposition 7, h is set to be proportional to $(S_t - \underline{S})^2$, so it shrinks as S_t approaches the circuit breaker.

Comparing $p_h(S_t)$ and $\widehat{p}_h$, shown by the solid blue line and the red dash line, respectively, we obtain the following results. First, the conditional hitting probability with the circuit breaker is always higher than that without the circuit breaker. This result comes largely from the fact that the circuit breaker tends to increases price volatility, which increases the hitting probability. Next, when S_t is sufficiently far from the circuit breaker threshold, the conditional hitting probabilities with and without the circuit breaker are both essentially zero. Third, the gap between the two quickly widens as the stock price moves closer to the threshold. The conditional hitting probability with the circuit breaker starts to increase

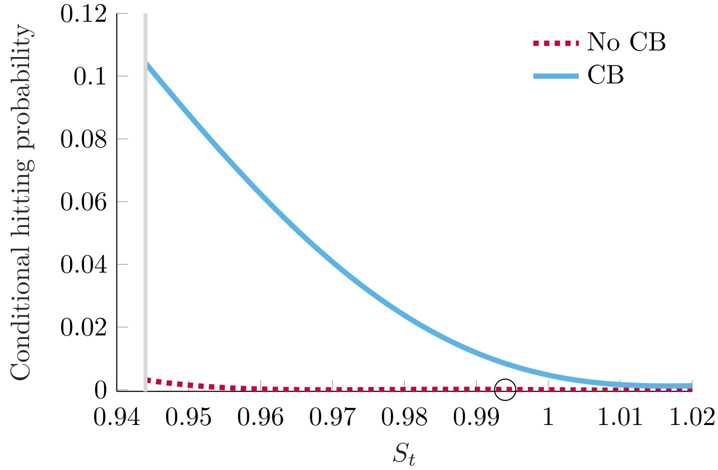

Figure 2: The "magnet effect". In the case of time-varying disagreement, we plot the conditional probabilities for the stock price to reach the circuit breaker limit within the next h minutes from $t = 0.1$. Here h is proportional to $(S_t - \underline{\mathrm{S}})^2$: h is roughly 0.5 minute when $S_t = 0.95$, 9.5 minutes when $S_t = 0.97$, and 62 minutes when $S_t = 1.01$. The black circle represents S_0. The gray vertical bar denotes the circuit breaker threshold $\underline{\mathrm{S}}$. The parameters are $\mu = 0.04\%$, $\sigma = \nu = 3\%$, $\kappa = 0$, $\delta_0 = \bar{\delta} = 0$, $\omega = 0.9$, and $\alpha = 5\%$.

substantially when S_t is still far away from the circuit breaker, at 1, while the conditional hitting probability with circuit breaker remains essentially zero and only picks up slightly when S_t is sufficiently close to the circuit breaker, around 0.95. This is the magnet effect as defined above. This effect arises from the fact that as the stock price is getting closer to the circuit breaker threshold, its volatility increases significantly, as shown in Proposition 6 and Figure 1, which then leads to higher hitting probability.[22]

The quantitative analysis above shows that under more general specifications of time-varying disagreement, the impact of circuit breaker on market behavior is consistent with the analytical results obtained for the constant disagreement case. In addition, the impact on price level, volatility and expected return can be substantial in magnitude.

4.3 Key model implications

As demonstrated in the theoretical analysis above, our model yields concrete predictions about how the circuit breaker affects the dynamics of the stock price, as captured by its conditional moments and trading behavior. To connect more directly with our empirical analysis later, we simulate the model with time-varying disagreement and examine the relationship between

[22]From Panel C of Figure 1, we also see that as the price moves closer to the circuit breaker, its drift $\mu^A_{S,t}$ also becomes more positive, which tends to pull the price away from the trigger threshold. However, as we show in the proof of Proposition 7 in Appendix, the volatility effect always dominates in the neighborhood of the threshold.

several *realized moments* of returns and the distance of the stock price from the circuit breaker threshold. Returns are measured over one-minute intervals and the distance to the circuit breaker is defined as $S_t/S_0 - (1-\alpha)$. The results are plotted in Figure 3, shown by the blue solid lines. For comparison, we have also shown the results without the circuit breaker in red dotted lines. We have the following predictions:

i. Realized return volatility increases dramatically as the distance to circuit breaker falls, and the rate of increase accelerates substantially as the distance to circuit breaker falls (until it is close to zero).[23] This differs substantially from the situation without the circuit breaker, in which the volatility increases very mildly.

ii. Realized return skewness becomes negative and further decreases as the distance to circuit breaker falls. This decline is only reversed when the distance to circuit breaker is sufficiently small, at which point the fact that prices are bounded from below starts to have a significant effect on skewness.

iii. Expected return rises as the distance to circuit breaker becomes very small.

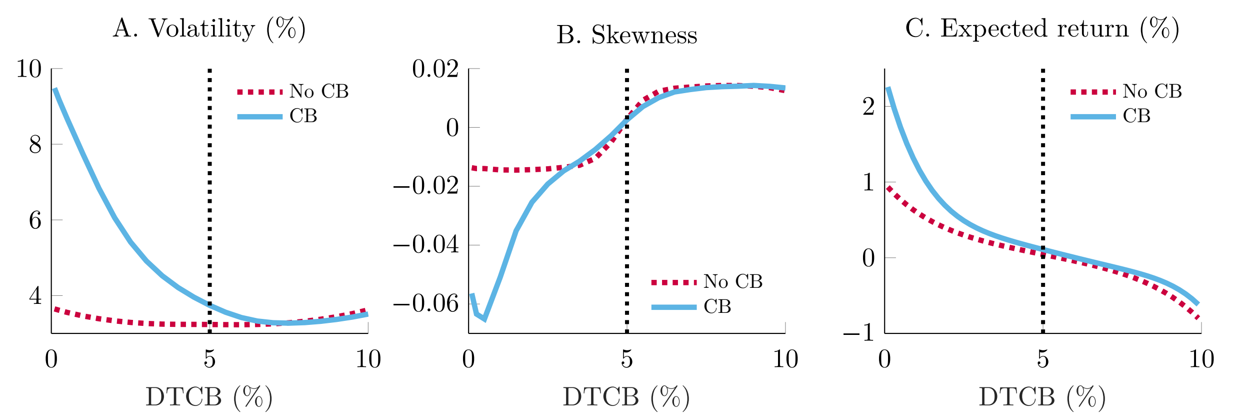

Figure 3: Model predicted 1-minute conditional volatility, conditional skewness, and conditional expected return as a function of the distance to the circuit breaker threshold (DTCB, defined as $S_t/S_0 - (1-\alpha)$) in the case of time-varying disagreement, shown in blue solid lines. Both volatility and return are normalized to one day. The parameters are $\mu = 0.04\%$, $\sigma = \nu = 3\%$, $\kappa = 0$, $\delta_0 = \bar{\delta} = 0$, $\omega = 0.9$, and $\alpha = 5\%$.

Comparing with the behavior of return moments without the circuit breaker, we see sharp differences. With no circuit breaker, the volatility exhibit much smaller variations overall, increasing only mildly when the price level drops; the skewness turns mildly negative as the

[23]Note that Panel A of Figure 3 is different from Panel B of Figure 1 due to a change in variable on the x-axis from D_t to DTCB.

price drops below its initial level) and then flattens out. These are in stark contrast to the fast and accelerating rise in volatility and negative skewness as well as its final reversal when the market moves closer to the circuit breaker when it is present.

Although our diffusion-based model does not yield a proper measure of trading volume, it does predict that the more optimistic investors will significantly reduce their stock positions at the instance when the circuit breaker is triggered. In the presence of transaction costs and execution delays, these investors will want to cut their stock positions earlier instead of waiting until the last moment. While we do not have account-level data to examine the directions of trades of different investors, we can examine the following prediction about total trading volume (from a generalized version of our model):

iv. Trading volume increases as the distance to circuit breaker falls.

In Section 6, we will examine the above predictions on the behavior of price dynamics and trading activity empirically.

5 Impact of Circuit Breaker on Welfare

So far, we have focused on how the circuit breaker changes market behavior, such as price level and dynamics. In this section, we examine how the circuit breaker impacts the welfare of market participants, which becomes possible in a general equilibrium setting like ours.

The welfare implication of a circuit breaker very much depends on the agents' trading motives and the welfare measure used. If we interpret agent B's preference as a form of a state-dependent utility function under the objective probability measure, as in (7), then there is no ambiguity about the welfare criteria. In this case, trading is for risk sharing, and the introduction of a circuit breaker reduces welfare due to the potential loss of risk sharing opportunities when the market is shut down.

If, instead, we interpret agent B's preference as a state-independent log utility function, as in (3), but under the subjective belief given in (5), then the welfare implication of the circuit breaker is less clear cut. While the literature has proposed different welfare criteria for economies with heterogeneous beliefs,[24] we take the perspective of a paternalistic planner/regulator who considers welfare under the objective probability measure. A main

[24]See e.g., Brunnermeier, Simsek, and Xiong (2014); Gilboa, Samuelson, and Schmeidler (2014).

rationale for policy makers to introduce a circuit breaker is to curb irrational trading, in particular "panic selling," under sudden market downturns. This is captured by the subjective belief deviation δ_t, which is on average zero but can become significant when the market moves substantially within a day. Under such an interpretation, a paternalistic planner/regulator will view the trading motive of agent B as speculation driven by irrational beliefs, and will see welfare gains from introducing a circuit breaker to curb such trading activities.

Welfare measure. To formally examine these issues, we define an agent's welfare by the ex-ante certainty equivalent wealth CE. For agent A, we have:

$$\ln(CE_0^A) = \mathbb{E}_0\left[u^A(W_T^A)\right] = \mathbb{E}_0\left[\ln(W_T^A)\right]. \tag{32}$$

For agent B, under the interpretation of state-dependent utility, we have:

$$\ln(CE_0^B) \equiv \mathbb{E}_0\left[\tilde{u}^B(W_T^B, \eta_T)\right] = \mathbb{E}_0\left[\eta_T \ln(W_T^B)\right], \tag{33}$$

where the second equality follows from the definition of the state-dependent utility $\tilde{u}^B$ in (7). However, under the alternative interpretation of state-independent utility and irrational belief, agent B's certainty-equivalent wealth changes to:

$$\ln(CE_0^B) \equiv \mathbb{E}_0\left[u^B(W_T^B)\right] = \mathbb{E}_0\left[\ln(W_T^B)\right]. \tag{34}$$

Here, the planner calculates agent B's expected utility under the objective probability measure $\mathbb{P}$ instead of his own belief $\mathbb{P}^B$.

The impact of the circuit breaker on welfare can then be defined by the change in agents' certainty equivalent wealth due to its introduction. In particular, we have:

$$\Delta CE^i \equiv CE_0^i - \widehat{CE}_0^i, \quad i = A,\, B, \quad \textit{and} \quad \Delta CE \equiv \omega\, \Delta CE^A + (1-\omega)\, \Delta CE^B, \tag{35}$$

where $\widehat{CE}^i$ denotes agent i's certainty equivalent wealth without the circuit breaker, ΔCE^i its change with the circuit breaker, and ΔCE the change in total welfare of the market, given by the weighted average of individual agents' welfare changes.

In Figure 4, we use percentage change in certainty equivalent wealth to represent the impact of a circuit breaker on the agents' welfare, under the two alternative interpretations of agent B's trading motives. To better demonstrate the results, we set ν to be twice the value of σ in (5), which increases the degree of heterogeneity between agents and potential gains from trading.

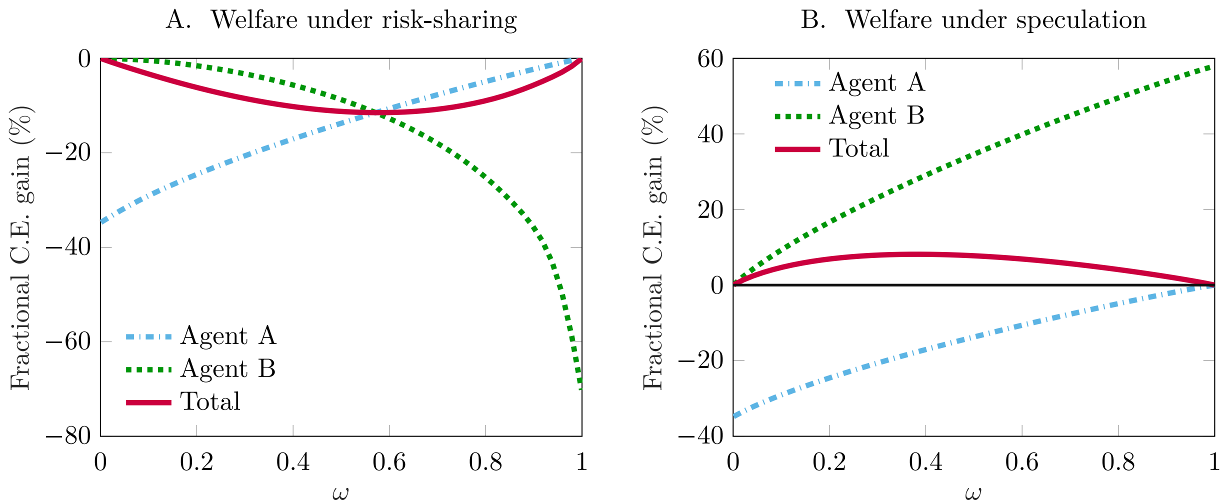

Figure 4: Agents' gain in certainty equivalent wealth in a market with circuit breaker versus a market without for different initial wealth share ω of agent A (the case of time-varying disagreement). The blue dashed line is for agent A and the green dotted line is for agent B. The red solid line is for the whole market, with proper weights for the agents. In Panel A, the certainty equivalent wealth is defined under agents' own beliefs. In Panel B, it is defined under the objective probability measure. The parameters are $\mu = 0.04\%$, $\sigma = 3\%$, $\nu = 6\%$, $\kappa = 0$, $\delta_0 = \bar{\delta} = 0$, $\omega = 0.9$, and $\alpha = 5\%$.

Welfare under risk-sharing. Panel A of Figure 4 plots the circuit breaker's impact on the agents' individual and total welfare. The blue dashed line shows the welfare change for agent A, the green dotted line shows the welfare change for agent B, and the red solid line plots the total welfare change.

For both agents, the introduction of the circuit breaker reduces welfare. This is a straightforward result from our model. In the absence of other market imperfections, a dynamically complete financial market allows efficient risk sharing. This is indeed the case in our setting without the circuit breaker. The presence of the circuit breaker, however, destroys market completeness, lowers the efficiency in risk sharing, and hence reduces welfare. This is what we refer to as the dark side of the circuit breaker, a topic we return to shortly.

In general, an agent's gain from risk sharing increases with the relative size of her/his counter party, who will demand a smaller premium. Thus, the welfare loss from the circuit breaker for agent A reaches its maximum when her wealth share ω approaches 0, it diminishes when ω increases and becomes zero when ω approaches 1. For agent B, the reverse is true – his welfare loss increases with ω. It also worth noting that for small values of ω (when it is close to 0), the loss of welfare for agent B is almost zero. This is because with little wealth, agent A will demand a very large premium from agent B to share his risk, who wants to trade large amounts. In equilibrium, the actual amount of trading is negligible even when the market is open. Thus, the welfare loss from a circuit breaker is limited. This property for risk sharing trades is important when we consider both risk-sharing and speculation trades.

For the market as a whole, the welfare loss is maximized when ω is somewhere in the middle. This is not surprising given the welfare loss for the two agents and their relative weights in the market. The maximum welfare loss can exceed 10% of the total wealth.

Welfare under speculation. Panel B of Figure 4 plots the circuit breaker's welfare impact when agent B is viewed as having a (time-varying) irrational belief. The blue dashed line shows the welfare loss of agent A, as measured by ΔCE^A. As in Panel A, it is always negative and decreases in magnitude with her initial wealth share ω. The green dotted line shows agent B's welfare change, now under the objective probability. Contrast to the welfare loss under his subjective measure, agent B's welfare now increases when a circuit breaker is in place. This is expected. With his incorrect belief, agent B in general incurs losses from his trading with agent A. Therefore, the circuit breaker, when triggered, helps to limit his trading and reduce his loss. Moreover, a larger wealth share of agent A, ω, allows her to take on larger positions against B, leading to larger potential losses for agent B. Consequently, imposing the circuit breaker leads to larger welfare gains for agent B when ω is larger.

For the market as a whole, the net welfare gain is a weighted average of the two agents' welfare gains/loses. In general, it can be positive or negative. However, as shown in panel B of Figure 4 by the red solid line, the net welfare gain from the circuit breaker is always positive. This is because that under the objective probability measure (and the same state-independent log utility function for both agents), the optimal allocation between the two agents is simply no trade. Any trading will then lead to less efficient allocations and lower total welfare. Since the circuit breaker reduces the agents' trading opportunities, especially those for agent B based on an irrational belief, it will increase welfare.

We thus conclude that in a complete market setting, if trading is only driven by agents' irrational beliefs, the circuit breaker helps to curtail the irrational trades and increases welfare for the market as a whole. Under this interpretation, circuit breaker only brings a bright side.

Welfare under general trading motives. The two welfare effects considered above take extreme stands: agents trade either for risk sharing due to heterogeneous preferences or for speculation due to irrational beliefs. As a result, the welfare implication of a circuit breaker is clear cut: bad in the former case and good in the latter case.

In general, both trading motives are present in the market, which leads to richer welfare implications, depending on the equilibrium trade-off between the two effects. We thus extend our model to this more general case. Suppose there are three agents, A, B_1, and B_2. Agent A is the same as in the original model; agent B_1 has the state-dependent utility function as in (7) under the objective probability measure; agent B_2 has the state-independent log utility function as in (3) under the subjective belief given in (5). Furthermore, the initial wealth share of agent A is still ω, while B_1 and B_2 have ω^B and $1-\omega^B$ shares of the remaining wealth. The changes in total welfare due to a circuit breaker is then given by:

$$\Delta CE = \omega\,\Delta CE^A + (1-\omega)\left[\omega^B \Delta CE^{B_1} + (1-\omega^B)\Delta CE^{B_2}\right], \tag{36}$$

where ΔCE^i is the change of certainty equivalent wealth for $i = A, B_1, B_2$, and CE^{B_1} and CE^{B_2} are given by (33) and (34), respectively.

The left panel of Figure 5 shows the changes in total welfare as a function of the wealth share of agent A, ω, and the relative wealth share of agent B_1, ω^B. As expected, the welfare implication of the circuit breaker depends on both the wealth share of agent A as well as the relative wealth share of B_1 and B_2. Such a dependence exhibits rich patterns.

When ω^B is large (close to 1) the rational trader with state-dependent utility is dominant, the circuit breaker reduces the total welfare because the gain from risk-sharing exceeds the loss from irrational speculation. This corresponds to the case shown in Figure 4, panel A. The opposite is true when irrational traders are dominant, i.e., when ω^B is small (close to zero), which corresponds to the case in Figure 4, panel B.

For a given interior value of ω, the net welfare impact of the circuit breaker decreases with ω_B. It starts at a positive value when ω_B is small and more trading is driven by irrational speculation, and turns negative when ω_B becomes sufficiently large and more trading is driven by risk sharing. The turning point, however, depends on the value of ω, the wealth share of agent A. In particular, for a smaller value of ω, the critical value of ω_B for the circuit breaker's net welfare impact to turn negative is larger. This dependence is shown by the zero-value line on the $x-y$ plane. This result is driven by the fact that when ω is small, the welfare loss from a circuit breaker in reduced risk sharing is negligible, as discussed above,

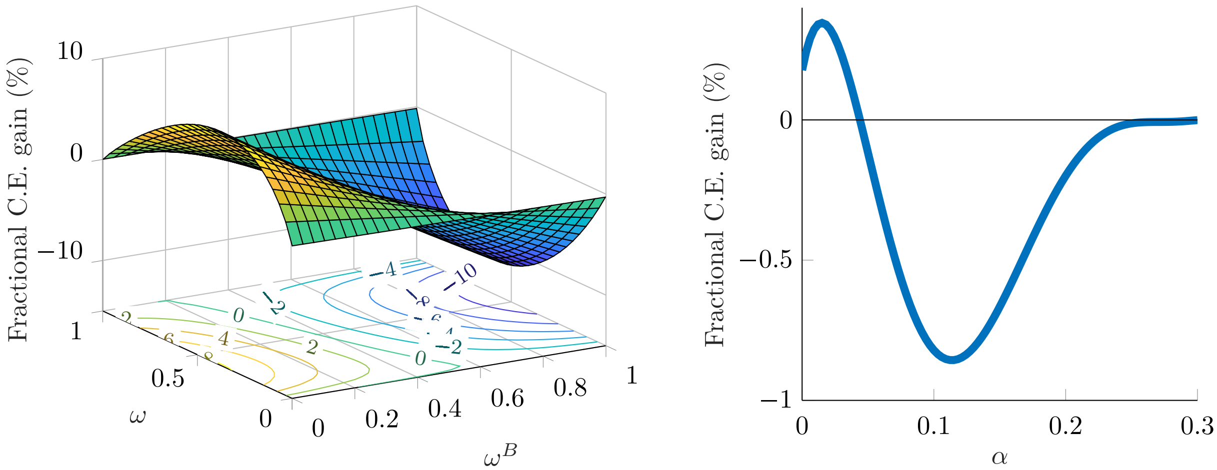

Figure 5: Welfare tradeoffs in the case of time-varying disagreement. The left panel shows changes in total welfare due to the circuit breaker in the presence of two types of trading motives: risk-sharing and irrational speculation. The right panel shows the dependence of changes in total welfare on the circuit breaker level. The parameters are $\mu = 0.04\%$, $\sigma = 3\%$, $\nu = 6\%$, $\kappa = 0$, and $\delta_0 = \bar{\delta} = 0$. For the left panel, $\alpha = 5\%$; for the right panel, $\omega = 0.50$ and $\omega^B = 0.41$.

and the welfare gain in reduced speculation is significant and dominates. For larger values of ω, however, the welfare effect from risk sharing becomes significant and important in influencing the tradeoff.

Optimal circuit breaker design. So far, we have taken the circuit breaker, defined by α, as given and analyze its welfare impact under different distribution of agents with different trading motives. We can also take the distribution of agents as given and examine how different choices of α influence welfare. As an example, the right panel of Figure 5 plots the total welfare gain from a circuit breaker for different values of α, with $\omega = 0.50$ and $\omega^B = 0.41$. The initial wealth shares of the three agents, A, B_1 and B_2, are then 0.5, $(0.5)(0.41) = 0.205$, and $(0.5)(0.59) = 0.295$, respectively. As discussed above, a circuit breaker protects investors with irrational belief from speculation, but also hinders investors with state-dependent utility from full risk sharing. For small values of α the former effect dominates, yielding a welfare gain. For large values of α, the latter effect dominates, yielding a welfare loss.

The tradeoff between these two effects is in general not monotonic and can exhibit rich patterns. In the current case, for example, as α increases from zero, corresponding to the case of no trade, the welfare gain from risk sharing exceeds the welfare loss from irrational speculation. The overall welfare is maximized at an interior α value of 1.2%. Beyond this point, the welfare begins to decrease and turns negative. As α further increases, the welfare loss actually starts to decrease. This example shows that quantitatively the tradeoff between the welfare gain and loss from a circuit breaker can be quite complex.[25]

To the extent that irrationality can also be interpreted as a form of market imperfection, our model can be a useful setup to illustrate the basic mechanisms behind potential pros and cons of circuit breakers. Nonetheless, we will restrain from making specific statements concerning practical welfare and policy implications of circuit breakers. These statements should be based on a model that can quantitatively captures the agents' trading motives and behavior as well as other important imperfections in the actual market.

6 Empirical Analysis

In Section 4, we have examined the model predictions on how market dynamics changes as it approaches the circuit breaker and highlighted four key predictions on the behavior of return volatility, skewness, expected return, and trading volume. In this section, we explore the data from the U.S. stock market and present a set of empirical results that are consistent with these predictions.

6.1 Data and Variables

The sample period we consider is from 05/01/2013 to 12/31/2020. The choice for the start date is due to the fact that the current version of the market-wide circuit breaker (MWCB) for the U.S. was first implemented on April 8, 2013 (see Online Appendix OA.1 for an overview of the evolution of the MWCB rule in the U.S.).

To study market dynamics, we analyze transaction-level data for the E-mini S&P 500 futures from the Chicago Mercantile Exchange (CME). The E-mini S&P 500 futuresis one of the most actively traded financial products in the world, and it closely tracks the S&P 500 index, the reference index for the current MWCB. Our analysis uses the lead month contracts, which are usually the most active in terms of trading volume.[26] Although the E-mini futures are traded almost around the clock, we restrict our attention to regular trading hours (9:30 a.m. to 4:00 p.m. ET) when the MWCB is active.

[25]This example is only for demonstrative purposes. We did not try to search extensively in the parameter space for an "optimal" α with higher values, such as 5% or 7%. Given the parsimonious nature of the model, it is not intended for a quantitative calibration to the actual market. Please also see the discussion below.

[26]All of our main empirical results continue to hold when we follow a deterministic rollover strategy by switching to the next contract one week before maturity.

Since price dynamics can change rapidly as the distance to the circuit breaker changes, it is important to carry our analysis at a sufficiently high frequency, which we set at the minute level. We first compute second-level prices of the index futures using volume-weighted average transaction prices over each second, and then construct measures for (log) return, volatility, and skewness at the minute frequency. To reduce the impact of microstructure noise in the volatility measure, we apply the two-scales realized volatility (TSRV) measure of Zhang, Mykland, and Ait-Sahalia (2005).[27] Our skewness measure is the realized skewness introduced by Amaya et al. (2015).[28]

We also construct a measure of abnormal trading volume at the minute frequency, which is defined as follows:

$$\text{Abnormal volume}_t \equiv \frac{\text{volume in minute } t}{\text{average volume in the previous 6.5 trading hours}} - 1. \tag{37}$$

We exclude the abnormal volume from the first and last day of each contract because the trading volume tends to spike and fall, respectively, on these dates.

We construct a "Distance to Circuit Breaker" (DTCB) measure to determine how far away the market is from the nearest circuit breaker threshold. It is worth noting that, in addition to the MWCB, the CME also imposes downside price limits for the E-mini S&P 500 future, with three levels of price limits that exactly match the three levels of circuit breaker thresholds during regular trading hours,[29] and we need to take into account a subtle difference between the two rules. Whereas the MWCB focuses on the distance between the S&P 500 index and its closing value on the previous day, the price limits on the E-mini focus on the distance between the E-mini price and its volume-weighted average price (VWAP) between 3:59:30 p.m.–4:00 p.m. ET on the previous day. Although these two distances are nearly identical most of the time, it is possible for the E-mini price limit to be reached before the MWCB. To ensure consistency, we define $DTCB_t$ based on the E-mini price and filter

[27] The TSRV realized volatility is constructed using the difference between two estimates of 1-min integrated realized volatility: one based on 1-sec returns and the other based a scaled average of 1-min integrated realized volatility computed using 5-sec returns. Taking the difference between the two estimates helps remove microstructure noise. In Online Appendix OA.6, we show that our main results are robust to different measures of volatility, including TSRV with different sampling frequency or simply the raw volatility measure.

[28] Amaya et al. (2015) define the realized skewness as $\text{RDSkew}_t = \left(\frac{1}{N}\sum_{i=1}^{N} r_{t,i}^3\right) / \left(\frac{1}{N}\sum_{i=1}^{N} r_{t,i}^2\right)^{3/2}$, where N is the number of observations in a minute.

[29] See https://www.cmegroup.com/trading/equity-index/sp-500-price-limits-faq.html for more information on the price limits.

out observations for which the price limit was reached before the MWCB.[30] Specifically, at minute t,

$$DTCB_t \equiv \frac{P_t^{min} - (1-\alpha_t)\bar{P}_t^{close}}{\bar{P}_t^{close}}. \tag{38}$$

Here, P_t^{min} is the *minimum price level* for the E-mini during minute t; $\bar{P}_t^{close}$ denotes the volume-weighted average E-mini price for the last 30 seconds before 4:00 p.m. from the prior trading day; α_t is the nearest active circuit breaker threshold. Specifically, α_t is equal to 7% at the beginning of a trading day; if the Level 1 circuit breaker is triggered before 3:25 p.m., α_t is raised to 13% after the market reopens; finally, α_t is raised to 20% after 3:25 p.m. or after the Level 2 circuit breaker is triggered. To reduce the noise caused by the discrete change in the circuit breaker threshold at 3:25 p.m., we drop observations post 3:25 p.m. in our analysis.

Finally, when examining the relationship between DTCB and conditional return moments or trading volume, we want to control for the well-known leverage effect on return volatility, which refers to the potential negative correlation between stock volatility and price change (see Black, 1976; Yu, 2005; Ait-Sahalia, Fan, and Li, 2013, among others). For this purpose, we construct the following measure of how the average price level in the past 60 minutes compares to a moving average price level from the past 21 trading days:

$$Lev_t = \frac{\text{Average E-mini price over the past 60 minutes}}{\text{Average E-mini price over the past 21 trading days}}. \tag{39}$$

To capture potential nonlinearity in the leverage effect, we also introduce a quadratic leverage factor as $QLev_t = (Lev_t - 1)^2$.

The summary statistics of the main variables and their correlation matrix are reported in the Online Appendix (OA.6). Two points are worth noting. First, the first percentile of DTCB is 4.47%. This is because intraday movements in the SPX (or E-mini) of 2.5% or more are rare. However, with a sample size of 686,697, we still have 248 minute-level observations with DTCB below 2%, which give us a window to examine market behavior near a MWCB.[31] Second, the correlation between DTCB and Lev (QLev) is 0.194 (-0.026). Despite the fact both the distance to circuit breaker and leverage factor are driven by price movements, there are important differences between them. In Lev, we measure "current price levels" using

[30] This filter removes eight observations in total from two dates, August 24, 2015 and March 9, 2020.

[31] For robustness, we also examine the empirical results when the four days in March 2020 are excluded and they remain the same (see Online Appendix OA.6).

average prices over the past 60 minutes rather than the past minute. More importantly, the benchmark for the current price level in Lev is the average price level over the past 21 days. In contrast, in DTCB we reset the benchmark daily to the previous closing price. It is thus possible that, at some moments on two different trading days, the levels of the SPX are similar, resulting in similar values for Lev, yet the DTCB are drastically different due to different closing prices on the previous day.

6.2 Empirical Results

For a first look at the data, we run the following regressions of return moments (volatility, skewness, mean return) and abnormal volume on a set of DTCB dummies while controlling for time trend, intraday seasonality,[32] and leverage effects,

$$y_{d,m,t} = \sum_i \beta_i \times 1_{\{DTCB \in \mathrm{Bin}_i\}} + a_d + b_m + c\, Lev_t + d\, QLev_t + \varepsilon_{dmt}, \tag{40}$$

where $y_{d,m,t}$ is the volatility, skewness, abnormal volume on day d, minute m, time t, all of which are contemporaneous with the right-hand-side variables, or the realized return in the next minute. The bins are defined as follows: $\mathrm{Bin}_1 = (0\%, 0.75\%]$, $\mathrm{Bin}_2 = (0.75\%, 1.25\%]$, $\mathrm{Bin}_3 = (1.25\%, 2\%]$, and $\mathrm{Bin}_i = ((i-1)\%, i\%]$ for $i = 4, \ldots, 14$. The controls include daily and minute-of-the-day fixed effects (a_d, b_m), as well as the linear and quadratic leverage factor (Lev_t, $QLev_t$). We then plot the coefficients β_i to examine how the variables of interest vary with the distance to circuit breaker.

As Figure 6 shows, both realized volatility (panel A) and expected return (panel C) for E-mini rise as DTCB becomes smaller, and their speeds of rise accelerate (the slopes steepen) when DTCB approaches zero. These patterns are consistent with the model predictions discussed in Section 4 (see panels A and C of Figure 3). Panel B shows that realized skewness first declines as DTCB becomes smaller and then starts to rise when DTCB is below 2%. This pattern is qualitatively the same as the model prediction in panel B of Figure 3, although the uptick in skewness when DTCB is near zero appears stronger in the data. Panel D of Figure 6 shows that abnormal volume steadily rises as DTCB decreases, then drops when DTCB reaches the last bin before zero. The final dip in volume is likely due to the deterioration of market liquidity when the market approaches the circuit breaker threshold, a feature that

[32]For example, both volatility and trading volume have well-documented U-shaped intraday patterns (see e.g., Hong and Wang, 2000).

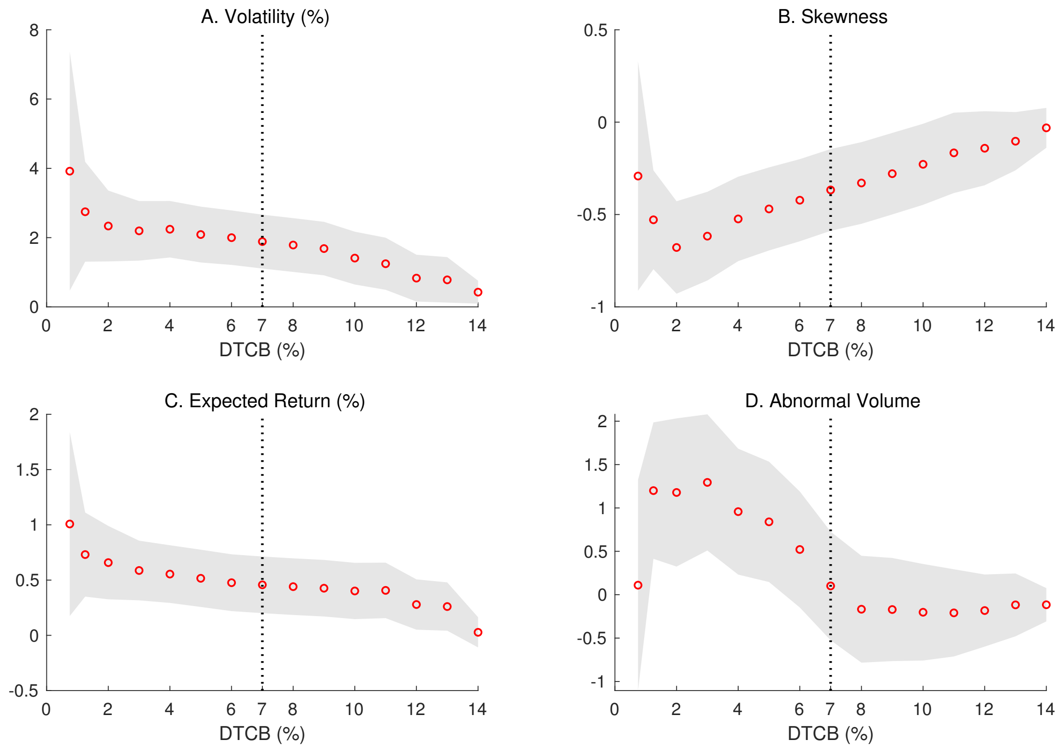

Figure 6: This figure shows the β regression coefficients of (40) for volatility (two scales realized volatility), skewness, expected return, and abnormal volume. $\text{Bin}_1 = (0\%, 0.75\%]$, $\text{Bin}_2 = (0.75\%, 1.25\%]$, $\text{Bin}_3 = (1.25\%, 2\%]$, and $\text{Bin}_i = ((i-1)\%, i\%]$ for $i = 4, \ldots, 14$. The red dots indicate β coefficients and grey lines represent two standard deviation around.

our model abstracts from. Finally, notice that the confidence bands on the β estimates widen significantly when DTCB is near zero due to limited observations in this region.

We further examine the statistical significance of the empirical patterns demonstrated in Figure 6 using the following regression specification:

$$y_{d,m,t} = \beta \times DTCB_t + \gamma \times \min(DTCB_t - 2\%, 0) \\ + a_d + b_m + c\, Lev_t + d\, QLev_t + Y_t'\delta + \varepsilon_{dmt}, \tag{41}$$

where $y_{d,m,t}$ is the realized volatility, skewness, and abnormal volume on day d, minute m, and time t; when examining the prediction about expected return, we replace $y_{d,m,t}$ with $r_{d,m,t+1}$, the realized E-mini log return in the next minute. For controls, similar to (40), we include daily and minute-of-day fixed effects (a_d, b_m), as well as the linear and quadratic leverage factor (Lev_t, $QLev_t$). In addition, we include Y_t, a vector of lagged values of $y_{d,m,t}$, to control for serial correlations between y at different lags. To determine the number of lags to include in the regressions, we run an AR model selection based on AIC.[33] The main

[33]For the AR models with optimally selected lags, we verify that no remaining AR or MA structure is

coefficients of interest are β, which captures the baseline relationship between $y_{d,m,t}$ and $DTCB_t$, and γ, which captures the potential non-linearity when DTCB is below 2%. Table 2 reports the regression results.

There is indeed a highly significant and sizable negative relationship between DTCB and realized volatility. On average, when the SPX E-mini price moves 1% closer to the circuit breaker, the realized volatility in the E-mini futures rises by 17.5 bps (normalized to daily). Moreover, the slope coefficient on DTCB nearly doubles in magnitude as DTCB drops below 2%. The coefficient estimate is highly significant based on OLS standard errors, but insignificant based on Newey-West standard errors. This is because there are a total of 248 observations for which DTCB is below 2%. The majority of them are concentrated in a few windows from a few days in our sample, for which the correction for autocorrelation across observations makes a big difference.

Notice that when we exclude DTCB from the regression, the coefficient on the leverage factor is negative and statistically significant, suggesting that there is indeed a leverage effect in our sample. However, when DTCB is included, the coefficient on the leverage factor turns positive, likely because DTCB has subsumed the leverage effect from the leverage factor. However, there are a few reasons for why DTCB is not just a leverage factor in disguise. First, the magnitude of the leverage effect is much smaller. When DTCB is not included in the regression, a one-standard deviation decrease in the leverage factor corresponds to a 7 bps increase in the realized volatility, about half the size of the effect from a one-standard deviation decrease in DTCB. Second, the slope coefficient on DTCB becoming more negative as DTCB declines is a unique prediction of our model on volatility dynamics in the presence of a circuit breaker, which is present in the data even after controlling for nonlinearity in the leverage factor.

Next, the skewness of E-mini returns is positively related to DTCB. When the SPX moves 1% closer to the active circuit breaker, the realized skewness drops by 0.12. Consistent with Figure 6, the coefficient γ is large and positive, enough to turn the overall relation between DTCB and skewness positive when when DTCB is below 2%. This pattern is qualitatively consistent with what is shown in Panel B of Figure 3 and Prediction (ii).

There is also a negative relationship between DTCB and the following 1-minute return

detected in the residuals.

Table 2: Regression results. This table reports the results of the regression

$$y_{d,m,t} = \beta \times DTCB_t + \gamma \times \min(DTCB_t - 2\%, 0) + a_d + b_m + c\, Lev_t + d\, QLev_t + Y_t'\delta + \varepsilon_{dmt},$$

where $y_{d,m,t}$ is the realized volatility, skewness, and abnormal volume on day d, minute m, and time t or return on time $t+1$, and Y is a vector of lagged values of $y_{d,m,t}$. OLS and Newey-West standard errors (with 29 lags) are reported in parentheses and square brackets, respectively. Sample period: 2013/05/01 – 2020/12/31. The total number of observations is 686,697.

	Volatility			Skewness			Return			Abnormal volume		
Lev	−2.843	8.555	8.441	−8.401	−16.761	−17.149	−2.648	1.116	1.039	0.166	24.004	24.687
	(0.170)***	(0.226)***	(0.229)***	(0.237)***	(0.328)***	(0.331)***	(0.076)***	(0.106)***	(0.107)***	(0.449)	(0.664)***	(0.669)***
	[0.304]***	[0.888]***	[0.744]***	[0.329]***	[0.672]***	[0.642]***	[0.144]***	[0.380]***	[0.316]***	[0.570]	[1.458]***	[1.392]***
QLev	−6.439	2.075	2.590	−17.945	−23.863	−22.293	−1.911	1.104	1.406	9.391	26.328	24.074
	(1.280)***	(1.280)	(1.287)	(1.804)***	(1.809)***	(1.819)***	(0.572)***	(0.574)*	(0.577)**	(3.442)***	(3.453)***	[3.463]***
	[6.675]	[6.563]	[6.594]	[2.300]***	[4.104]***	[3.993]***	[3.026]	[2.872]	[2.920]	[3.235]***	[8.665]***	[8.426]***
DTCB		−17.454	−17.263		12.085	12.701		−5.375	−5.254		−28.290	−29.277
(Distance to CB)		(0.230)***	(0.236)***		(0.328)***	(0.336)***		(0.107)***	(0.109)***		(0.582)***	(0.594)***
		[1.256]***	[1.050]***		[0.698]***	[0.628]***		[0.474]***	[0.381]***		[1.561]***	[1.454]***
$\min\{DTCB - 2\%, 0\}$			−12.684			−40.883			−7.656			69.168
			(3.443)***			(4.892)***			(1.543)***			(8.351)***
			[44.433]			[11.597]***			[15.880]			[17.508]***
r_t							−0.015	−0.012	−0.012			
							(0.001)***	(0.001)***	(0.001)***			
							[0.004]***	[0.004]***	[0.004]***			
Control for lags	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Date + Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.664	0.667	0.667	0.003	0.005	0.006	0.002	0.006	0.006	0.591	0.593	0.593

Note: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

in the SPX E-mini futures, which is again consistent with the model prediction that the expected return rises when the market approaches the circuit breaker. Like volatility, the negative relationship between DTCB and expected return strengthens when DTCB is below 2%. The coefficient estimate more than doubles from -5.254 when DTCB is above 2% to -12.91 when DTCB is below 2%, but the difference is statistically insignificant based on the Newey-West standard error.

Finally, we find a negative and statistically significant relationship between DTCB and abnormal trading volume, although the relation appears to flatten when when DTCB drops below 2%. Although we are not able to directly test the prediction about deleveraging by some investors, the findings on trading volume are consistent with elevated investor trading demand as the market approaches the circuit breaker.

7 Model Robustness

In order to better illustrate how a circuit breaker impacts the market, our analysis has focused on a parsimonious version of our model. For example, the bond supply is assumed to be zero. The form of the circuit breaker is one-sided and only determined by price levels. Our set-up, however, allows more general specifications. In this section, we examine the robustness of our results when the bond is in positive supply.[34]

Positive bond supply can potentially change the equilibrium behavior because, upon triggering the circuit breaker, the two agents are no longer required to be fully invested in the stock for the market to clear. For example, if the optimistic agent is sufficiently wealthy, she could hold the entire stock market without taking on leverage at the market closure. Consequently, the pessimistic investor need not be the only marginal investor at that instant.

Instead, depending on the wealth distribution and the amount of bond supply, there are four possible scenarios at the market closure time τ: (i) the optimistic agent faces binding leverage constraint, while the pessimistic agent is unconstrained; (ii) the optimistic agent is unconstrained, while the pessimistic agent faces binding short-sale constraint; (iii) the

[34]We have also considered other extensions of the model, including the differences between continuous-time and discrete-time trading and different types of circuited breakers such as two-sided circuit breakers, multi-tiered circuit breakers, and circuit breakers based on triggers beyond price levels. These results are available from the authors upon request. Contrast to trading halts triggered by price-level thresholds, like circuit breakers, markets typically have pre-scheduled trading halts like overnight closures. In general, their impact on market behavior is different. Such a comparison is provided in the Online Appendix (OA.5.

optimistic agent faces binding leverage constraint, while the pessimistic agent faces binding short-sale constraint; and (iv) neither agent is constrained. In particular, Scenario (ii) is the opposite of Scenario (i), in which case the equilibrium price level can become higher and volatility lower in the economy with a circuit breaker. We now examine the conditions (wealth distribution and size of bond supply) that determine which of the four scenarios occur in equilibrium, and how they affect our model's main prediction that the presence of circuit breakers amplify return volatility.

As in Section 3.2, we first consider the problem at the instant before market closure, which will provide us with much of the intuition about the effect of positive bond supply. Suppose the total supply of the bond is Δ. At time τ, the two agents have one last chance to trade and then have to hold onto their positions until time T. Similar equilibrium conditions as those given in Section 3.2 apply here, with the exception that $\phi_\tau^A + \phi_\tau^B = \Delta$.

For illustration, in Figure 7 we plot the equilibrium stock price as a function of the wealth share of agent A for $\tau = 0.25$ and $D_\tau = 0.97$ (blue solid line), and compare it to the stock price without the circuit breaker (red dotted line). The bond supply is assumed to be $\Delta = 0.17$. Notice that because of the one-to-one mapping between D_t and δ_t, we know that agent A is relatively more optimistic for this value of D_τ.

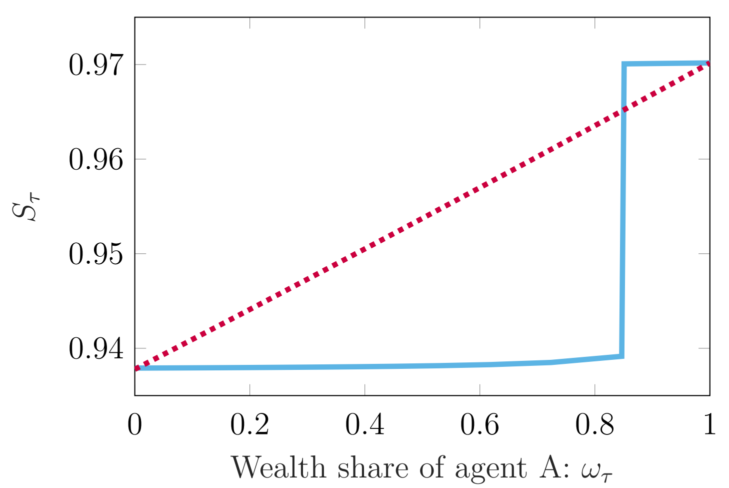

Figure 7: Stock price at time τ with positive bond supply. Blue solid lines are for the case with circuit breaker. Red dotted lines are for the case without circuit breaker. The parameters are $\tau = 0.25$, $D_\tau = 0.97$, and $\Delta = 0.17$.

Again, the stock price in the market without circuit breaker is a weighted average of the valuations of the optimist and pessimist, with the weight depending on their respective wealth shares. Since agent A is more optimistic, the stock price without circuit breaker naturally increases in her wealth share. When $\Delta = 0$, the stock price with circuit breaker will always be equal to the valuation of agent B (the pessimist) at time τ. However, this is no longer the case with $\Delta > 0$. As Figure 7 shows, when agent A's wealth share is not too high, the stock price with circuit breaker will be at or slightly above agent B's valuation. However, when her wealth share is sufficiently high, the stock price will rise sharply and reach the valuation of agent A (the optimist).

The intuition is as follows. When agent A's wealth share ω_τ^A is not too high, she will be fully invested in the stock at time τ but is not able to clear the stock market by herself. In this case, agent A's leverage constraint will be binding, agent B will hold all the riskless bond and the remaining stock, and the market clearing price has to agree with agent B's (the pessimist) valuation. This is Scenario (i) listed above. The reason that the pessimist valuation increases slightly with ω_τ^A is that, as agent A becomes wealthier, she will hold more of the stock, making agent B's portfolio at time τ less risky.

When agent A's wealth share becomes sufficiently high, she will be able to hold the entire stock market without borrowing. In such cases, agent A will hold all of the stock and potentially some bonds, while agent B will invest all of his wealth in the bonds. As long as the stock price is above agent B's private valuation, he would want to short the stock, but the short-sale constraint would be binding (an arbitrarily small short position can lead to negative wealth). Consequently, agent A (the optimist) becomes the marginal investor at τ, and the market clearing price has to agree with her valuation.[35] This equilibrium is qualitatively different. The switch of the marginal investor from the pessimist to the optimist means that the price-dividend ratio will be higher and conditional return volatility lower with circuit breaker.

The above analysis highlights the key differences between the economies with positive and zero bond supply. When $\Delta > 0$, the circuit breaker equilibrium is similar to the case with $\Delta = 0$ as long as agent A's wealth share is not dominant. However, if agent A's wealth share becomes sufficiently high, the property of the equilibrium changes qualitatively, with price level becoming higher and volatility lower in the presence of a circuit breaker.

[35] There is also a knife-edge case where agent A does not hold any bonds, the stock price is below her valuation, but she faces binding leverage constraint. In this case, the stock price and the wealth distribution have to satisfy the condition $\omega_\tau^A = \frac{S_\tau}{S_\tau + \Delta}$.

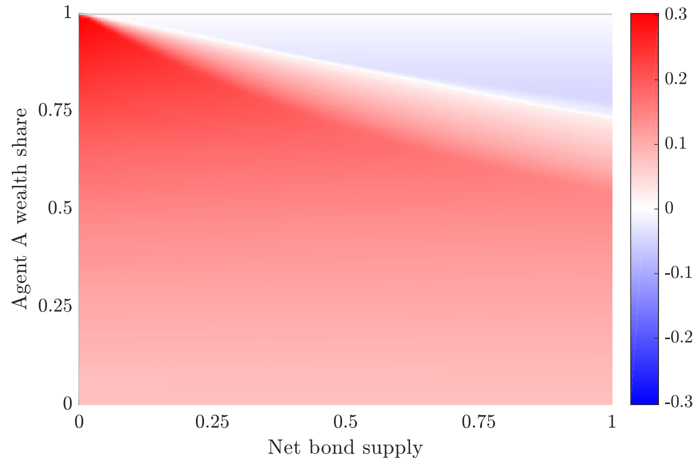

Figure 8: Heat map for the ratio of average daily volatilities in the economies with and without circuit breakers. For normalization, we subtract one from the ratios.

To comprehensively examine the different cases, in Figure 8 we present a heat map for the average daily return volatility in the economies with and without circuit breaker. It covers a wide range of bond supply (Δ) and initial wealth share for agent A (ω). The red region indicates the scenarios under which the introduction of a circuit breaker amplifies market volatility, whereas the blue region indicates the opposite. Quantitatively, the volatility amplification effect occurs for most parameter areas and is stronger when the net bond supply is small relative to the net supply of the stock, and when agent A's initial wealth share is not too low or too high.

For the actual market, the net supply of riskless bonds relative to the stock market is likely small. For example, the market for equity is about $41 trillion in 2020, while the total size of the U.S. corporate bond market is about $10 trillion, of which about $3.5 trillion are rated at A or above. If we treat these highly-rated bonds as essentially riskless, the relative size of the market for riskless bonds would be 9%. At such level of riskless bond supply, we still get volatility amplification for most values of ω.

8 Conclusion

In this paper, we build a dynamic model to examine the mechanism through which market-wide circuit breakers affect trading and price dynamics in the stock market. As we show, a downside circuit breaker tends to lower the price-dividend ratio, reduce daily price ranges, but increase conditional and realized volatility. It also raises the probability of the stock price reaching the circuit breaker limit as the price approaches the threshold (the "magnet effect").

The effects of circuit breakers can be further amplified when some agents' willingness to hold the stock is sensitive to recent shocks to fundamentals, which can be due to behavioral biases, institutional constraints, etc.

Our results demonstrate some of the negative impacts of circuit breakers even without any other market frictions, and they highlight the source of these effects, namely the tightening of leverage constraint when levered investors cannot rebalance their portfolios during trading halts. These results also shed light on the design of circuit breaker rules. Using historical price data from a period when circuit breakers have not been implemented can lead one to significantly underestimate the likelihood of triggering a circuit breaker, especially when the threshold is relatively tight.

References

Abel, A. B. 1990. Asset Prices under Habit Formation and Catching Up with the Joneses. *American Economic Review* 80:38–42.

Ackert, L. F., B. Church, and N. Jayaraman. 2001. An experimental study of circuit breakers: the effects of mandated market closures and temporary halts on market behavior. *Journal of Financial Markets* 4:185–208.

Ait-Sahalia, Y., J. Fan, and Y. Li. 2013. The leverage effect puzzle: Disentangling sources of bias at high frequency. *Journal of Financial Economics* 109:224–49.

Amaya, D., P. Christoffersen, K. Jacobs, and A. Vasquez. 2015. Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics* 118:135–67.

Bayrakter, E., Q. Song, and J. Yang. 2011. On the continuity of stochastic exit time control problems. *Stochastic Analysis and Applications* 29:1554–74.

Berkman, H., O. W. Steenbeek, et al. 1998. The influence of daily price limits on trading in nikkei futures. *Journal of Futures Markets* 18:265–79.

Bertero, E., and C. Mayer. 1990. Structure and performance: Global interdependence of stock markets around the crash of october 1987. *European Economic Review* 34:1155–80.

Black, F. 1976. Studies of stock price volatility changes. *In: Proceedings of the 1976 Meetings of the American Statistical Association* 171–81.

Blume, L., T. Cogley, D. Easley, T. Sargent, and V. Tsyrennikov. 2018. A case for incomplete markets. *Journal of Economic Theory* 178:191–221.

Boyarchenko, N., L. C. Larsen, and P. Whelan. 2020. The Overnight Drift. Staff Reports 917, Federal Reserve Bank of New York.

Brennan, M. J. 1986. A theory of price limits in futures markets. *Journal of Financial Economics* 16:213–33.

Brunnermeier, M. K., A. Simsek, and W. Xiong. 2014. A Welfare Criterion For Models With Distorted Beliefs. *The Quarterly Journal of Economics* 129:1753–97.

Chen, C., and J.-L. Jeng. 1996. The impact of price limits on foreign currency futures' price volatility and market efficiency. *Global Finance Journal* 7:13–25.

Chen, H., S. Joslin, and N.-K. Tran. 2010. Affine Disagreement and Asset Pricing. *American Economic Review* 100:522–6.

Chen, T., Z. Gao, J. He, W. Jiang, and W. Xiong. 2017. Daily price limits and destructive market behavior. *Forthcoming, Journal of Econometrics* .

Chen, Y.-M. 1993. Price limits and stock market volatility in taiwan. *Pacific-Basin Finance Journal* 1:139–53.

Christie, W. G., S. A. Corwin, and J. H. Harris. 2002. Nasdaq trading halts: The impact of market mechanisms on prices, trading activity, and execution costs. *The Journal of Finance* 57:1443–78.

Corwin, S. A., and M. L. Lipson. 2000. Order flow and liquidity around nyse trading halts. *The Journal of Finance* 55:1771–801.

Coursey, D. L., and E. A. Dyl. 1990. Price limits, trading suspensions, and the adjustment of prices to new information. *Review of Futures Markets* 9:342–60.

Darling, R., and E. Pardoux. 1997. Backwards sde with random terminal time and applications to semilinear elliptic pde. *Annals of Probability* 25:1135–59.

Detemple, J., and S. Murthy. 1994. Intertemporal Asset Pricing with Heterogeneous Beliefs. *Journal of Economic Theory* 62:294–320.

Gilboa, I., L. Samuelson, and D. Schmeidler. 2014. No-betting-pareto dominance. *Econometrica* 82:1405–42.

Goldstein, M. A., and K. A. Kavajecz. 2004. Trading strategies during circuit breakers and extreme market movements. *Journal of Financial Markets* 7:301–33.

Gomber, P., M. Haferkorn, M. Lutat, and K. Zimmermann. 2012. The effect of single-stock circuit breakers on the quality of fragmented markets. In *International Workshop on Enterprise Applications and Services in the Finance Industry*, 71–87. Springer.

Greenwald, B., and J. Stein. 1988. The task force report: The reasoning behind the recommendations. *The Journal of Economic Perspectives* 2:3–23.

Greenwald, B. C., and J. C. Stein. 1991. Transactional risk, market crashes, and the role of circuit breakers. *Journal of Business* 443–62.

Grossman, S. J. 1990. Introduction to nber symposium on the october 1987 crash. *The Review of Financial Studies* 3:1–3.

Hong, H., and J. Wang. 2000. Trading and Returns under Periodic Market Closures. *Journal of Finance* 55:297–354.

Jeanblanc, M., M. Yor, and M. Chesney. 2009. *Mathematical methods for financial markets*. Springer Science & Business Media.

Jiang, C., T. McInish, and J. Upson. 2009. The information content of trading halts. *Journal of Financial Markets* 12:703–26.

Karatzas, I., and S. E. Shreve. 1991. *Brownian motion and stochastic calculus*, vol. 113 of *Graduate Texts in Mathematics*. Second ed. Springer-Verlag, New York.

Kim, K. A., and S. Rhee. 1997. Price limit performance: evidence from the tokyo stock exchange. *the Journal of Finance* 52:885–901.

Kodres, L. E., and D. P. O'Brien. 1994. The existence of pareto-superior price limits. *The American Economic Review* 919–32.

Kuserk, G. J., P. R. Locke, and C. L. Sayers. 1992. The effects of amendments to rule 80a on liquidity, volatility, and price efficiency in the s&p 500 futures. *Journal of Futures Markets* 12:383–409.

Lauterbach, B., and U. Ben-Zion. 1993. Stock market crashes and the performance of circuit breakers: Empirical evidence. *The Journal of Finance* 48:1909–25.

Lee, C., M. J. Ready, and P. J. Seguin. 1994. Volume, volatility, and new york stock exchange trading halts. *The Journal of Finance* 49:183–214.

Ma, C. K., R. P. Rao, and R. S. Sears. 1989a. Limit moves and price resolution: the case of the treasury bond futures market. *Journal of Futures Markets* 9:321–35.

———. 1989b. Volatility, price resolution, and the effectiveness of price limits. *Journal of Financial Services Research* 3:165–99.

Roll, R. 1988. The international crash of october 1987. *Financial Analysts Journal* 44:19–35.

———. 1989. Price volatility, international market links, and their implications for regulatory policies. *Journal of Financial Services Research* 3:211–46.

Santoni, G. J., and T. Liu. 1993. Circuit breakers and stock market volatility. *Journal of Futures Markets* 13:261–77.

Securities and Exchange Commission. 1998. Trading analysis of october 27 and 28, 1997: A report by the division of market regulation u.s. securities and exchange commission. https://www.sec.gov/news/studies/tradrep.htm.

Slezak, S. 1994. A theory of dynamics of security returns around market closures. *The Journal of Finance* 49:1163–211.

Subrahmanyam, A. 1994. Circuit breakers and market volatility: A theoretical perspective. *The Journal of Finance* 49:237–54.

———. 1995. On rules versus discretion in procedures to halt trade. *Journal of Economics and Business* 47:1–16.

Wang, J. 1994. A model of competitive stock trading volume. *Journal of Political Economy* 102:127–68.

Yu, J. 2005. On leverage in a stochastic volatility model. *Journal of Econometrics* 127:165–78.

Zapatero, F. 1998. Effects of financial innovations on market volatility when beliefs are heterogeneous. *Journal of Economic Dynamics and Control* 22:597–626.

Zhang, L., P. A. Mykland, and Y. Ait-Sahalia. 2005. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* 100:1394–411.

Appendix

In this appendix, we present the key components of the proofs for the analytical results in Sections 3 and 4, while deferring some standard derivations and proofs for more technical claims to the Online Appendix.

A.1 Proof of Proposition 1

From the solution to the planner's problem (10), deriving the results simply utilizes the (exponential) affine structure of (D, δ), which is given in Online Appendix OA.2.1.

A.2 Proof of Proposition 2

We first prove that $S_\tau \leqslant \min\left\{\widehat{S}^A_\tau, \widehat{S}^B_\tau\right\}$, and then show that S_τ cannot be strictly less than the minimum value. Suppose the market closes at time $\tau < T$. The two agents' optimization problems at time τ are specified in (19), together with the portfolio constraints $\theta^i_\tau \geqslant 0$ and $\phi^i_\tau \geqslant 0$ as implied by the Inada condition. The Lagrangian for agent i is:

$$L = \mathbb{E}^i_\tau\left[\ln\left(\theta^i_\tau D_T + \phi^i_\tau\right)\right] + \zeta^i\left(W^i_\tau - \theta^i_\tau S_\tau - \phi^i_\tau\right) + \xi^i_\theta\,\theta^i_\tau + \xi^i_\phi\,\phi^i_\tau, \quad i \in A, B,$$

and the first order conditions with respect to θ^i_τ and ϕ^i_τ are:

$$0 = \mathbb{E}^i_\tau\left[\frac{D_T}{\theta^i_\tau D_T + \phi^i_\tau}\right] - \zeta^i S_\tau + \xi^i_\theta, \quad 0 = \mathbb{E}^i_\tau\left[\frac{1}{\theta^i_\tau D_T + \phi^i_\tau}\right] - \zeta^i + \xi^i_\phi. \tag{A.1}$$

Furthermore, the market clearing conditions at time τ are given in (20).

First, consider the events when agent A's valuation of the stock in a single-agent economy is higher than that of agent B, i.e., $\widehat{S}^A_\tau(\omega) \geqslant \widehat{S}^B_\tau(\omega)$, for some ω. This implies the condition $\delta_\tau \leqslant \underline{\delta}(\tau) = -\frac{a(\tau,T)}{b(\tau,T)}$ holds for these ω. The other case, where $\widehat{S}^A_\tau(\omega) < \widehat{S}^B_\tau(\omega)$ for some ω, can be proved similarly. We examine the following three scenarios:

1. Agent B (pessimist) is unconstrained and finds it optimal to hold all of his wealth in the stock; agent A (optimist) would like to put more than 100% of his wealth in the stock but faces a binding leverage constraint. In this case,

   $$\theta^{i\,*}_\tau = W_{i,\tau}/S_\tau, \quad \phi^{A*}_\tau = \phi^{B*}_\tau = 0; \quad \xi^A_\theta = \xi^B_\theta = 0, \quad \xi^A_\phi > 0, \quad \xi^B_\phi = 0,$$

   where $\cdot^*$ denotes the optimal holdings. Then from the FOC (A.1) of the unconstrained agent B, we get:

   $$S_\tau = \frac{\mathbb{E}^B_\tau[D_T/(\theta^{B*}_\tau D_T + \phi^{B*}_\tau)]}{\mathbb{E}^B_\tau\left[1/(\theta^{B*}_\tau D_T + \phi^{B*}_\tau)\right]} = \frac{1}{\mathbb{E}^B_\tau\left[1/D_T\right]} = \widehat{S}^B_\tau.$$

2. The price is so low that both agents would prefer to take levered positions in the stock. But the circuit breaker constrains both agents from borrowing. In this case, both agents submit demands proportional to their wealth:

   $$\theta^{i\,*}_\tau = W_{i,\tau}/S_\tau, \quad \phi^{A*}_\tau = \phi^{B*}_\tau = 0; \quad \xi^A_a = \xi^B_a = 0, \quad \xi^A_b > 0, \quad \xi^B_b > 0,$$

and the market for the stock clears. Hence,

$$S_\tau = \frac{\mathbb{E}_\tau^B[D_T/(\theta_\tau^{B*}D_T + \phi_\tau^{B*})]}{\mathbb{E}_\tau^B\left[1/(\theta_\tau^{B*}D_T + \phi_\tau^{B*})\right] + \xi_b^B} = \frac{1}{\mathbb{E}_\tau^B\left[1/D_T\right] + \xi_b^B\theta_\tau^{B*}} < \frac{1}{\mathbb{E}_\tau^B\left[1/D_T\right]} = \widehat{S}_\tau^B.$$

3. For any $S_\tau > \widehat{S}_\tau^B$, agent B will prefer to hold less than 100% of the wealth in the stock. Agent A will need to take levered position in order to clear the stock market but cannot because of the circuit breaker. Thus, this cannot be an equilibrium.

Next we prove that the market clearing price is given by (25) before τ. To this end, agent's indirect utility functions are

$$V^i(W^i_{\tau\wedge T}, \tau\wedge T) = \begin{cases} \ln(W_T^i), & \text{if} \quad \tau \geqslant T \\ V^i(W_\tau^i, \tau), & \text{if} \quad \tau < T. \end{cases} \tag{A.2}$$

Substituting them into the planner's problem (21) and taking the first order condition, we get the wealth of agent A at time $\tau\wedge T$:

$$W^A_{\tau\wedge T} = \frac{\omega S_{\tau\wedge T}}{\omega + (1-\omega)\,\eta_{\tau\wedge T}}. \tag{A.3}$$

Taking the equilibrium allocation $W^A_{\tau\wedge T}$, the state price density for agent A at time $\tau\wedge T$ can be expressed as his marginal utility of wealth times a constant ξ: $\pi^A_{\tau\wedge T} = \xi \left.\frac{\partial V^A(W,\tau\wedge T)}{\partial W}\right|_{W=W^A_{\tau\wedge T}}$. The price of the stock at any time $t \leqslant \tau\wedge T$ is then given by:

$$S_t = \mathbb{E}_t\left[\frac{\pi^A_{\tau\wedge T}}{\pi^A_t} S_{\tau\wedge T}\right], \tag{A.4}$$

where $\pi_t^A = \mathbb{E}_t\left[\pi^A_{\tau\wedge T}\right]$. Combining the previous two equations, we confirm (25).

Finally, we prove $S_\tau = \min\left\{\widehat{S}_\tau^A, \widehat{S}_\tau^B\right\}$ when $\tau < T$. Given the Brownian filtration, it follows from the martingale representation theorem (see e.g. Karatzas and Shreve (1991), Chap. 3, Th. 4.15) that both $\mathbb{E}_t\left[\pi^A_{\tau\wedge T} S_{\tau\wedge T}\right]$ and $\pi^A_t = \mathbb{E}_t[\pi^A_{\tau\wedge T}]$ in (A.4) can be represented as stochastic integrals with respect to the Brownian motion Z_t; in particular, they have continuous paths for $t \in [0, \tau\wedge T]$. Therefore, S in (A.4) is also continuous on the same time interval. Suppose that $S_\tau < \min\{\widehat{S}_\tau^A, \widehat{S}_\tau^B\}$ with positive probability under $\mathbb{P}$. Due to the path continuity of S, there exists $t < \tau\wedge T$, i.e., before the circuit breaker is triggered, such that $S_t < \min\{\widehat{S}_t^A, \widehat{S}_t^B\}$ with positive probability under $\mathbb{P}$. However, since neither agent is constrained before market closure, both would want to take on leverage to invest in the under-valued stock when the stock price is less than the pessimistic agent's valuation. This cannot be an equilibrium. Therefore $S_\tau = \min\{\widehat{S}_\tau^A, \widehat{S}_\tau^B\}$.

A.3 Proof of Proposition 3

Recall from (15), the definition of τ' in (28) is equivalent to:

$$\tau' = \inf\{t \geqslant 0 \,:\, \widehat{S}_t^{min} \leqslant \underline{S}\}. \tag{A.5}$$

We first show that the map F admits a fixed point. To this end, when $S_0 = 0$, $\underline{S} = (1-\alpha)S_0 = 0$, we have $\underline{D} \equiv 0$, hence $\tau' = \infty$ and $F(0) = \left[\omega\mathbb{E}[D_T^{-1}] + (1-\omega)\mathbb{E}^B[D_T^{-1}]\right]^{-1} > 0$. When S_0 is sufficiently large such that $\underline{D}(0, \delta_0) > D_0$, then $\tau' = 0$ a.s. As a result,

$$F(S_0) = \widehat{S}_0^{min} < \underline{D}(0, \delta_0) \cdot \begin{cases} e^{(\mu-\sigma^2)T} & \text{if} \quad \delta_0 \geqslant \underline{\delta}(0) \\ e^{(\mu-\sigma^2)T - a(0,T) - b(0,T)\delta_0}, & \text{if} \quad \delta_0 \leqslant \underline{\delta}(0) \end{cases} = \underline{S} \leqslant S_0.$$

Therefore we have shown that $F(0) > 0$, $F(S_0) < S_0$ for sufficiently large S_0. If F is also continuous, then there exists at least one fixed point S_0 such that $F(S_0) = S_0$. The continuity of F is proved in Online Appendix OA.2.2.

Finally, for a fixed point S_0, introduce S' as in (29), we prove that S' reaches the threshold $\underline{S}$ at τ' for the first time. Define:

$$\sigma = \inf\{t \geqslant 0 \,:\, S'_t \leqslant \underline{S}\}.$$

Given that $S'_{\tau'} = \underline{S}$ when $\tau' < T$, we have $\sigma \leqslant \tau'$. We only need to show that σ cannot be smaller than τ' with positive probability. Assuming otherwise, that is, $\sigma < \tau' < T$ happens with positive probability. Then, $S'_\sigma = \underline{S}$ when $\sigma < \tau' < T$. In this case, (29) implies that:

$$\text{either} \quad \mathbb{E}_\sigma[(\widehat{S}^{min}_{\tau'\wedge T})^{-1}] \geqslant \underline{S}^{-1} \quad \text{or} \quad \mathbb{E}^B_\sigma[(\widehat{S}^{min}_{\tau'\wedge T})^{-1}] \geqslant \underline{S}^{-1}. \tag{A.6}$$

However, $\widehat{S}^{min}_{\tau'\wedge T} \geqslant \underline{S}$ and this inequality is strict with positive probability (this situation happens when $\tau' = T$). Hence, $\mathbb{E}_\sigma[(\widehat{S}^{min}_{\tau'\wedge T})^{-1}] < \underline{S}^{-1}$ and $\mathbb{E}^B_\sigma[(\widehat{S}^{min}_{\tau'\wedge T})^{-1}] < \underline{S}^{-1}$, contradicting with (A.6). Therefore, σ and τ' must be the same, and τ' satisfies (8).

A.4 Proof of Proposition 4

Lemma 1 is proved in Online Appendix OA.2.3. In the proof, we first show that $\underline{\delta}(t) > 0$ for any $t < T$. Therefore, agent B must be the marginal agent if $\tau = 0$, because $\delta_0 = 0 < \underline{\delta}(t)$. When $0 < \tau < T$, using the explicit solution of (5) with $\delta_0 = \bar{\delta} = 0$, we prove that $\delta_\tau < \underline{\delta}(\tau)$. Hence agent B is the marginal agent in this case as well.

We now prove Proposition 4 first for the constant δ case then for the stochastic δ case.

Constant δ. Let us consider the case where δ is a negative constant, i.e., agent B is more pessimistic. It is proved in Online Appendix OA.2.4 that the map F is decreasing, hence there can be only one fixed point for F.

Stochastic δ. We first claim the following result, proved in Online Appendix OA.2.5.

Lemma A.2. *Suppose that $\bar{\delta} = \delta_0 = 0$ and $\omega < 1 - \alpha$. Then, $F'(S_0) < 1$ for* any *fixed point S_0 of F.*

With the previous result, now suppose that F admits more that one fixed point, say, S_0 and $\widetilde{S}_0$. Because $F'(S_0)$ and $F'(\widetilde{S}_0) < 1$, the continuous function $S \mapsto F(S)$ must cross the 45 degree line $S \mapsto S$ at a point S_0^{mid} between S_0 and $\widetilde{S}_0$. The point S_0^{mid} is another fixed point for F, but $F'(S_0^{mid})$ must be at least 1 to cross the 45 degree line. This contradicts with $F'(S_0^{mid}) < 1$. Therefore, there can be only one fixed point for F.

A.5 Proofs for Section 4

PROOF OF PROPOSITION 5

Comparing (16) and (30), the statement is equivalent to $\mathbb{E}_t\big[\mathbb{E}^B_{\tau\wedge T}[D_T^{-1}]\big] > \mathbb{E}_t[D_T^{-1}]$. Due to the tower property of conditional expectation, it suffices to prove:

$$\mathbb{E}^B_{\tau\wedge T}[D_T^{-1}] \geqslant \mathbb{E}_{\tau\wedge T}[D_T^{-1}], \tag{A.7}$$

where the inequality is strict with positive probability. We will show later in the proof of Proposition 6 that stock volatility is strictly positive, hence the circuit breaker is triggered before T with positive probabilities under both $\mathbb{P}$ and $\mathbb{P}^B$. Due to (15), (A.7) is equivalent to $\widehat{S}^B_{\tau\wedge T} \leqslant \widehat{S}^A_{\tau\wedge T}$, which holds for the negative constant δ or the stochastic δ due to Lemma 1. Therefore (A.7) is confirmed.

PROOF OF PROPOSITION 6

Other than the results stated in the proposition, we also have:

$$\lim_{D_t\uparrow\infty} \widehat{S}_t - S_t = 0 \quad \text{and} \quad \lim_{D_t\uparrow\infty} \sigma^S_t = \lim_{D_t\uparrow\infty} \sigma^{\widehat{S}}_t = \sigma, \quad \text{for any } t < T.$$

We will prove these additional results in what follows as well.

Statements in (i). The statements and $\lim_{D_t\uparrow\infty} \widehat{S}_t - S_t = 0$ will be proved in Online Appendix OA.2.6 by examining the limiting behavior of $D_t \uparrow \infty$ and $D_t \downarrow \underline{D}_t$.

To prove the statements in (ii) and (iii), we introduce:

$$y_t \equiv \left(\frac{\sigma}{2} - \frac{\delta}{\sigma}\right)t + Z_t \quad \text{and} \quad d \equiv \frac{1}{\sigma}\log\left[(1-\alpha)S_0 e^{-(\mu-\sigma^2+\delta)T}\right].$$

Then, $D_t = e^{(\mu-\sigma^2+\delta)t+\sigma y_t}$, $\underline{D}_t = e^{(\mu-\sigma^2+\delta)t+\sigma d}$, and $\tau = \inf\{s \geqslant t : y_s \leqslant d\}$ is the circuit breaker triggering time. Denote:

$$f^A(y_t,t) \equiv \mathbb{E}_t\big[\mathbb{E}^B_{\tau\wedge T}[D_T^{-1}]\big], \quad \widehat{f}^A(y_t,t) \equiv \mathbb{E}_t[D_T^{-1}], \quad \text{and} \quad f^B(y_t,t) \equiv \mathbb{E}^B_t[D_T^{-1}]. \tag{A.8}$$

It then follows from (16) and (30) that

$$\widehat{S}_t = \left[\omega^A \widehat{f}^A(y_t,t) + \omega^B f^B(y_t,t)\right]^{-1}, \quad S_t = \left[\omega^A f^A(y_t,t) + \omega^B f^B(y_t,t)\right]^{-1}. \tag{A.9}$$

The functions $\widehat{f}^A$ and f^B have explicit expressions:

$$\widehat{f}^A(y_t,t) = \mathbb{E}_t[D_T^{-1}] = e^{-(\mu-\sigma^2+\delta)T}\mathbb{E}_t\big[e^{-\sigma y_T}\big] = e^{-(\mu-\sigma^2+\delta)T-\sigma y_t+\delta(T-t)}, \tag{A.10}$$

$$f^B(y_t,t) = e^{-(\mu-\sigma^2+\delta)T-\sigma y_t}. \tag{A.11}$$

The function f^A also admits a closed-form expression presented in Proposition OA.1.

Statements in (ii). Because the volatility of y_t is 1, the volatility of $\widehat{S}$ is $\partial_y \ln(\widehat{S}_t)$. It follows from (A.9) that:

$$\begin{aligned}
\sigma_t^{\widehat{S}} = \frac{\partial_y \widehat{S}_t}{\widehat{S}_t} &= -\frac{\partial_y\Big(\omega_t^A \widehat{f}^A(y_t,t) + (1-\omega_t^A) f^B(y_t,t)\Big)}{\omega_t^A \widehat{f}^A(y_t,t) + (1-\omega_t^A) f^B(y_t,t)} \\
&= -\frac{\partial_y \omega_t^A (\widehat{f}^A - f^B)}{\omega_t^A \widehat{f}^A + (1-\omega_t^A) f^B} - \frac{\omega_t^A \widehat{f}^A}{\omega_t^A \widehat{f}^A + (1-\omega_t^A) f^B}\frac{\partial_y \widehat{f}^A}{\widehat{f}^A} - \frac{(1-\omega_t^A) f^B}{\omega_t^A \widehat{f}^A + (1-\omega_t^A) f^B}\frac{\partial_y f^B}{f^B} \\
&= \frac{(1-\omega_t^A)\omega_t^A f^B}{\omega_t^A \widehat{f}^A + (1-\omega_t^A) f^B}\frac{\delta}{\sigma}\Big(e^{\delta(T-t)} - 1\Big) + \sigma, \qquad \text{(A.12)}
\end{aligned}$$

where the third identity follows from: $-\partial_y \omega_t^A = \omega_t^A(1-\omega_t^A)\frac{\delta}{\sigma}, -\frac{\partial_y \widehat{f}^A}{\widehat{f}^A} = -\frac{\partial_y f^B}{f^B} = \sigma$, and $\frac{\widehat{f}^A}{f^B} = e^{\delta(T-t)}$. Note that $\delta(e^{\delta(T-t)} - 1) > 0$ when $t < T$, we have from (A.12) that: $\lim_{y \downarrow d} \sigma_t^{\widehat{S}} > \sigma$, for $t < T$. When $y \uparrow \infty$, $f^B \downarrow 0$. Therefore $\sigma^{\widehat{S}} \to \sigma$.

To prove the statements about σ^S in (ii), we collect several properties of functions $f^A, \widehat{f}^A$, and f^B in the following result, whose proofs are given in Online Appendix OA.2.8.

Lemma A.3. *When $\delta < 0$,*

(i) $f^A(d,t) = f^B(d,t)$ *for any* $t \leqslant T$*;*

(ii) $f^B(y,t) > f^A(y,t)$ *for any* $y > d$ *and* $t < T$, $\lim_{y\uparrow\infty} f^A(y,t) = \lim_{y\uparrow\infty} \widehat{f}^A(y,t)$*;*

(iii) $\frac{\partial_y f^B(y,t)}{f^B(y,t)} \equiv -\sigma$ *and* $\lim_{y\uparrow\infty} \frac{\partial_y f^A(y,t)}{f^A(y,t)} = -\sigma$*;*

(iv) If $\delta \leqslant -\frac{\sigma^2}{2}$*, then* $\partial_y f^B(y,t)|_{y=d} > \partial_y f^A(y,t)|_{y=d} > -\big(\sigma - \frac{2\delta}{\sigma}\big) f^A(y,t)|_{y=d}$ *for any* $t < T$*;*

(v) If $\delta \leqslant -\frac{\sigma^2}{2}$, $\partial_y\big(\frac{\partial_y f^A(y,t)}{f^A(y,t)}\big)\big|_{y=d} > 0$ *for any* $t < T$*.*

Limit of σ^S. Using (30) and following similar derivations in (A.12), we obtain

$$\sigma_t^S = \frac{\partial_y S_t}{S_t} = \sigma - \frac{\partial_y \omega_t^A (f^A - f^B)}{\omega_t^A f^A + (1-\omega_t^A) f^B} + \frac{\omega_t^A f^A}{\omega_t^A f^A + (1-\omega_t^A) f^B}\Big(-\frac{\partial_y f^A}{f^A} - \sigma\Big). \qquad \text{(A.13)}$$

Sending $y_t \downarrow d$ (equivalently $D_t \downarrow \underline{D}_t$), Lemma A.3 (i) implies:

$$\lim_{D_t \downarrow \underline{D}_t} \sigma_t^S = \sigma + \lim_{y_t \downarrow d} \frac{\omega_t^A f^A}{\omega_t^A f + (1-\omega_t^A) f^B}\Big(-\frac{\partial_y f^A}{f^A} - \sigma\Big). \qquad \text{(A.14)}$$

Comparing the previous expression with (A.12), the statement $\lim_{D_t \downarrow \underline{D}_t} \sigma_t^S > \lim_{D_t \downarrow \underline{D}_t} \sigma_t^{\widehat{S}}$ is equivalent to:

$$\lim_{y \downarrow d} \frac{\omega_t^A f^A}{\omega_t^A f^A + (1-\omega_t^A) f^B}\Big(-\frac{\partial_y f^A}{f^A} - \sigma\Big) > \lim_{y \downarrow d} \frac{(1-\omega_t^A)\omega_t^A f^B}{\omega_t^A \widehat{f}^A + (1-\omega_t^A) f^B}\frac{\delta}{\sigma}\Big(e^{\delta(T-t)} - 1\Big). \qquad \text{(A.15)}$$

This is satisfied when $\delta \leqslant -\sigma^2$, proved in Online Appendix OA.2.9.

When $y_t \uparrow \infty$ (equivalently $D_t \uparrow \infty$), $\eta_t \to 1$ and $\omega_t^A \to 1, \partial_y \omega_t^A \to 0$. Therefore $\lim_{y\uparrow\infty} \frac{\partial_y f^A}{f^A} = -\sigma$ in Lemma A.3 (iii) and (A.13) combined imply $\lim_{D_t\uparrow\infty} \sigma_t^S = \sigma$.

Statements in (iii). We first prove the statements in the limit $\omega \to 1$. Then apply a continuity argument to obtain the statements when ω is sufficiently close to 1.

As $\omega \to 1$, we obtain from (OA.20) from the online appendix that $\widehat{S}_t/S_t \to e^{-\delta(T-t)} > 1$ due to $\delta < 0$.

As $\omega \to 1$, $\widehat{S}_t = e^{(\mu-\sigma^2)(T-t)} D_t$, hence $\sigma_t^{\widehat{S}} \equiv \sigma$. We claim that σ_t^S increasingly converges to a limit higher than σ as $D_t \downarrow \underline{D}_t$. To prove this claim, observe that when $\omega \to 1$, $\omega^A \to 1$, equation (A.9) yields $S_t = (f^A(y_t, t))^{-1}$. Then:

$$\lim_{\omega \to 1} \sigma_t^S = -\frac{\partial_y f^A(y_t, t)}{f^A(y_t, t)}.$$

Thanks to Lemma A.3 (v), $\partial_y \big(\partial_y f^A / f^A\big)|_{y=d} > 0$. Therefore, in the limit $\omega \to 1$, σ_t^S increases to its limit as $y_t \downarrow d$. Lemma A.3 (i), (iii), and (iv) imply that:

$$\sigma = -\frac{\partial_y f^B(y,t)|_{y=d}}{f^B(d,t)} < -\frac{\partial_y f^A(y,t)|_{y=d}}{f^A(d,t)} = \lim_{\omega \to 1} \sigma_t^S.$$

Finally, we prove the statements when ω is close to 1 using a continuity argument in Online Appendix OA.2.10.

Proof of Proposition 7 Denote $\Delta S_t = S_t - \underline{S}$. Then $h = (\Delta S_t)^2$ from the choice of h. We present the following results whose proofs are given in Online Appendix OA.2.11.

Lemma A.4. *When $\delta \leqslant 0$ and $S_t > \underline{S}$, we have the following statements for any $t < T$.*

$$\lim_{S_t \downarrow \underline{S}} p_{(\Delta S_t)^2}(S_t) = 2N\left(-\frac{1}{\underline{S}\sigma_t^S|_{S_t=\underline{S}}}\right), \tag{A.16}$$

where $N[\cdot]$ is the distribution function of standard normal. For the complete market, the same statement holds with the volatility replaced by $\sigma_t^{\widehat{S}}|_{\widehat{S}_t=\underline{S}}$ on the right-hand side. Moreover,

$$\partial_{\Delta S_t} p_{(\Delta S_t)^2}(S_t)\big|_{S_t=\underline{S}} = -\frac{\sigma - \frac{2\delta}{\sigma}}{\underline{S}\sigma_t^S|_{S_t=\underline{S}}} N\left(-\frac{1}{\underline{S}\sigma_t^S|_{S_t=\underline{S}}}\right). \tag{A.17}$$

In the limit $\omega \to 1$, $\partial_{\Delta S_t} \hat{p}_{(\Delta S_t)^2}(S_t)\big|_{S_t=\underline{S}} = -\frac{1}{\underline{S}} N\left(-\frac{1}{\sigma \underline{S}}\right)$.

Because $\delta \leqslant 0$, (A.17) shows that $\partial_{\Delta S_t} p_{(\Delta S_t)^2}(S_t)\big|_{S_t=\underline{S}} < 0$, the statement in (i) immediately follows. For the statement in (ii), equation (A.16) shows that the limit of hitting probability is increasing in the volatility at the circuit breaker $\underline{S}$. We have seen from Proposition 6 (ii) that volatility at the threshold is higher with circuit breaker than without. Therefore $\lim_{S_t \downarrow \underline{S}} p_{(\Delta S_t)^2}(S_t) > \lim_{S_t \downarrow \underline{S}} \hat{p}_{(\Delta S_t)^2}(S_t)$. Then the continuity of hitting probabilities in S_t implies that the previous inequality still holds in the pre-limit in a neighborhood of $\underline{S}$. When ω is sufficiently close to 1, the statement is proved in Online Appendix OA.2.12.

Online Appendix

The Dark Side of Circuit Breakers

Hui Chen   Anton Petukhov   Jiang Wang   Hao Xing

February 27, 2023

This online appendix provides supplemental materials for the paper. Section OA.1 provides additional institutional details on the circuit breakers in the U.S. Section OA.2 supplies the details of the proofs. Section OA.3 presents the quantitative analysis on the impact of circuit breaker for the constant disagreement case, and Section OA.4 examines how a circuit-breaker's market impact varies during the trading session. Section OA.5 compares the impact of pre-scheduled trading halts with that of circuit breakers. Section OA.6 contains more robustness results on the empirical analysis.

OA.1 A Brief History of the MWCB Mechanism in the U.S.

- **October 1988:** In response to the "Black Monday" of October 19, 1987, when the DJIA declined 22.6%, the Presidential Task Force ("Brady Commission") proposed the first circuit breaker rule for stocks, stock options, and stock index futures, which was implemented in 1988 (Securities Exchange Act Release No. 26198). There are two trigger levels:
  - Level 1: 250 points decline in DJIA from the previous day's close; the markets will halt for one hour.
  - Level 2: 400 points decline in DJIA from the previous day's close; the markets will halt for two hours.
- **July 1996:** The durations of the 250 and 400 point halts were reduced to 30 minutes and 60 minutes from one hour and two hours, respectively (Securities Exchange Act Release No. 37457).
- **January 1997:** The two trigger values were increased to 350 and 550 points (Securities Exchange Act Release No. 38221).
- **October 1997:** On October 27, 1997, at 2:36 p.m., a 350-point (4.54%) decline in the DJIA led to a 30-minute trading halt on stocks, equity options, and index futures. After trading resumed at 3:06 p.m., prices fell rapidly to reach the second-level 550-point circuit breaker at 3:30 p.m., leading to the early market closure for the day. For a detailed review of this event, see Securities and Exchange Commission (1998).

- **January 1998:** To address the concerns rising from the triggering of circuit breakers on October 27, 1997, the exchanges adopted the following interim changes:
  - Level 1: 350 points decline in DJIA; results in a halt of 30 minutes (if triggered at or before 3:00 p.m.), or no halt (after 3:00 p.m.).
  - Level 2: 550 points decline in DJIA; results in a halt of 60 minutes (if triggered at or before 2:00 p.m.), 30 minutes (if triggered between 2:00 p.m. and 3:00 p.m.), or market closure for the day (if reached at or after 3:00 p.m.).
- **April 1998:** The MWCB trigger levels were changed from declines in DJIA of specific number of points from previous day's close to percentage declines. The reference values for the percentage declines were calculated using average closing values of the DJIA for the last month of a quarter and updated at the beginning of each quarter (Securities Exchange Act Release No. 39846). There were three trigger levels instead of two:
  - Level 1: 10% decline in DJIA; results in a halt of 1 hour (if triggered before 2:00 p.m.), 30 minutes (if triggered between 2:00 p.m. and 2:30 p.m.), or no halt (after 2:30 p.m.).
  - Level 2: 20% decline in DJIA; results in a halt of 2 hours (if triggered before 1:00 p.m.), 1 hour (if triggered between 1:00 p.m. and 2:00 p.m.), or market closure for the day (after 2:00 p.m.).
  - Level 3: 30% decline in DJIA; results in market closure for the day.
- **April 2013:** The exchanges replaced the DJIA with the SPX for a broader representation of the market, and changed the updating frequency of the percentage thresholds from quarterly to daily (Securities Exchange Act Release No. 67090).
  - Level 1: 7% decline in SPX; results in a halt of 15 minutes (if triggered before 3:25 p.m.) or no halt (after 3:25 p.m.).
  - Level 2: 13% decline in SPX; results in a halt of 15 minutes (if triggered before 3:25 p.m.) or no halt (after 3:25 p.m.).
  - Level 3: 20% decline in SPX; results in market closure for the day.

  This version of the MWCB was first implemented on April 8, 2013 and has remained substantively unchanged since then.
- **March 2020:** The circuit breaker was triggered four times during the month as the COVID-19 pandemic started hitting the U.S., as shown in Figure OA.1.

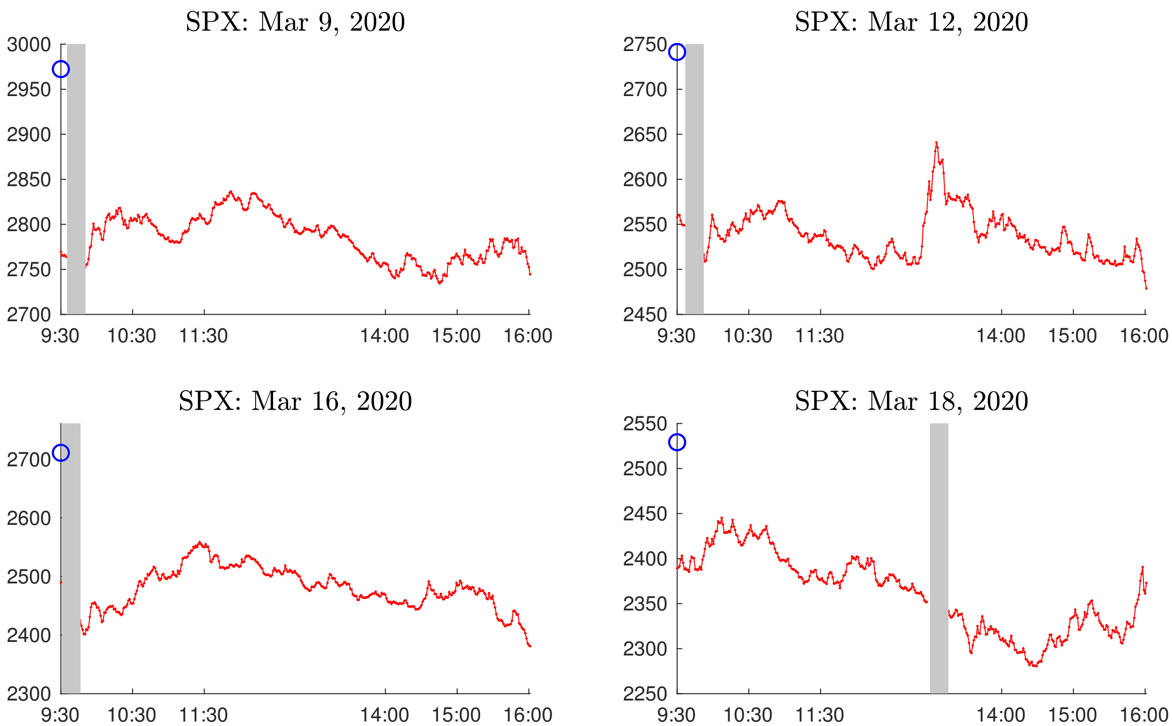

Figure OA.1: Four market-wide circuit breaker triggered days in March 2020. Level 1 circuit breaker is triggered after a 7% drop in SPX from the previous day's close (blue circle). Level 1 circuit breaker was triggered at 9:35 a.m, 9:36 a.m, 9:30 a.m, and 12:57 p.m on March 9, 12, 16, 18 of 2020, respectively. The market is then paused for 15 minutes before its reopen, which is marked by the gray bar in each panel.

OA.2 Additional Proofs

OA.2.1 Proof of Proposition 1

Starting from the planner's problem (10), we obtain the following result:

Lemma OA.5. *Without the circuit breaker, the equilibrium allocation between the two agents is given by:*

$$\widehat{W}_T^A = \frac{\omega}{\omega + (1-\omega)\eta_T} D_T, \quad \widehat{W}_T^B = \frac{(1-\omega)\eta_T}{\omega + (1-\omega)\eta_T} D_T. \tag{OA.1}$$

Clearly, in equilibrium, agent B will obtain a bigger share of the aggregate dividend when realized value of the Radon-Nikodym derivative η_T is higher, i.e., under those paths that agent B considers to be more likely.

Given the equilibrium allocation, we can further obtain the state price density under agent A's belief, which corresponds the objective probability measure $\mathbb{P}$:

$$\widehat{\pi}_t^A = \mathbb{E}_t\left[\xi u'(\widehat{W}_T^A)\right] = \mathbb{E}_t\left[\xi(\widehat{W}_T^A)^{-1}\right], \quad 0 \leqslant t \leqslant T \tag{OA.2}$$

for some constant ξ. Using the state price density, we obtain the stock price:

$$\widehat{S}_t = \mathbb{E}_t\left[\frac{\widehat{\pi}_T^A D_T}{\mathbb{E}_t\left[\widehat{\pi}_T^A\right]}\right] = \frac{\mathbb{E}_t\left[\omega + (1-\omega)\,\eta_T\right]}{\mathbb{E}_t\left[D_T^{-1}\left(\omega + (1-\omega)\,\eta_T\right)\right]} = \frac{\omega + (1-\omega)\,\eta_t}{\omega\mathbb{E}_t\left[D_T^{-1}\right] + (1-\omega)\,\mathbb{E}_t\left[D_T^{-1}\eta_T\right]}, \quad \text{(OA.3)}$$

where

$$\mathbb{E}_t\left[D_T^{-1}\right] = D_t^{-1}e^{-\left(\mu-\sigma^2\right)(T-t)}, \quad \mathbb{E}_t\left[D_T^{-1}\eta_T\right] = \eta_t\mathbb{E}_t\left[D_T^{-1}\frac{\eta_T}{\eta_t}\right] = \eta_t\mathbb{E}_t^B\left[D_T^{-1}\right]. \quad \text{(OA.4)}$$

Our model fits into the affine disagreement framework of Chen, Joslin, and Tran (2010). Define the log dividend $x_t = \log D_t$. Under measure $\mathbb{P}^B$, the processes for x_t and δ_t are:

$$dx_t = \left(\mu - \frac{\sigma^2}{2} + \delta_t\right)dt + \sigma dZ_t^B, \quad d\delta_t = \left[\kappa\bar{\delta} + \left(\frac{\nu}{\sigma} - \kappa\right)\delta_t\right]dt + \nu dZ_t^B. \quad \text{(OA.5)}$$

Define $X_t = [x_t\ \delta_t]'$, then X_t follows an affine process,

$$dX_t = \left(K_0 + K_1 X_t\right)dt + \sigma_X dZ_t^B, \quad \text{(OA.6)}$$

with

$$K_0 = \begin{bmatrix}\mu - \frac{\sigma^2}{2}\\ \kappa\bar{\delta}\end{bmatrix},\ K_1 = \begin{bmatrix}0 & 1\\ 0 & \frac{\nu}{\sigma} - \kappa\end{bmatrix},\ \sigma_X = \begin{bmatrix}\sigma\\ \nu\end{bmatrix}. \quad \text{(OA.7)}$$

We are interested in computing:

$$g\left(t, X_t\right) = \mathbb{E}_t^B\left[e^{\rho_1' X_T}\right], \quad \text{with } \rho_1 = [-1\ \ 0]'. \quad \text{(OA.8)}$$

By applying standard results for the conditional moment-generating functions of affine processes, we get:

$$g\left(t, X_t\right) = \exp\left(A\left(t,T\right) + B\left(t,T\right)' X_t\right), \quad \text{(OA.9)}$$

where

$$0 = \dot{B} + K_1' B, \quad B\left(T,T\right) = \rho_1 \quad \text{(OA.10a)}$$

$$0 = \dot{A} + B' K_0 + \frac{1}{2}tr\left(BB'\sigma_X\sigma_X'\right), \quad A\left(T,T\right) = 0. \quad \text{(OA.10b)}$$

Solving for the ODEs gives:

$$B(t,T) = \begin{bmatrix}-1 & \dfrac{1 - e^{\left(\frac{\nu}{\sigma}-\kappa\right)(T-t)}}{\frac{\nu}{\sigma} - \kappa}\end{bmatrix}', \quad \text{(OA.11)}$$

and

$$A(t,T) = \left[\mu - \sigma^2 - \frac{\kappa\bar{\delta} - \sigma\nu}{\frac{\nu}{\sigma} - \kappa} - \frac{\nu^2}{2\left(\frac{\nu}{\sigma} - \kappa\right)^2}\right](t-T) - \frac{\nu^2}{4\left(\frac{\nu}{\sigma} - \kappa\right)^3}\left[1 - e^{2\left(\frac{\nu}{\sigma}-\kappa\right)(T-t)}\right]$$
$$+ \left[\frac{\kappa\bar{\delta} - \sigma\nu}{\left(\frac{\nu}{\sigma} - \kappa\right)^2} + \frac{\nu^2}{\left(\frac{\nu}{\sigma} - \kappa\right)^3}\right]\left[1 - e^{\left(\frac{\nu}{\sigma}-\kappa\right)(T-t)}\right]. \quad \text{(OA.12)}$$

After plugging the above results back into (OA.3) and reorganizing the terms, we get:

$$S_t = \frac{\omega + (1-\omega)\eta_t}{\omega + (1-\omega)\eta_t H(t,\delta_t)} D_t e^{\left(\mu-\sigma^2\right)(T-t)}, \tag{OA.13}$$

where

$$H(t,\delta_t) = e^{a(t,T)+b(t,T)\delta_t}, \tag{OA.14a}$$

$$\begin{aligned} a(t,T) = & \left[\frac{\kappa\bar{\delta}-\sigma\nu}{\frac{\nu}{\sigma}-\kappa} + \frac{\nu^2}{2\left(\frac{\nu}{\sigma}-\kappa\right)^2}\right](T-t) - \frac{\nu^2}{4\left(\frac{\nu}{\sigma}-\kappa\right)^3}\left[1-e^{2\left(\frac{\nu}{\sigma}-\kappa\right)(T-t)}\right] \\ & + \left[\frac{\kappa\bar{\delta}-\sigma\nu}{\left(\frac{\nu}{\sigma}-\kappa\right)^2} + \frac{\nu^2}{\left(\frac{\nu}{\sigma}-\kappa\right)^3}\right]\left[1-e^{\left(\frac{\nu}{\sigma}-\kappa\right)(T-t)}\right], \end{aligned} \tag{OA.14b}$$

$$b(t,T) = \frac{1-e^{\left(\frac{\nu}{\sigma}-\kappa\right)(T-t)}}{\frac{\nu}{\sigma}-\kappa}. \tag{OA.14c}$$

Finally, to compute the conditional volatility of stock returns, we have:

$$\begin{aligned} d\widehat{S}_t &= \widehat{\mu}_{S,t}\widehat{S}_t dt + \widehat{\sigma}_{S,t}\widehat{S} dZ_t \\ &= o(dt) + \widehat{S}_t \frac{dD_t}{D_t} + \eta_t D_t e^{\left(\mu-\sigma^2\right)(T-t)} \frac{\omega(1-\omega)[1-H(t,\delta_t)]}{[\omega+(1-\omega)\eta_t H(t,\delta_t)]^2}\frac{d\eta_t}{\eta_t} \\ &\quad - D_t e^{\left(\mu-\sigma^2\right)(T-t)} \frac{[\omega+(1-\omega)\eta_t](1-\omega)\eta_t H(t,\delta_t) b(t,T)}{[\omega+(1-\omega)\eta_t H(t,\delta_t)]^2} d\delta_t. \end{aligned}$$

After collecting the diffusion terms, we get:

$$\begin{aligned} \widehat{\sigma}_{S,t} = \sigma + \frac{D_t e^{\left(\mu-\sigma^2\right)(T-t)}}{\widehat{S}_t} & \left\{ \frac{\omega(1-\omega)\left[1-H(t,\delta_t)\right]}{[\omega+(1-\omega)\eta_t H(t,\delta_t)]^2}\frac{\delta_t\eta_t}{\sigma} \right. \\ & \left. - \frac{[\omega+(1-\omega)\eta_t]\,(1-\omega)\eta_t b(t,T) H(t,\delta_t)}{[\omega+(1-\omega)\eta_t H(t,\delta_t)]^2}\nu \right\}. \end{aligned} \tag{OA.15}$$

OA.2.2 The map F is continuous

We first consider the case of constant δ. In this case, explicit expressions of $\mathbb{E}\big[(\widehat{S}^{min}_{\tau'\wedge T})^{-1}\big]$ and $\mathbb{E}^B\big[(\widehat{S}^{min}_{\tau'\wedge T})^{-1}\big]$ are given in Proposition OA.1 later, which show the continuity of these two expectations with respect to $\underline{S}$. For stochastic δ, we prove the continuity of $\mathbb{E}\big[(\widehat{S}^{min}_{\tau'\wedge T})^{-1}\big]$ with respect to $\underline{S}$ in what follows, the proof for the continuity of $\mathbb{E}^B\big[(\widehat{S}^{min}_{\tau'\wedge T})^{-1}\big]$ is similar.

Introduce two auxiliary variables:

$$X^A_t = \ln \widehat{S}^A_t - \ln \underline{S} \quad \text{and} \quad X^B_t = \ln \widehat{S}^B_t - \ln \underline{S}.$$

In terms of these two auxiliary variables, τ' can be equivalently transformed into:

$$\tau^{x^A,x^B} = \inf\{t \geqslant 0 \,:\, X^A_t \leqslant 0\} \wedge \inf\{t \geqslant 0 \,:\, X^B_t \leqslant 0\},$$

where $X^A_0 = x^A$ and $X^B_0 = x^B$. Define

$$v(x^A, x^B) = \mathbb{E}\big[(\widehat{S}^{min}_{\tau^{x^A,x^B}\wedge T})^{-1} \,|\, X^A_0 = x^A, X^B_0 = x^B\big].$$

It suffices to prove that the function v is continuous in (x^A, x^B). It follows from Darling and Pardoux (1997) that $v(x^A, x^B)$ can be represented as the initial value for the solution of a backward stochastic differential equation with the random horizon $\tau^{x^A,x^B} \wedge T$. Suppose that the set $\Gamma = \{x^A = 0 \text{ or } x^B = 0 : \mathbb{P}(\tau^{x^A,x^B} > 0) = 0\}$ is closed, (Darling and Pardoux, 1997, Prop. 6.3) shows that $(x^A, x^B) \mapsto \tau^{x^A,x^B}$ is continuous a.s.. Then the continuity of v in (x^A, x^B) follows from the previous continuity of the stopping time and a dominated convergence theorem. Because the volatility of (X^A, X^B) is nondegenerate, the assumption that Γ is closed is ensured by (Bayrakter, Song, and Yang, 2011, Prop. 3.1).

OA.2.3 Proof of Lemma 1

We first prove that $a(t,T) > 0$ and $b(t,T) < 0$ for $t < T$. To this end, it follows from the form of $b(t,T)$ in (12b) that $b(t,T) < 0$, no matter whether $\frac{\nu}{\sigma} - \kappa > 0$ or $\frac{\nu}{\sigma} - \kappa < 0$. For $a(t,T)$, plug $\bar{\delta} = 0$ into (12a), we obtain:

$$\begin{aligned} a(t,T) = &\frac{-\nu^2/2 + \sigma\nu\kappa}{(\nu/\sigma - \kappa)^2}(T-t) + \frac{\sigma\nu\kappa}{(\nu/\sigma-\kappa)^3}\big[1 - e^{(\nu/\sigma-\kappa)(T-t)}\big] \\ &- \frac{1}{4}\frac{\nu^2}{(\nu/\sigma-\kappa)^3}\big[1 - e^{2(\nu/\sigma-\kappa)(T-t)}\big]. \end{aligned}$$

Define:

$$h(t) = \frac{-\nu^2/2 + \sigma\nu\kappa}{(\nu/\sigma - \kappa)^2}t + \frac{\sigma\nu\kappa}{(\nu/\sigma-\kappa)^3}\big[1 - e^{(\nu/\sigma-\kappa)t}\big] - \frac{1}{4}\frac{\nu^2}{(\nu/\sigma-\kappa)^3}\big[1 - e^{2(\nu/\sigma-\kappa)t}\big].$$

Then $a(t,T) = h(T-t)$. To show that $a(t,T) > 0$, it suffices to show $h(0) = 0$ and $h'(t) > 0$ for any $t > 0$. The first claim $h(0) = 0$ is clear from the definition of h. For the second claim,

$$\begin{aligned} h'(t) &= \frac{-\nu^2/2 + \sigma\nu\kappa}{(\nu/\sigma - \kappa)^2} - \frac{\sigma\nu\kappa}{(\nu/\sigma-\kappa)^2}e^{(\nu/\sigma-\kappa)t} + \frac{\nu^2}{2(\nu/\sigma-\kappa)^2}e^{2(\nu/\sigma-\kappa)t} \\ &= \frac{1}{2(\nu/\sigma-\kappa)^2}\Big(\sigma\kappa - \nu e^{(\nu/\sigma-\kappa)t}\Big)^2 - \frac{1}{2(\nu/\sigma-\kappa)^2}(\sigma\kappa - \nu)^2. \end{aligned}$$

When $\nu/\sigma - \kappa > 0$ and $t > 0$, $\nu e^{(\nu/\sigma-\kappa)t} - \sigma\kappa > \nu - \sigma\kappa > 0$, then $h'(t) > 0$. When $\nu/\sigma - \kappa < 0$, $\sigma\kappa - \nu e^{(\nu/\sigma-\kappa)t} > \sigma\kappa - \nu > 0$, then $h'(t) > 0$ as well. Therefore $h'(t) > 0$ always holds for $t > 0$. Consequently, $a(t,T) > 0$ for any $t < T$.

Thanks to $a(t,T) > 0$ and $b(t,T) < 0$, it follows $\underline{\delta}(t) = -\frac{a(t,T)}{b(t,T)} > 0$. When $\tau = 0$, $\delta_0 < \underline{\delta}(0)$ is satisfied because $\delta_0 = 0$. Therefore agent B is the marginal agent. Next, we prove $\delta_\tau < \underline{\delta}(\tau)$ when $0 < \tau < T$. To this end, because $a(t,T) + b(t,T)\delta_t \geqslant 0$ if and only if $\delta_t \leqslant \underline{\delta}(t)$, we have

$$\begin{aligned} \underline{D}(t, \delta_t) &= \begin{cases} \underline{S}e^{-(\mu-\sigma^2)(T-t)}, & \delta_t > \underline{\delta}(t) \\ \underline{S}e^{-(\mu-\sigma^2)(T-t)+a(t,T)+b(t,T)\delta_t}, & \delta_t \leqslant \underline{\delta}(t). \end{cases} \\ &= \max\Big\{\underline{S}e^{-(\mu-\sigma^2)(T-t)}, \underline{S}e^{-(\mu-\sigma^2)(T-t)+a(t,T)+b(t,T)\delta_t}\Big\}. \end{aligned} \tag{OA.16}$$

With $\delta_0 = \bar{\delta} = 0$, δ, following $d\delta_t = -\kappa\delta_t dt + \nu dZ_t$, admits an explicit solution:

$$\delta_t = \int_0^t e^{\kappa(s-t)}\nu dZ_s.$$

Suppose that $\delta_\tau \geqslant \underline{\delta}(\tau)$ when $0 < \tau < T$, we will derive a contradiction. Because τ is the first time such that $D_t \leqslant \underline{D}(t, \delta_t)$, we have:

$$\begin{aligned} D_t &> \underline{S}e^{-(\mu-\sigma^2)(T-t)}, \quad t < \tau, \\ D_\tau &= \underline{S}e^{-(\mu-\sigma^2)(T-\tau)}. \end{aligned} \tag{OA.17}$$

where the inequality follows from (OA.16) and the equality is due to $\delta_\tau \geqslant \underline{\delta}(\tau)$. Using the expression that $D_t = D_0 e^{(\mu-\frac{1}{2}\sigma^2)t+\sigma Z_t}$, we obtain an equivalent form of (OA.17):

$$\begin{aligned} Z_t - \frac{1}{\sigma}\log\frac{\underline{S}e^{-(\mu-\sigma^2)T}}{D_0} + \frac{1}{2}\sigma t &> 0, \quad t < \tau, \\ Z_\tau - \frac{1}{\sigma}\log\frac{\underline{S}e^{-(\mu-\sigma^2)T}}{D_0} + \frac{1}{2}\sigma\tau &= 0. \end{aligned} \tag{OA.18}$$

Now,

$$\begin{aligned} \delta_\tau &= \int_0^\tau e^{\kappa(s-\tau)}\nu dZ_s = \nu\Big(Z_\tau - \int_0^\tau Z_s \kappa e^{\kappa(s-\tau)}ds\Big) \\ &= \nu\Big[Z_\tau - \int_0^\tau \Big(Z_s - \frac{1}{\sigma}\log\frac{\underline{S}e^{-(\mu-\sigma^2)T}}{D_0} + \frac{1}{2}\sigma s\Big)\kappa e^{\kappa(s-\tau)}ds \\ &\quad + \int_0^\tau \Big(-\frac{1}{\sigma}\log\frac{\underline{S}e^{-(\mu-\sigma^2)T}}{D_0} + \frac{1}{2}\sigma s\Big)\kappa e^{\kappa(s-\tau)}ds\Big] \\ &< \nu\Big[Z_\tau + \int_0^\tau \Big(-\frac{1}{\sigma}\log\frac{\underline{S}e^{-(\mu-\sigma^2)T}}{D_0} + \frac{1}{2}\sigma s\Big)\kappa e^{\kappa(s-\tau)}ds\Big] \\ &= \nu\Big[Z_\tau + \Big(-\frac{1}{\sigma}\log\frac{\underline{S}e^{-(\mu-\sigma^2)T}}{D_0} + \frac{1}{2}\sigma\tau\Big)(1-e^{-\kappa\tau})\Big] \\ &< \nu\Big[Z_\tau - \frac{1}{\sigma}\log\frac{\underline{S}e^{-(\mu-\sigma^2)T}}{D_0} + \frac{1}{2}\sigma\tau\Big] \\ &= 0. \end{aligned}$$

Here, the second equality follows from the integration by part, the first inequality follows from the first inequality in (OA.18). Because $\tau > 0$ and $\delta_0 = 0$, then $D_0 > \underline{D}(0,0) \geqslant \underline{S}e^{-(\mu-\sigma^2)T}$. Therefore, $\log\frac{\underline{S}e^{-(\mu-\sigma^2)T}}{D_0} < 0$, hence the second inequality above holds. Finally, the second equation in (OA.18) yields the last equality above. However, $\delta_\tau < 0$ contradicts with the assumption $\delta_\tau \geqslant \underline{\delta}(\tau) > 0$ above. Therefore, $\delta_\tau < \underline{\delta}(\tau)$ when $0 < \tau < T$, i.e., agent B is the marginal agent at τ.

OA.2.4 The map F is decreasing in the constant disagreement case

Because agent B is more pessimistic, $\widehat{S}^{min} = S^B$ and

$$(S_t^B)^{-1} = D_0 e^{(\mu-\sigma^2+\delta)T} e^{\delta t} e^{-\frac{\sigma^2}{2}t-\sigma Z_t} = D_0 e^{(\mu-\sigma^2+\delta)T} e^{-\frac{\sigma^2}{2}t-\sigma Z_t^B}.$$

Hence, $\delta < 0$ implies that $(S^B)^{-1}$ is a super-martingale under $\mathbb{P}$ and a martingale under $\mathbb{P}^B$.

For any given $\widetilde{S}_0 > S_0$, we have $(1-\alpha)\widetilde{S}_0 = \underset{\sim}{\widetilde{S}} > \underset{\sim}{S} = (1-\alpha)S_0$. Let:

$$\widetilde{\tau}' = \inf\{t \geqslant 0 \,:\, \widehat{S}_t^{min} \leqslant \underset{\sim}{\widetilde{S}}\}.$$

Then, $\widetilde{\tau}' \leqslant \tau'$ a.s.. It then follows from the optional sampling theorem that:

$$\mathbb{E}\big[(S^B_{\widetilde{\tau}'\wedge T})^{-1}\big] > \mathbb{E}\big[(S^B_{\tau'\wedge T})^{-1}\big] \quad \text{and} \quad \mathbb{E}^B\big[(S^B_{\widetilde{\tau}'\wedge T})^{-1}\big] = \mathbb{E}^B\big[(S^B_{\tau'\wedge T})^{-1}\big].$$

Therefore,

$$\begin{aligned} F(\widetilde{S}_0) &= \Big(\omega\mathbb{E}\big[(S^B_{\widetilde{\tau}'\wedge T})^{-1}\big] + (1-\omega)\mathbb{E}^B\big[(S^B_{\widetilde{\tau}'\wedge T})^{-1}\big]\Big)^{-1} \\ &< \Big(\omega\mathbb{E}\big[(S^B_{\tau'\wedge T})^{-1}\big] + (1-\omega)\mathbb{E}^B\big[(S^B_{\tau'\wedge T})^{-1}\big]\Big)^{-1} = F(S_0), \end{aligned}$$

implying that F is decreasing.

OA.2.5 Proof of Lemma A.2

We have seen that B is the marginal agent when the conditions of Lemma 1 are satisfied. Thus, $\widehat{S}^{min}_{\tau'} = \widehat{S}^B_{\tau'}$ when $\tau' < T$ and $\widehat{S}^{min}_T = D_T^{-1} = \widehat{S}^B_T$ when $\tau' > T$. (29) then implies:

$$F(S_0) = \Big(\omega\mathbb{E}\big[(\widehat{S}^B_{\tau'\wedge T})^{-1}\big] + (1-\omega)\mathbb{E}^B\big[(\widehat{S}^B_{\tau'\wedge T})^{-1}\big]\Big)^{-1}.$$

Recall from (15) that $(\widehat{S}^B_t)^{-1} = \mathbb{E}^B_t[D_T^{-1}]$. It then follows from the tower property of conditional expectation that $\mathbb{E}^B\big[(\widehat{S}^B_{\tau'\wedge T})^{-1}\big] = \mathbb{E}^B\big[\mathbb{E}^B_{\tau'\wedge T}[D_T^{-1}]\big] = \mathbb{E}^B[D_T^{-1}]$, which is independent of the choice of S_0. Therefore, for any fixed point S_0 for F,

$$\frac{d}{dS_0}F(S_0) = -\big(F(S_0)\big)^2\omega\frac{d}{dS_0}\mathbb{E}\big[(\widehat{S}^B_{\tau'\wedge T})^{-1}\big] = -S_0^2\omega(1-\alpha)\frac{d}{d\underset{\sim}{S}}\mathbb{E}\big[(\widehat{S}^B_{\tau'\wedge T})^{-1}\big].$$

Suppose that

$$\frac{d}{d\underset{\sim}{S}}\mathbb{E}\big[(\widehat{S}^B_{\tau'\wedge T})^{-1}\big] \geqslant -\underset{\sim}{S}^{-2}, \tag{OA.19}$$

then we would confirm our claim:

$$\frac{d}{dS_0}F(S_0) \leqslant S_0^2\omega(1-\alpha)\underset{\sim}{S}^{-2} = \frac{\omega}{1-\alpha} < 1,$$

where the last inequality follows from the assumption $\omega < 1-\alpha$.

It then remains to prove (OA.19). To this end, observe that:

$$\mathbb{E}\big[(\widehat{S}^B_{\tau'\wedge T})^{-1}\big] = \underset{\sim}{S}^{-1}\mathbb{P}(\tau' < T) + \mathbb{E}\big[D_T^{-1}1_{\{\tau'\geqslant T\}}\big] = \underset{\sim}{S}^{-1} - \mathbb{E}\Big[\big(\underset{\sim}{S}^{-1} - D_T^{-1}\big)1_{\{\tau'\geqslant T\}}\Big].$$

For the second term on the right hand side, when the threshold $\underset{\sim}{S}$ increases, the set $\{\tau' \geqslant T\}$

becomes smaller. For each path of D in this set, $\underline{S}^{-1} - D_T^{-1}$ is positive but becomes smaller because $\underline{S}^{-1}$ reduces as $\underline{S}$ increases. Therefore, $\frac{d}{d\underline{S}}\mathbb{E}\Big[\big(\underline{S}^{-1} - D_T^{-1}\big)1_{\{\tau' \geqslant T\}}\Big] \leqslant 0$ and (OA.19) is confirmed.

OA.2.6 Proof of Proposition 6 (i)

As $D_t \to \infty$, both $\mathbb{P}(\tau < T)$ and $\mathbb{P}^B(\tau < T) \to 0$. Then $\mathbb{E}_{\tau\wedge T}[D_T^{-1}] \to \mathbb{E}^B_{\tau\wedge T}[D_T^{-1}]$, hence $\lim_{D_t \uparrow\infty} S_t - \widehat{S}_t = 0$.

As $D_t \to \underline{D}_t$, $\tau \to t$ and $S_t \to (1-\alpha)S_0$, due to positive stock volatility. Without the circuit breaker, as $D_t \to \underline{D}_t$,

$$\lim_{D_t \to \underline{D}_t} \eta_t = \lim_{D_t \to \underline{D}_t} e^{\left(-\frac{\delta^2}{2\sigma^2} - \frac{\delta\mu}{\sigma^2} + \frac{\delta}{2}\right)t} D_t^{\frac{\delta}{\sigma^2}} = e^{\left(-\frac{\delta^2}{2\sigma^2} - \frac{\delta\mu}{\sigma^2} + \frac{\delta}{2}\right)t} \underline{D}_t^{\frac{\delta}{\sigma^2}},$$

where $\underline{D}_t = (1-\alpha)S_0 e^{-(\mu-\sigma^2+\delta)(T-t)}$. Therefore,

$$\begin{aligned}\widehat{S}_t \to & \frac{1}{\underline{\omega}_t^A e^{-(\mu-\sigma^2)(T-t)}\underline{D}_t^{-1} + (1-\underline{\omega}_t^A)e^{-(\mu-\sigma^2+\delta)(T-t)}\underline{D}_t^{-1}} \\ &= \frac{(1-\alpha)S_0}{\underline{\omega}_t^A\, e^{\delta(T-t)} + (1-\underline{\omega}_t^A)} > (1-\alpha)S_0, \quad \text{as } D_t \downarrow \underline{D}_t,\end{aligned} \tag{OA.20}$$

where $\underline{\omega}_t^A = \omega\big/\omega + (1-\omega)\, e^{\left[-\delta^2/(2\sigma^2) - \delta\mu/(\sigma^2) + \delta/2\right] t}\underline{D}_t^{\delta/\sigma^2}$.

OA.2.7 Stock price in the constant disagreement case

We first present a closed form expression for the stock price in the case of constant disagreement, $\delta_t = \delta$. Without loss of generality, we focus on the case where agent B is relatively more pessimistic, $\delta \leqslant 0$. The results are summarized below, whose proof utilizes the explicit expressions for hitting time distribution of a Brownian motion with drift.

Proposition OA.1. *Take S_0 as given, With $\delta \leqslant 0$, the stock price at time $t \leqslant \tau \wedge T$ is:*

$$S_t = \left(\omega_t^A \mathbb{E}_t[S_{\tau\wedge T}^{-1}] + \omega_t^B \mathbb{E}_t^B[D_T^{-1}]\right)^{-1}, \tag{OA.21}$$

where

$$\begin{aligned}\mathbb{E}_t[S_{\tau\wedge T}^{-1}] &= f^A(y_t, t) \\ &= \frac{1}{(1-\alpha)S_0}\left\{N\left[\frac{d_t - \left(\frac{\sigma}{2} - \frac{\delta}{\sigma}\right)(T-t)}{\sqrt{T-t}}\right] + e^{(\sigma - \frac{2\delta}{\sigma})d_t} N\left[\frac{d_t + (\frac{\sigma}{2} - \frac{\delta}{\sigma})(T-t)}{\sqrt{T-t}}\right]\right\} \\ &\quad + D_t^{-1}e^{-(\mu-\sigma^2)(T-t)}\left\{N\left[-\frac{d_t + (\frac{\delta}{\sigma} + \frac{\sigma}{2})(T-t)}{\sqrt{T-t}}\right]\right. \\ &\quad \left. - e^{-(\sigma + \frac{2\delta}{\sigma})d_t} N\left[\frac{d_t - (\frac{\delta}{\sigma} + \frac{\sigma}{2})(T-t)}{\sqrt{T-t}}\right]\right\},\end{aligned} \tag{OA.22}$$

and

$$d_t = \frac{1}{\sigma}\left[\log\left(\frac{(1-\alpha)S_0}{D_t}\right) - (\mu - \sigma^2 + \delta)(T-t)\right] = d - y_t. \tag{OA.23}$$

Proof. As show in (A.9), the stock price at time $t \leqslant \tau \wedge T$ follows (OA.21). We only need to prove (OA.22). Taking S_0 as given and imposing the condition for stock price at the circuit breaker trigger, we have

$$\mathbb{E}_t\left[S_{\tau\wedge T}^{-1}\right] = \frac{1}{(1-\alpha)S_0} P_t(\tau \leqslant T) + \mathbb{E}_t\left[D_T^{-1}\mathbb{1}_{\{\tau>T\}}\right]. \tag{OA.24}$$

The following standard results about hitting times of Brownian motions are helpful for deriving the expressions for the expectations in (OA.24) (see e.g., Jeanblanc, Yor, and Chesney, 2009, chap 3). Let Z^μ denote a drifted Brownian motion, $Z_t^\mu = \mu t + Z_t$, with $Z_0^\mu = 0$. Let $\mathcal{T}_y^\mu = \inf\{t \geqslant 0 : Z_t^\mu = y\}$ for $y < 0$. Then:

$$\Pr\left(T_y^\mu \leqslant t\right) = N\left(\frac{y-\mu t}{\sqrt{t}}\right) + e^{2\mu y} N\left(\frac{y+\mu t}{\sqrt{t}}\right), \tag{OA.25}$$

$$E\left[e^{-\lambda T_y^\mu}\mathbb{1}_{\{T_y^\mu \leqslant t\}}\right] = e^{(\mu-\gamma)y} N\left(\frac{y-\gamma t}{\sqrt{t}}\right) + e^{(\mu+\gamma)y} N\left(\frac{y+\gamma t}{\sqrt{t}}\right), \tag{OA.26}$$

where $\gamma = \sqrt{2\lambda + \mu^2}$.

Recall the definition of the stopping time τ in Equation (28), which simplifies in the case with constant disagreement,

$$\tau = \inf\left\{t \geqslant 0 : D_t \leqslant (1-\alpha)S_0 e^{-\left(\mu-\sigma^2+\delta\right)(T-t)}\right\}. \tag{OA.27}$$

Through a change of variables, we can redefine τ as the first hitting time of a drifted Brownian motion for a constant threshold. Specifically, define:

$$y_t = \frac{1}{\sigma}\log\left(e^{-\left(\mu-\sigma^2+\delta\right)t} D_t\right), \tag{OA.28}$$

then $y_0 = 0$, and

$$y_t = Z_t^{\frac{\sigma}{2}-\frac{\delta}{\sigma}} = \left(\frac{\sigma}{2} - \frac{\delta}{\sigma}\right)t + Z_t. \tag{OA.29}$$

Moreover,

$$\mathcal{T}_d^{\frac{\sigma}{2}-\frac{\delta}{\sigma}} = \inf\{t \geqslant 0 : y_t \leqslant d\} \overset{a.s.}{=} \tau, \tag{OA.30}$$

where the threshold is constant over time,

$$d = \frac{1}{\sigma}\log\left((1-\alpha)S_0 e^{-\left(\mu-\sigma^2+\delta\right)T}\right). \tag{OA.31}$$

Conditional on y_t and the fact that the circuit breaker has not been triggered up to time t, the result from (OA.25) implies:

$$\begin{aligned} P_t(\tau \leqslant T) &= P_t\left(\mathcal{T}_{d_t}^{\frac{\sigma}{2}-\frac{\delta}{\sigma}} \leqslant T - t\right) \\ &= N\left[\frac{d_t - \left(\frac{\sigma}{2}-\frac{\delta}{\sigma}\right)(T-t)}{\sqrt{T-t}}\right] + e^{\left(\sigma-\frac{2\delta}{\sigma}\right)d_t} N\left[\frac{d_t + \left(\frac{\sigma}{2}-\frac{\delta}{\sigma}\right)(T-t)}{\sqrt{T-t}}\right], \end{aligned} \tag{OA.32}$$

where

$$d_t = d - y_t = \frac{1}{\sigma}\left[\log\left(\frac{(1-\alpha)S_0}{D_t}\right) - \left(\mu - \sigma^2 + \delta\right)(T-t)\right]. \tag{OA.33}$$

The threshold d is normalized with respect to y_t so as to start the drifted Brownian motion $Z^{\frac{\sigma}{2}-\frac{\delta}{\sigma}}$ from 0 at time t. Next,

$$\begin{aligned}
\mathbb{E}_t\left[D_T^{-1}\mathbb{1}_{\{\tau>T\}}\right] &= D_t^{-1}e^{-\left(\mu-\frac{\sigma^2}{2}\right)(T-t)}\mathbb{E}_t\left[e^{-\sigma(Z_T-Z_t)}\mathbb{1}_{\{\tau>T\}}\right] = D_t^{-1}e^{-\left(\mu-\sigma^2\right)(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[\mathbb{1}_{\{\tau>T\}}\right] \\
&= D_t^{-1}e^{-\left(\mu-\sigma^2\right)(T-t)}\left\{N\left[-\frac{d_t+\left(\frac{\sigma}{2}+\frac{\delta}{\sigma}\right)(T-t)}{\sqrt{T-t}}\right]\right. \\
&\quad \left. - e^{-(\sigma+\frac{2\delta}{\sigma})d_t}N\left[\frac{d_t-\left(\frac{\sigma}{2}+\frac{\delta}{\sigma}\right)(T-t)}{\sqrt{T-t}}\right]\right\}.
\end{aligned} \tag{OA.34}$$

The second equality follows from Girsanov's Theorem, and the third equality again follows from (OA.25). Under $\mathbb{Q}$, $Z_t^\sigma = Z_t + \sigma t$ is a standard Brownian motion, and

$$y_t = -\left(\frac{\sigma}{2}+\frac{\delta}{\sigma}\right)t + Z_t^\sigma. \qquad \square$$

OA.2.8 Proof of Lemma A.3

(i) When $y_t = d$, $\tau = t$ and the statement follows from (A.8).

(ii) Setting $t = \tau \wedge T$ in (A.11), we obtain $\mathbb{E}^B_{\tau\wedge T}[D_T^{-1}] = e^{-(\mu-\sigma^2+\delta)T - \frac{1}{2}\sigma^2 T\wedge\tau - \sigma Z^\delta_{T\wedge\tau}}$, hence:

$$\begin{aligned}
f^A(y_t,t) &= \mathbb{E}_t\Big[\mathbb{E}^B_{\tau\wedge T}[D_T^{-1}]\Big] = e^{-(\mu-\sigma^2+\delta)T}\mathbb{E}_t\Big[e^{-\frac{1}{2}\sigma^2 T\wedge\tau-\sigma Z^\delta_{T\wedge\tau}}\Big] \\
&= e^{-(\mu-\sigma^2+\delta)T}e^{-\frac{1}{2}\sigma^2 t-\sigma Z_t}\mathbb{E}_t^{\hat{A}}\big[e^{\delta(T\wedge\tau)}\big] \;=\; e^{-(\mu-\sigma^2+\delta)T}e^{-\frac{1}{2}\sigma^2 t-\sigma Z_t+\delta t}\mathbb{E}_t^{\hat{A}}\Big[e^{\delta(T\wedge\tau-t)}\Big] \\
&= e^{-(\mu-\sigma^2+\delta)T-\sigma y_t}\mathbb{E}_t^{\hat{A}}\Big[e^{\delta(T\wedge\tau-t)}\Big],
\end{aligned} \tag{OA.35}$$

where $Z_t^\delta = Z_t - \frac{\delta}{\sigma}t$ is a $\mathbb{P}^B$-Brownian motion and $\frac{d\mathbb{P}^{\hat{A}}}{d\mathbb{P}} = e^{-\frac{1}{2}\sigma^2 T\wedge\tau-\sigma Z_{T\wedge\tau}}$. Then $f^B(y,t) > f^A(y,t)$ follows from comparing (A.11), (OA.35), and using $\delta < 0$ together with $T\wedge\tau > t$ when $y_t > d$. When $y \uparrow \infty$, $\tau \to \infty$. Then $\lim_{y\uparrow\infty} f^A(y,t) = \lim_{y\uparrow\infty}\hat{f}^A(y,t)$ follows from comparing (A.10) and (OA.35).

(iii) The explicit form of f^A is obtained in (OA.22) where $d_t = d - y_t$. Take derivative with respect to y and using $D_t = e^{(\mu-\sigma^2+\delta)t+\sigma y_t}$ together with the definition of d, we obtain:

$$
\begin{aligned}
&\partial_y f^A(y,t) \\
&= -\, e^{-(\mu-\sigma^2+\delta)T-\sigma d} \left\{ \phi\Big[\frac{d_t - \big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\big)(T-t)}{\sqrt{T-t}}\Big]\frac{1}{\sqrt{T-t}} \right. \\
&\quad + e^{(\sigma-\frac{2\delta}{\sigma})d_t}\phi\Big[\frac{d_t + \big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\big)(T-t)}{\sqrt{T-t}}\Big]\frac{1}{\sqrt{T-t}} \\
&\quad \left. + \Big(\sigma - \frac{2\delta}{\sigma}\Big)e^{(\sigma-\frac{2\delta}{\sigma})d_t}N\Big[\frac{d_t + \big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\big)(T-t)}{\sqrt{T-t}}\Big]\right\} \\
&\quad -\; e^{-(\mu-\sigma^2+\delta)T-\sigma y_t+\delta(T-t)} \left\{ \sigma N\Big[-\frac{d_t + \big(\frac{\sigma}{2} + \frac{\delta}{\sigma}\big)(T-t)}{\sqrt{T-t}}\Big] \right. \\
&\quad - \sigma\, e^{-(\sigma+\frac{2\delta}{\sigma})d_t}N\Big[\frac{d_t - \big(\frac{\sigma}{2} + \frac{\delta}{\sigma}\big)(T-t)}{\sqrt{T-t}}\Big] \\
&\quad - \phi\Big[-\frac{d_t + \big(\frac{\sigma}{2} + \frac{\delta}{\sigma}\big)(T-t)}{\sqrt{T-t}}\Big]\frac{1}{\sqrt{T-t}} - e^{-(\sigma+\frac{2\delta}{\sigma})d_t}\phi\Big[\frac{d_t - \big(\frac{\sigma}{2} + \frac{\delta}{\sigma}\big)(T-t)}{\sqrt{T-t}}\Big]\frac{1}{\sqrt{T-t}} \\
&\quad \left. + \Big(\sigma + \frac{2\delta}{\sigma}\Big)e^{-(\sigma+\frac{2\delta}{\sigma})d_t}N\Big[\frac{d_t - \big(\frac{\sigma}{2} + \frac{\delta}{\sigma}\big)(T-t)}{\sqrt{T-t}}\Big]\right\},
\end{aligned}
\tag{OA.36}
$$

where $\phi(\cdot)$ is the standard normal density. Sending $y \uparrow \infty$ (hence $d_t \downarrow -\infty$), we obtain from (OA.36) that

$$
\lim_{y_t\uparrow\infty}\frac{\partial_y f^A(y_t,t)}{f^A(y_t,t)} = \lim_{y_t\uparrow\infty}\frac{-\sigma e^{-(\mu-\sigma+\delta)^2 T-\sigma y_t+\delta(T-t)}}{f^A(y_t,t)} = -\sigma,
\tag{OA.37}
$$

where the identity $\frac{\partial_y f^B}{f^B} \equiv -\sigma$ follows from (A.11) directly.

(iv) Sending $y \downarrow d$ (equivalently $d_t \uparrow 0$), we obtain from (OA.36) that:

$$
\begin{aligned}
&\partial_y f^A(y,t)\big|_{y=d} \\
&= -\, e^{-(\mu-\sigma^2+\delta)T-\sigma d}\left\{2\phi\Big[\Big(\frac{\sigma}{2}-\frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big]\frac{1}{\sqrt{T-t}} + \Big(\sigma-\frac{2\delta}{\sigma}\Big)N\Big[\Big(\frac{\sigma}{2}-\frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big)\Big]\right\} \\
&\quad -\, e^{-(\mu-\sigma^2+\delta)T-\sigma d+\delta(T-t)}\left\{-2\phi\Big[-\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big]\frac{1}{\sqrt{T-t}} \right. \\
&\quad \left. + \Big(\sigma+\frac{2\delta}{\sigma}\Big)N\Big[-\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big]\right\} \\
&= -\, e^{-(\mu-\sigma^2+\delta)T-\sigma d}\left\{\Big(\sigma-\frac{2\delta}{\sigma}\Big)N\Big[\Big(\frac{\sigma}{2}-\frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big)\Big] \right. \\
&\quad \left. + e^{\delta(T-t)}\Big(\sigma+\frac{2\delta}{\sigma}\Big)N\Big[-\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big]\right\},
\end{aligned}
\tag{OA.38}
$$

where the second identity follows from:

$$
\phi\Big[\Big(\frac{\sigma}{2}-\frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big] = e^{\delta(T-t)}\phi\Big[-\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big].
\tag{OA.39}
$$

We claim that:

$$h(T-t) = \Big(\sigma - \frac{2\delta}{\sigma}\Big)N\Big[\Big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big)\Big] + e^{\delta(T-t)}\Big(\sigma + \frac{2\delta}{\sigma}\Big)N\Big[-\Big(\frac{\sigma}{2} + \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big]$$
$$> \sigma, \tag{OA.40}$$

if $\delta \leqslant -\frac{\sigma^2}{2}$. To see this, because $\delta \leqslant -\frac{\sigma^2}{2}$, we have:

$$\begin{aligned}
h(T-t) \ &> \ \Big(\sigma - \frac{2\delta}{\sigma}\Big)N\Big[\Big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big)\Big] + \Big(\sigma + \frac{2\delta}{\sigma}\Big)N\Big[-\Big(\frac{\sigma}{2} + \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big] \\
&= \ \sigma\Big\{N\Big[\Big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big)\Big] + N\Big[-\Big(\frac{\sigma}{2} + \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big]\Big\} \\
&\quad - \frac{2\delta}{\sigma}\Big\{N\Big[\Big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big)\Big] - N\Big[-\Big(\frac{\sigma}{2} + \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big]\Big\} \\
&> \ \sigma,
\end{aligned}$$

where the first inequality follows from $\delta < 0$ and $\sigma + \frac{2\delta}{\sigma} \leqslant 0$ and the second inequality holds due to:

$$N\Big[\Big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big)\Big] + N\Big[-\Big(\frac{\sigma}{2} + \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big] > 1,$$
$$N\Big[\Big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big)\Big] - N\Big[-\Big(\frac{\sigma}{2} + \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big] > 0.$$

Therefore the claim (OA.40) is confirmed.

Combining (OA.38) and (OA.40), we arrive at:

$$\partial_y f^A(y,t)\big|_{y=d} < -\sigma e^{-(\mu-\sigma^2)T-\sigma d} = \partial_y f^B(y,t)\big|_{y=d}. \tag{OA.41}$$

Meanwhile, it follows from (OA.38) and $\frac{2\delta}{\sigma} + \sigma \leqslant 0$ that:

$$\partial_y f^A(y,t)|_{y=d} > -e^{-(\mu-\sigma^2+\delta)T-\sigma d}\Big(\sigma - \frac{2\delta}{\sigma}\Big) = -\Big(\sigma - \frac{2\delta}{\sigma}\Big)f^A(d,t).$$

(v) Feynman-Kac formula implies that $f^A(y,t)$ satisfies the following boundary value problem:

$$\begin{aligned}
\partial_t f^A + \frac{1}{2}\partial^2_{yy} f^A + \Big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\Big)\partial_y f^A \ &= \ 0, \quad (y,t) \in (d,\infty) \times [0,T), \\
f^A(d,t) \ &= \ e^{-(\mu-\sigma^2+\delta)T-\sigma d}, \quad t \in [0,T], \\
f^A(y,T) \ &= \ e^{-(\mu-\sigma^2+\delta)T-\sigma y}, \quad y \in [d,\infty).
\end{aligned}$$

Here $\frac{1}{2}\partial^2_{yy} + \big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\big)\partial_y$ is the infinitesimal generator of y_t under $\mathbb{P}$. Define $g^A(y,t) = \log f^A(y,t)$. Then, g satisfies:

$$\begin{aligned}
\partial_t g^A + \frac{1}{2}\partial^2_{yy} g^A + \frac{1}{2}\big(\partial_y g^A\big)^2 + \Big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\Big)\partial_y g^A \ &= \ 0, \quad (y,t) \in (d,\infty) \times [0,T), \qquad \text{(OA.42)} \\
g^A(d,t) \ &= \ -(\mu-\sigma^2+\delta)T - \sigma d, \quad t \in [0,T], \\
g^A(y,T) \ &= \ -(\mu-\sigma^2+\delta)T - \sigma y, \quad y \in [d,\infty).
\end{aligned}$$

We claim that $\lim_{y\downarrow d} \partial^2_{yy} g(y,t) > 0$. To see this, it follows from item (iv) that $-\sigma > \partial_y g|_{y=d} >$

$-(\sigma - \frac{2\delta}{\sigma})$, hence:

$$\frac{1}{2}\partial_y g^A\Big(\partial_y g^A + \sigma - \frac{2\delta}{\sigma}\Big)\Big|_{y=d} < 0.$$

On the other hand, $\partial_t g^A(d,t) = 0$ for any $t < T$, because $g^A(d,t)$ is independent of t. Taking the limit $y \downarrow d$ in the first equation of (OA.42), we conclude $\lim_{y\downarrow d} \partial^2_{yy} g^A(y,t) > 0$. The statement in (v) follows from observing that $\partial_y\big(\frac{\partial_y f^A}{f^A}\big) = \partial^2_{yy} g^A$.

OA.2.9 A technical inequality

Lemma OA.6. *Suppose that $\delta \leqslant -\sigma^2$. Then,* (A.15) *holds.*

Proof. Combining (A.10), (A.11), and Lemma A.3 (i), we obtain:

$$\begin{aligned}
&\lim_{y\downarrow d} \frac{\omega^A_t f^A}{\omega^A_t f^A + (1-\omega^A_t) f^B} = \lim_{y\downarrow d} \omega^A_t, \\
&\lim_{y\downarrow d} \frac{(1-\omega^A_t)\omega^A_t f^B}{\omega^A_t \widehat{f}^A + (1-\omega^A_t) f^B} = \lim_{y\downarrow d} \omega^A_t \frac{1-\omega^A_t}{\omega^A_t e^{\delta(T-t)} + 1 - \omega^A_t}.
\end{aligned} \tag{OA.43}$$

Plugging (OA.43) into (A.15) and using the fact $\frac{1-\omega^A_t}{\omega^A_t e^{\delta(T-t)}+1-\omega^A_t} < 1$, we observe that (A.15) is satisfied when the following inequality holds.

$$\lim_{y\downarrow d}\Big(-\frac{\partial_y f^A}{f^A} - \sigma\Big) > \frac{\delta}{\sigma}\big(e^{\delta(T-t)} - 1\big). \tag{OA.44}$$

It remains to prove (OA.44). Define

$$g(T-t) = \lim_{y\downarrow d}\Big(-\frac{\partial_y f^A}{f^A} - \sigma\Big) - \frac{\delta}{\sigma}\big(e^{\delta(T-t)} - 1\big).$$

We need to show $g(T-t) > 0$ when $t < T$. We have from (OA.38) later that:

$$\begin{aligned}
g(T-t) \;\;=\;\; & \sigma\Big\{N\Big[\Big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big)\Big] + e^{\delta(T-t)} N\Big[-\Big(\frac{\sigma}{2} + \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big] - 1\Big\} \\
& -\frac{2\delta}{\sigma}\Big\{N\Big[\Big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big)\Big] - e^{\delta(T-t)} N\Big[-\Big(\frac{\sigma}{2} + \frac{\delta}{\sigma}\Big)\sqrt{T-t}\Big]\Big\} - \frac{\delta}{\sigma}\big(e^{\delta(T-t)} - 1\big).
\end{aligned}$$

Note that $g(0) = 0$. In order to show $g(s) > 0$ for any $s > 0$, it suffices to prove $g'(s) > 0$ for any $s > 0$. To this end, calculation shows that:

$$\begin{aligned}
g'(s) &= \sigma\Big\{\phi\Big[\Big(\frac{\sigma}{2}-\frac{\delta}{\sigma}\Big)\sqrt{s}\Big]\frac{1}{2}\Big(\frac{\sigma}{2}-\frac{\delta}{\sigma}\Big)s^{-\frac{1}{2}} + \delta e^{\delta s}N\Big[-\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)\sqrt{s}\Big] \\
&\quad - e^{\delta s}\phi\Big[-\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)\sqrt{s}\Big]\frac{1}{2}\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)s^{-\frac{1}{2}}\Big\} \\
&\quad - \frac{2\delta}{\sigma}\Big\{\phi\Big[\Big(\frac{\sigma}{2}-\frac{\delta}{\sigma}\Big)\sqrt{s}\Big]\frac{1}{2}\Big(\frac{\sigma}{2}-\frac{\delta}{\sigma}\Big)s^{-\frac{1}{2}} - \delta e^{\delta s}N\Big[-\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)\sqrt{s}\Big] \\
&\quad + e^{\delta s}\phi\Big[-\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)\sqrt{s}\Big]\frac{1}{2}\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)s^{-\frac{1}{2}}\Big\} - \frac{\delta^2}{\sigma}e^{\delta s} \\
&= \sigma\Big\{\phi\Big[\Big(\frac{\sigma}{2}-\frac{\delta}{\sigma}\Big)\sqrt{s}\Big]\frac{-\delta}{\sigma}s^{-\frac{1}{2}} + \delta e^{\delta s}N\Big[-\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)\sqrt{s}\Big]\Big\} \\
&\quad - \frac{2\delta}{\sigma}\Big\{\phi\Big[\Big(\frac{\sigma}{2}-\frac{\delta}{\sigma}\Big)\sqrt{s}\Big]\frac{\sigma}{2}s^{-\frac{1}{2}} - \delta e^{\delta s}N\Big[-\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)\sqrt{s}\Big]\Big\} - \frac{\delta^2}{\sigma}e^{\delta s} \\
&= -\frac{2\delta}{\sqrt{2\pi}}e^{\delta s}e^{-\frac{1}{2}(\frac{\delta}{\sigma}+\frac{\sigma}{2})^2 s}s^{-\frac{1}{2}} + \Big(\frac{2\delta}{\sigma}+\sigma\Big)\delta e^{\delta s}N\Big[-\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)\sqrt{s}\Big] - \frac{\delta^2}{\sigma}e^{\delta s} \\
&= e^{\delta s}h(s),
\end{aligned}$$

where the second identity follows from (OA.39), the third identity holds due to $\phi\Big[\Big(\frac{\sigma}{2}-\frac{\delta}{\sigma}\Big)\sqrt{s}\Big] = \frac{1}{\sqrt{2\pi}}e^{\delta s}e^{-\frac{1}{2}(\frac{\delta}{\sigma}+\frac{\sigma}{2})^2 s}$, and

$$h(s) = -\frac{2\delta}{\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{\delta}{\sigma}+\frac{\sigma}{2})^2 s}s^{-\frac{1}{2}} + \Big(\frac{2\delta}{\sigma}+\sigma\Big)\delta N\Big[-\Big(\frac{\sigma}{2}+\frac{\delta}{\sigma}\Big)\sqrt{s}\Big] - \frac{\delta^2}{\sigma}. \tag{OA.45}$$

When $\delta \leqslant -\sigma^2$, we have $\lim_{s\uparrow\infty} h(s) = \big(\frac{\delta}{\sigma}+\sigma\big)\delta \geqslant 0$. Moreover, because $\delta < 0$,

$$h'(s) = \frac{\delta}{\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{\delta}{\sigma}+\frac{\sigma}{2})^2 s}s^{-\frac{3}{2}} < 0.$$

Therefore, h is a decreasing positive function.

Coming back to the expression of $g'(s)$, we confirm that $g'(s) > 0$ for any $s > 0$. Therefore $g(s) > 0$ for any $s > 0$, and (OA.44) is verified. □

OA.2.10 A continuity argument

The expression of $\sigma^{\widehat{S}}$ is given in (A.12), where $\widehat{f}^A$ and f^B are given in (A.10) and (A.11), $\omega_t^A = \frac{\omega}{\omega+(1-\omega)\eta_t}$ with $\eta_t = e^{(\frac{\delta}{\sigma}y_t - \frac{\delta}{2}t + \frac{\delta^2}{2\sigma^2}t}$. As a result, $\sigma^{\widehat{S}}$ is a continuous differentiable function of (ω, y, t) in the domain $[0,1]\times[d,\infty)\times[0,T)$. Using $\lim_{\omega\uparrow 1}\omega_t^A = 1$ and $\lim_{\omega\uparrow 1}\partial_y\omega_t^A = 0$, we obtain from calculation that:

$$\lim_{\omega\uparrow 1}\partial_y\sigma_t^{\widehat{S}} = 0, \qquad \text{for any } y \in [d,\infty). \tag{OA.46}$$

Meanwhile, the expression of σ^S is given in (A.13), where f^A is given by (OA.22) in the Online Appendix. Then, σ^S is also a continuous differentiable function of (ω, y, t) in the domain $[0,1]\times[d,\infty)\times[0,T)$. Using $\lim_{\omega\uparrow 1}\omega_t^A = 1$ and $\lim_{\omega\uparrow 1}\partial_y\omega_t^A = 0$, we obtain from calculation that:

$$\lim_{\omega\uparrow 1}\partial_y\sigma_t^S = -\partial_y\Big(\frac{\partial_y f^A}{f^A}\Big), \qquad \text{for any } y\in[d,\infty). \tag{OA.47}$$

Using Lemma A.3 (v), we obtain from (OA.46) and (OA.47) that:

$$\lim_{\omega\uparrow 1}\partial_y\big(\sigma_t^S-\sigma_t^{\hat S}\big)\Big|_{y=d}<0. \tag{OA.48}$$

Because $\partial_y(\sigma_t^S-\sigma_t^{\hat S})$ is a continuous function in (ω,y), $\partial_y(\sigma_t^S-\sigma_t^{\hat S})<0$ in a neighborhood of $(\omega,y)=(1,d)$. Therefore, the statement follows from the limiting case that $\lim_{y\downarrow d}\sigma_t^S > \lim_{y\downarrow d}\sigma_t^{\hat S}$ as $\omega\uparrow 1$.

OA.2.11 Proof of Lemma A.4

Recall that the market clearing price follows:

$$dS_t=\mu(t,S_t)dt+\sigma(t,S_t)dZ_t,$$

where the expected return μ and the volatility σ are functions of time and price. Define an equivalent measure $\mathbb{Q}$ via $\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t}=\exp\big(-\int_0^t\lambda(u,S_u)dZ_u-\frac12\int_0^t\lambda^2(u,S_u)du\big)$, where $\lambda(u,S_u)=\frac{\mu(u,S_u)}{\sigma(u,S_u)}$. Then,

$$dS_t=\sigma(t,S_t)dZ_t^{\mathbb{Q}},$$

where $dZ_t^{\mathbb{Q}}=dZ_t+\lambda(t,S_t)dt$ is a Brownian motion under $\mathbb{Q}$.

Consider the density of threshold trigger time $\mathbb{P}(\tau\in du|\mathcal{F}_t)$ with $u\in[t,t+h]$,

$$\begin{aligned}\mathbb{P}(\tau\in du|\mathcal{F}_t)&=\Big(\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_t}\Big)\mathbb{E}^{\mathbb{Q}}\Big[\Big(\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_u}\Big)^{-1}1_{\{\tau\in du\}}\Big|\mathcal{F}_t\Big]\\&=\mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(\int_t^u\lambda(v,S_v)dZ_v+\frac12\int_t^u\lambda^2(v,S_v)dv\Big)1_{\{\tau\in du\}}\Big|\mathcal{F}_t\Big]\end{aligned}$$

As $h\downarrow 0$, $u\to t$, the right-hand side is approximated by the following expression, with an error $o(u-t)$,

$$\begin{aligned}&\mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(\frac{S_u-S_t}{\sigma(u,S_u)}\lambda(u,S_u)-\frac12\lambda^2(u,S_u)(u-t)\Big)1_{\{\tau\in du\}}\Big|\mathcal{F}_t\Big]\\&=e^{-a\lambda(u,\underline{S})-\frac12\lambda^2(u,\underline{S})(u-t)}\mathbb{Q}(\tau\in du|\mathcal{F}_t),\end{aligned}$$

where $a=\frac{S_t-\underline{S}}{\sigma(u,\underline{S})}$ and the equality follows from the fact that $S_u=\underline{S}$ when $\tau\in du$. Now recall the hitting time density of a Brownian motion (see e.g. Karatzas and Shreve, 2.8.3) $\mathbb{Q}(\tau\in du|\mathcal{F}_t)=\frac{a}{\sqrt{2\pi(u-t)^3}}e^{-\frac{a^2}{2(u-t)}}du$. We obtain from the previous two equations that:

$$\mathbb{P}(\tau\in du|\mathcal{F}_t)=\frac{a}{\sqrt{2\pi(u-t)^3}}e^{-\frac{(a+\lambda(u,\underline{S})(u-t))^2}{2(u-t)}}du+o(u-t).$$

Integrating the density on $[t,t+h]$, we obtain:

$$\mathbb{P}(\tau\in du|\mathcal{F}_t)=\int_t^{t+h}\frac{a}{\sqrt{2\pi(u-t)^3}}e^{-\frac{(a+\lambda(u,\underline{S})(u-t))^2}{2(u-t)}}du+o(h). \tag{OA.49}$$

We claim that the right-hand side of (OA.49) converges to:

$$\frac{2}{\sqrt{2\pi}}\int_{1/\sigma(t,\underline{S})}^{\infty} e^{-\frac{x^2}{2}}dx, \quad \text{as } h \downarrow 0. \tag{OA.50}$$

Then the statement in (i) follows from $\sigma(t,\underline{S}) = \underline{S}\sigma_t^S|_{S_t=\underline{S}}$ and the symmetry of N.

To show the convergence of the right-hand side of (OA.49), introduce $\theta \in [0,1]$ such that $u - t = \theta h$. Using the assumption $h = (S_t - \underline{S})^2$, we obtain $a = \frac{\sqrt{h}}{\sigma(u,\underline{S})}$ and

$$\int_t^{t+h} \frac{a}{\sqrt{2\pi(u-t)^3}} e^{-\frac{(a+\lambda(u,\underline{S})(u-t))^2}{2(u-t)}}du = \int_0^1 \frac{1}{\sigma(u,\underline{S})\sqrt{2\pi\theta^3}} e^{-\frac{\left(\frac{1}{\sigma(u,\underline{S})}+\lambda(u,\underline{S})\theta h\right)^2}{2\theta}}d\theta.$$

Sending $h \downarrow 0$, the right-hand side converges to:

$$\int_0^1 \frac{1}{\sigma(t,\underline{S})\sqrt{2\pi\theta^3}} e^{-\frac{1}{2\theta\sigma^2(t,\underline{S})}}d\theta.$$

Changing the variable to $x = \frac{\theta^{-\frac{1}{2}}}{\sigma(t,\underline{S})}$ in the previous integral, we obtain (OA.50).

For the statement in (ii), the explicit expression of $p_h(S_t)$ is given in (OA.32), which is a function of h and y_t. To calculate $\partial_{\Delta S_t} p_{(\Delta S_t)^2}(S_t)$, we first calculate $\partial_{y_t} p_h(S_t)$. Consider $p_h(S_t)$ as a function of $\sqrt{h}$ and y_t, h depends on S_t, hence it also depends on y_t, we obtain from (OA.32) that:

$$\begin{aligned}
\partial_{y_t} p_h(S_t) =& \phi\Big[\frac{d - y_t - \big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\big)h}{\sqrt{h}}\Big]\big(-\frac{1}{\sqrt{h}}\big) \\
&+ e^{(\sigma - \frac{2\delta}{\sigma})(d-y_t)}\phi\Big[\frac{d - y_t + \big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\big)h}{\sqrt{h}}\Big]\big(-\frac{1}{\sqrt{h}}\big) \\
&- \big(\sigma - \frac{2\delta}{\sigma}\big)e^{(\sigma - \frac{2\delta}{\sigma})(d-y_t)}N\Big[\frac{d - y_t + \big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\big)h}{\sqrt{h}}\Big] \\
&+ \phi\Big[\frac{d - y_t - \big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\big)h}{\sqrt{h}}\Big]\Big[-\frac{d-y_t}{h} - \big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\big)\Big]\partial_{y_t}\sqrt{h} \\
&+ e^{(\sigma - \frac{2\delta}{\sigma})(d-y_t)}\phi\Big[\frac{d - y_t + \big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\big)h}{\sqrt{h}}\Big]\Big[-\frac{d-y_t}{h} + \big(\frac{\sigma}{2} - \frac{\delta}{\sigma}\big)\Big]\partial_{y_t}\sqrt{h}.
\end{aligned} \tag{OA.51}$$

Meanwhile, it follows from Taylor's expansion and the first equality in (A.13) that:

$$\begin{aligned}
\sqrt{h} = \Delta S_t &= S_t(y_t,t) - S_t(d,t) \\
&= \partial_y S_t(d,t)(y_t - d) + o(y_t - d) = \underline{S}\sigma_t^S|_{S_t=\underline{S}}(y_t - d) + o(y_t - d).
\end{aligned} \tag{OA.52}$$

Moreover,

$$\partial_{y_t}\sqrt{h}\big|_{y_t=d} = \partial_{y_t}\Delta S_t\big|_{y_t=d} = \underline{S}\sigma_t^S|_{S_t=\underline{S}}. \tag{OA.53}$$

Combining (OA.51), (OA.52), (OA.53), and sending $y_t \downarrow d$, we obtain:

$$\partial_{y_t} p_h(S_t)\big|_{y_t=d} = -\Big(\sigma - \frac{2\delta}{\sigma}\Big)N\Big[-\frac{1}{\underline{S}\sigma^S|_{S_t=\underline{S}}}\Big]. \tag{OA.54}$$

Finally,

$$\partial_{\Delta S_t} p_{(\Delta S_t)^2}(S_t)\Big|_{S_t=\underline{S}} = \partial_{\Delta S_t} y_t\, \partial_{y_t} p_h(S_t)\big|_{y_t=d} = -\frac{\sigma-\frac{2\delta}{\sigma}}{\underline{S}\sigma_t^S|_{S_t=\underline{S}}} N\Big[-\frac{1}{\underline{S}\sigma_t^S|_{S_t=\underline{S}}}\Big].$$

For the statement in (iii) about the complete market, in the limit $\omega \to 1$,

$$\widehat{S}_t = \big(\widehat{f}^A(y_t,t)\big)^{-1} = e^{(\mu-\sigma^2+\delta)T+\sigma y_t-\delta(T-t)}.$$

Therefore the hitting time $\widehat{\tau} = \inf\{t \geqslant 0 : \widehat{S}_t \leqslant (1-\alpha)S_0\}$ is equivalent to $\widehat{\tau} = \inf\{t \geqslant 0 : y_t + \frac{\delta}{\sigma}t \leqslant d + \frac{\delta}{\sigma}T\}$. Define:

$$\widehat{y}_t = y_t + \frac{\delta}{\sigma}t = \frac{\sigma}{2}t + Z_t, \quad \widehat{d} = d + \frac{\delta}{\sigma}T = \frac{1}{\sigma}\log\big((1-\alpha)S_0\big) - \frac{1}{\sigma}(\mu-\sigma^2)T.$$

Then, $\widehat{S}_t = e^{(\mu-\sigma^2)T+\sigma\widehat{y}_t}$. We follow the same proof leading to (OA.32) that:

$$\hat{p}_h(\widehat{S}_t) = P_t(\widehat{\tau} \leqslant h) = N\Big[\frac{\widehat{d}-\widehat{y}_t-\frac{\sigma}{2}h}{\sqrt{h}}\Big] + e^{\sigma(\widehat{d}-\widehat{y}_t)}N\Big[\frac{\widehat{d}-\widehat{y}_t+\frac{\sigma}{2}h}{\sqrt{h}}\Big]. \tag{OA.55}$$

We then follow the argument leading to (OA.54) to obtain:

$$\partial_{\widehat{y}_t}\hat{p}_h(\widehat{S}_t)\big|_{\widehat{y}_t=\widehat{d}} = -\sigma N\Big[-\frac{1}{\sigma\underline{S}}\Big]. \tag{OA.56}$$

Finally, combining the previous equality with $\partial_{\widehat{y}_t}\widehat{S}_t\big|_{\widehat{y}_t=\widehat{d}} = \underline{S}\sigma$, we conclude that:

$$\partial_{\Delta\widehat{S}_t}\hat{p}_{(\Delta\widehat{S}_t)^2}(\widehat{S}_t)\Big|_{\widehat{S}_t=S} = \partial_{\Delta\widehat{S}_t}\widehat{y}_t\,\partial_{\widehat{y}_t}\hat{p}_h(S_t)\big|_{\widehat{y}_t=\widehat{d}} = -\frac{1}{\underline{S}}N\Big[-\frac{1}{\sigma\underline{S}}\Big].$$

OA.2.12 Hitting probability when ω is close to 1

To prove the statement of Proposition 7 when ω is sufficiently close to 1, we first prove the statement in the limit of $\omega \to 1$, then apply a continuation argument.

First, Lemma A.3. (iv) implies that:

$$\sigma < -\frac{\partial_y f^A(y,t)|_{y=d}}{f^A(d,t)} < \sigma - (2\delta/\sigma).$$

Then combining the previous inequalities with (A.14), we have:

$$\sigma < \sigma_t^S|_{S_t=\underline{S}} < \sigma - (2\delta/\sigma). \tag{OA.57}$$

Therefore:

$$N\left(-\frac{1}{\sigma\underline{S}}\right) < N\left(-\frac{1}{\underline{S}\sigma_t^S|_{S_t=\underline{S}}}\right) \quad \text{and} \quad -\frac{1}{\underline{S}} > -\frac{\sigma-(2\delta/\sigma)}{\underline{S}\sigma_t^S|_{S_t=\underline{S}}}. \tag{OA.58}$$

Combining Lemma A.4 (ii) and (iii) with (OA.58), we obtain in the limit $\omega \to 1$ that

$$\partial_{\Delta S_t} p_{(\Delta S_t)^2}(S_t)\Big|_{S_t=\underline{S}} < \partial_{\Delta S_t}\hat{p}_{(\Delta S_t)^2}(S_t)\Big|_{S_t=\underline{S}} < 0.$$

This implies that $p_{(\Delta S_t)^2}(S_t) - \hat{p}_{(\Delta S_t)^2}(S_t)$ increases as $S_t \downarrow \underline{S}$ in a neighborhood of $\underline{S}$. We have seen that $\lim_{S_t\downarrow\underline{S}} p_{(\Delta S_t)^2}(S_t) > \lim_{S_t\downarrow\underline{S}} \hat{p}_{(\Delta S_t)^2}(S_t)$. Therefore the limit of $p_{(\Delta S_t)^2}(S_t) - \hat{p}_{(\Delta S_t)^2}(S_t)$ as $S_t \downarrow \underline{S}$ is a positive constant.

When ω is sufficiently close to 1, we apply a continuity argument, similar to the proof of Proposition 6 (iii), to obtain the same statement. The neighborhood in different items of Proposition 7 may not be the same. However, taking the intersection of all these different neighborhood, all statements of the proposition hold in a common neighborhood.

OA.3 Market Behavior with Constant Disagreement

In addition to the analytical results presented in the paper for the constant disagreement case, we now examine quantitatively the impact of the circuit breaker on price and trading behavior over the whole state space. Overall, the quantitative results are consistent with the analytical results obtained under more stringent conditions.

The calibration is the same as in the paper for the case of time-varying disagreement: $T = 1$ to denote one trading day, the dividend growth $\mu = 10\%/250 = 0.04\%$, its (daily) volatility $\sigma = 3\%$. The circuit breaker threshold is $\alpha = 5\%$. Initially, agent A owns 90% of total wealth ($\omega = 0.9$ at $t = 0$). For constant disagreement, we set $\delta = -2\%$. This means that agent B is relatively pessimistic about dividend growth, and his private assessment of the stock at $t = 0$, $\widehat{S}_0^B$, will be 2% lower than that of agent A, $\widehat{S}_0^A$, which is fairly modest. Since the impact of the circuit breaker changes during the day, we choose $t = 0.1$ to fix ideas. The results for different t are given in the next section.

Figure OA.2 plots the equilibrium price-dividend ratio S_t/D_t (panel A), the conditional volatility of returns $\sigma_{S,t}$ (panel B), the conditional expected return $\mu_{S,t}^A$ (panel C), and the stock holding for agent A (panel D), respectively. In each panel, the solid blue line denotes the solution for the case with the circuit breaker, while the red dotted line denotes the case without the circuit breaker.

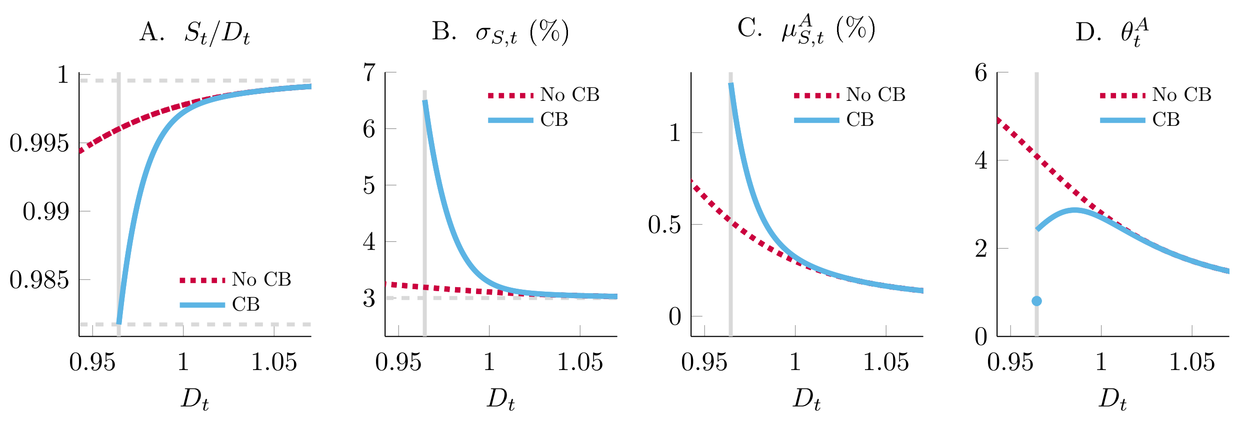

Figure OA.2: Price-dividend ratio, conditional return volatility, conditional expected return, and agent A's (relatively optimistic) portfolio holding at $t = 0.1$ in the case of constant disagreement. Blue solid lines are for the case with the circuit breaker. Red dotted lines are for the case without the circuit breaker. The gray vertical bars denote the circuit breaker threshold $\underline{D}(t)$. The parameters are $\mu = 0.04\%$, $\sigma = 3\%$, $\delta = -2\%$, $\omega = 0.9$, and $\alpha = 5\%$.

Price-dividend ratio. Let's start with the price-dividend ratio, shown in panel A. Under the calibration above, the price-dividend ratio is $1 = \widehat{S}_t^A/D_t$ for any $t \in [0, T]$ under agent A's belief, and it is $e^{\delta(T-t)} = \widehat{S}_t^A/D_t \leqslant 1$ under agent B's belief. These two limiting values are denoted by the upper and lower horizontal dash lines.

When both agents are present, the price of the stock without the circuit breaker, shown by the red dotted line, is the (harmonic) average of the two limiting price levels, weighted by agents' wealth shares. Indeed, it lies between the two gray dash lines. Since agent A is relatively more optimistic, she will hold levered position in the stock, and her share of total wealth will become higher following positive dividend shocks. Thus, as dividend D_t rises (falls), the share of total wealth owned by agent A increases (decreases), which makes the equilibrium price-dividend ratio approach the value $\widehat{S}_t^A/D_t$ ($\widehat{S}_t^B/D_t$).

With the circuit breaker, the price-dividend ratio, shown by the blue solid line, still lies between the price-dividend ratios from the two representative agent economies, but it is always below the price-dividend ratio without the circuit breaker for any given level of dividend D_t. The gap between the two price-dividend ratios is negligible when D_t is sufficiently high, but it widens as D_t decreases. Moreover, the price-dividend ratio drops drastically as D_t approaches the circuit breaker threshold $\underline{D}_t$, hitting $\widehat{S}_t^B/D_t$ at the threshold. As explained in the paper, this is simply because B, the more pessimistic agent, becomes the marginal investor at the market closure. The accelerating drop in price near the circuit breaker when D_t decreases implies increasing price sensitivity to dividend changes.

Conditional return volatility. The higher sensitivity of the price-dividend ratio to dividend shocks in the presence of the circuit breaker leads to higher conditional return volatility, as shown in panel B of Figure OA.2. Without the circuit breaker, the conditional return volatility (red dotted line) peaks at about 3.3%, only slightly higher than the fundamental volatility of $\sigma = 3\%$. This small amount of excess volatility comes from the time variation in the wealth distribution between the two agents. With the circuit breaker, the conditional volatility (blue solid line) becomes substantially higher as D_t approaches $\underline{D}_t$. In particular, the conditional volatility reaches 6.5% at the circuit breaker threshold, almost twice as high as the return volatility without circuit breaker.

Conditional expected return. The conditional expected return $\mu_{S,t}^A$ on the stock varies with the level of D_t as shown in panel C of Figure OA.2. The red dotted line shows the case without the circuit breaker. Clearly, as D_t decreases, agent B, who is the more pessimistic agent, becomes more dominant in the market and depresses the stock price. As a result, $\mu_{S,t}^A$ increases. When the circuit breaker is present, the above process is substantially magnified. Especially, when the market approaches the circuit breaker, the stock price drops quickly towards the threshold, which significantly increases its conditional expected return, as shown by the blue solid line.

Agents' stock holdings. We can also analyze the impact of the circuit breaker on the market by considering how it influences the agents' stock holdings. In the case without the circuit breaker, shown by the red dotted line in panel D of Figure OA.2, the stock holding of agent A, $\widehat{\theta}_t^A$, continues to rise as D_t falls to $\underline{D}_t$ and beyond. As discussed in the paper, this is because with lower D_t, the stock price is lower, agent A then has higher expected return on the stock and is willing to take on more levered positions, and agent B (who is shorting the stock) becomes wealthier and thus more capable of lending to agent A.

With the circuit breaker, while the stock holding θ_t^A takes on similar values as in the case without the circuit breaker, shown by the red dotted line, for large values of D_t, it becomes visibly lower as D_t approaches the circuit breaker threshold, and eventually starts to decrease as D_t continues to drop, as shown by the blue solid line. This is because agent A becomes increasingly concerned with the rising return volatility at lower D_t, which eventually dominates the effect of higher expected return. Finally, θ_t^A takes a discrete drop when $D_t = \underline{D}_t$. With the leverage constraint binding, agent A will hold all of his wealth in the stock, which means θ_t^A will be equal to his wealth share ω_t^A. The stock price in equilibrium has to fall enough such that agent A has no incentive to sell more of his stock holding. This is indeed the case as panel A shows.

OA.4 Time-of-the-Day Effect

The impact of the circuit breaker on the market also depends on the time of the day during the trading session. Intuitively, we expect the impact to be larger earlier in the day. When close the end of day, i.e., t is close to T, the circuit breaker should have a small impact since the market will close at T anyway.

In Figure OA.3, we compare the market behavior for two different times in the day, $t = 0.25$ and $t = 0.75$. The top four panels from left to right plot, respectively, the price-dividend ratio, conditional stock return volatility, conditional expected stock return, and agent A's stock holdings for different values of the fundamental D_t. Clearly, the impact of circuit breaker on price-dividend ratio, return volatility and expected return weakens as t becomes closer to T. For example, at $t = 0.25$, the price-dividend ratio with the circuit breaker can be as much as 1.2% lower than the level without the circuit breaker, and the conditional return volatility peaks at 6%. In contrast, at $t = 0.75$, the gap in price-dividend ratio is at most 0.3%, and the peak return volatility is 4.5%. The stock holdings for the two different times, on the other hand, seem quite similar, as shown in the two right panels. The above "time-of-the-day" effect in price behavior mostly reflects the fact that the impact of the circuit breaker is larger when there is more disagreement on the terminal payoff, which increases with the remaining horizon.

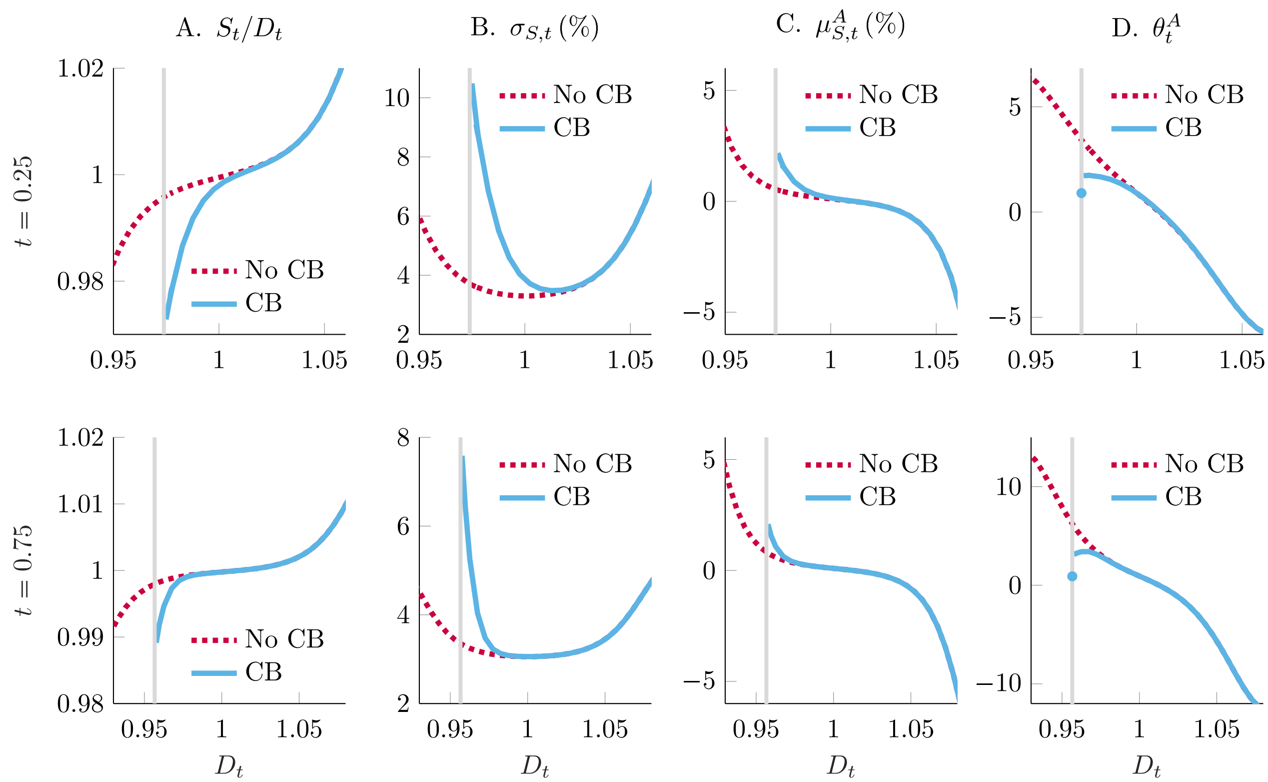

Figure OA.3: Time-of-the-day effect in the case of time-varying disagreements. Blue solid lines are for the case with circuit breaker. Red dotted lines are for the case without circuit breaker. The variables of interest are conditional on t and D_t. The parameters are $\mu = 0.04\%$, $\sigma = \nu = 3\%$, $\kappa = 0$, $\delta_0 = \bar{\delta} = 0$, $\omega = 0.9$, and $\alpha = 5\%$.

Comparing the two left panels for $t = 0.25$ and $t = 0.75$, respectively, we also observe that the circuit breaker threshold $\underline{D}_t$ becomes lower as t increases. That is, the dividend needs to drop more to trigger the circuit breaker later in the day. This is because the price-dividend ratio for any given D_t becomes higher as t increases (since the impact of more pessimistic agent decreases). This leads to an interesting implication from the model: Everything else equal, the circuit breaker is more likely to be triggered earlier in the day.

It is interesting to note that out of the four days in March of 2020 when the circuit breaker was triggered for the U.S. stock market, for three times it occurred within minutes of the market open. Only for one time (March 18, 2022), it occurred in the afternoon.

OA.5 Circuit Breaker vs. Pre-scheduled Trading Halt

Like the price-based circuit breaker, pre-scheduled trading halts, such as daily market closures considered by Slezak (1994) and Hong and Wang (2000), will also prevent investors from rebalancing their portfolios for an extended period of time. However, the implications of such pre-scheduled trading halts on trading behavior and price dynamics are quite different from those of circuit breakers. The key difference is that, in the case of a circuit breaker, the trigger of trading halt endogenously depends on the underlying fundamental (D_t). A negative shock to the fundamental not only reduces the price-dividend ratio through its impact on the wealth distribution (as in the case without the circuit breaker), but also drives the price-dividend ratio closer to the level based on the pessimist's belief by moving the market closer to the trading halt threshold.

This second effect is absent in the case of a pre-scheduled trading halt. As t approaches the pre-scheduled time of market closure $T' < T$, the price-dividend ratio converges to the pessimist valuation for all levels of dividend (which only occurs when D_t approaches $\underline{D}_t$ in the case with circuit breaker). Away from $T' < T$, the price level is lower for all levels of dividend due to the expectation of trading halt, but there is no additional sensitivity of the price-dividend ratio to fundamental shocks, hence no significant volatility amplification.

To illustrate these differences, Figure OA.4 plots the price-dividend ratio, the conditional return volatility, and agent A's portfolio holding in the case when the market is scheduled to close at $T' = 0.5$ and remains closed until $T = 1$ (red dotted lines). We then compare these results against the case with a 5% circuit breaker (blue solid lines) as well as the case without any trading halts (gray dotted lines).

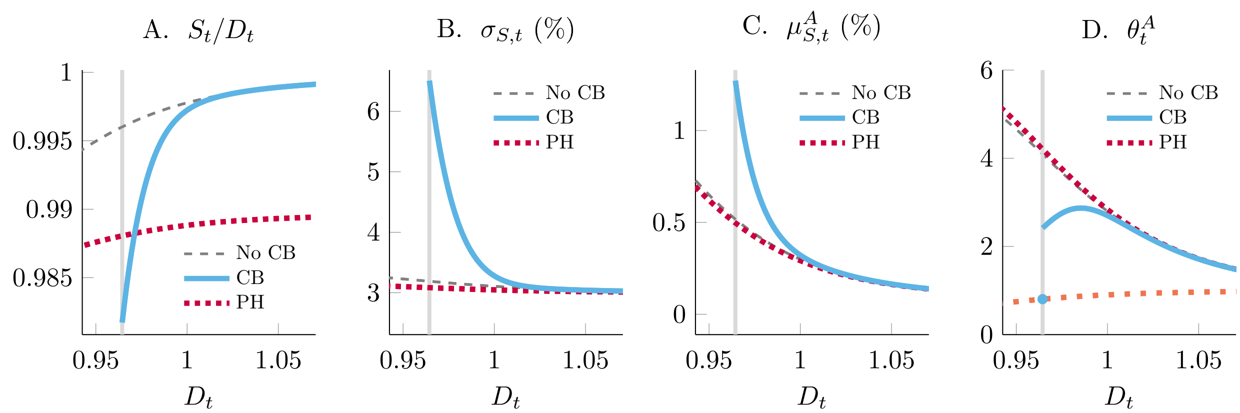

Figure OA.4: Circuit breaker vs. pre-scheduled trading halt (the case of constant disagreement). Blue solid lines are for the case with circuit breaker. Red dotted lines are for the case with pre-scheduled trading halt at $T' = 0.5$. Gray dashed lines are for the case without trading halts. The gray vertical bars denote the circuit breaker threshold $\underline{D}(t)$. The orange dotted line in Panel D denotes θ^A_t for $t = 0.49$. The parameters are $\mu = 0.04\%$, $\sigma = 3\%$, $\delta = -2\%$, $\omega = 0.9$, and $\alpha = 5\%$.

The price-dividend ratio has the most sensitivity to changes in dividend in the case of a circuit breaker; consequently, the conditional return volatility is the highest in that case. Interestingly, the conditional return volatility is the lowest with the pre-scheduled trading halt, and it almost does not change with the dividend. Moreover, unlike in the circuit breaker case, there is no preemptive deleveraging with pre-scheduled trading halt – agent A continues to take levered positions in the stock market as t approaches T', and only delevers at the instant of market closure (see the purple dotted line – agent A's stock holding at $t = 0.49$ – and the red dotted line – stock holding at $t = 0.5$ – in the bottom right panel).

OA.6 Robustness of Empirical Results

Table OA.1 reports the summary statistics of the main variables. The average 1-minute return (normalized to daily) during regular trading hours is slightly negative, consistent with earlier findings that the average intraday return for U.S. equity futures are slightly negative while the average overnight return is positive (see for example, Boyarchenko, Larsen, and Whelan, 2020). The average 1-minute volatility (normalized to daily scale) is 0.66%. The median DTCB is about 7.05 %; as a reference, DTCB would be 7% if the market opens at the same level as the closing price from the previous day. The median leverage factor is around 1.

Table OA.1: Summary statistics. Return: 1-minute E-mini return (normalized to daily); Volatility: 1-minute two scales realized volatility of S&P 500 E-mini futures (normalized to daily); Skewness: 1-minute realized skewness based on second-level E-mini return; Abnormal Volume: 1-minute volume normalized by average min-level volume in the previous 6.5 trading hours minus 1. DTCB: (minimum E-mini price over a minute – active CB threshold)/volume-weighted average E-mini price for the last 30 seconds before close from the previous trading day. Sample period: 2013/05/01 – 2020/12/31. The total number of observations is 686,697.

	Mean	STD	$Q_{1\%}$	$Q_{25\%}$	$Q_{50\%}$	$Q_{75\%}$	$Q_{99\%}$
Return (%)	−0.051	16.195	−46.101	−5.869	0.000	5.810	45.716
Volatility (%)	0.655	0.621	0.026	0.360	0.487	0.708	3.489
Skewness	0.001	0.514	−1.414	−0.229	0.000	0.232	1.423
Abnormal volume	0.045	1.281	−0.959	−0.591	−0.279	0.258	4.680
DTCB (%)	7.046	0.808	4.468	6.731	7.052	7.413	9.117
Lev	1.005	0.026	0.919	0.995	1.008	1.019	1.058

Given that both Lev and DTCB rely on certain past reference prices, together with the current price level, they will be correlated. However, the fact that Lev is with smoothing and DTCB is not makes the correlation between them rather small (0.194). Table OA.2 below reports the correlation matrix for the main empirical variables:

Table OA.2: Correlation Matrix for Main Variables

	Skewness	Return (%)	Abnormal Volume	DTCB	Lev	QLev
Volatility (%)	-0.003	0.004	0.141	-0.140	-0.456	0.483
Skewness		-0.009	-0.013	0.016	-0.005	0.006
Return (%)			-0.001	-0.009	-0.002	-0.002
Abnormal Volume				-0.095	-0.023	0.013
DTCB					0.194	-0.026
Lev						-0.492

In the rest of this section, we present several robustness checks for our empirical results.

The four circuit breaker triggering days in March 2020 gave us a rare window to observe the market dynamics when the triggering event is imminent. However, without these four days, our empirical results remain the same qualitatively. In Figure OA.5, we present the regression results of (40) when these four days are excluded from our sample. The number of minute-level observations with DTCB below 2% is 118, around 50% less than the observations when the 4 days are included.

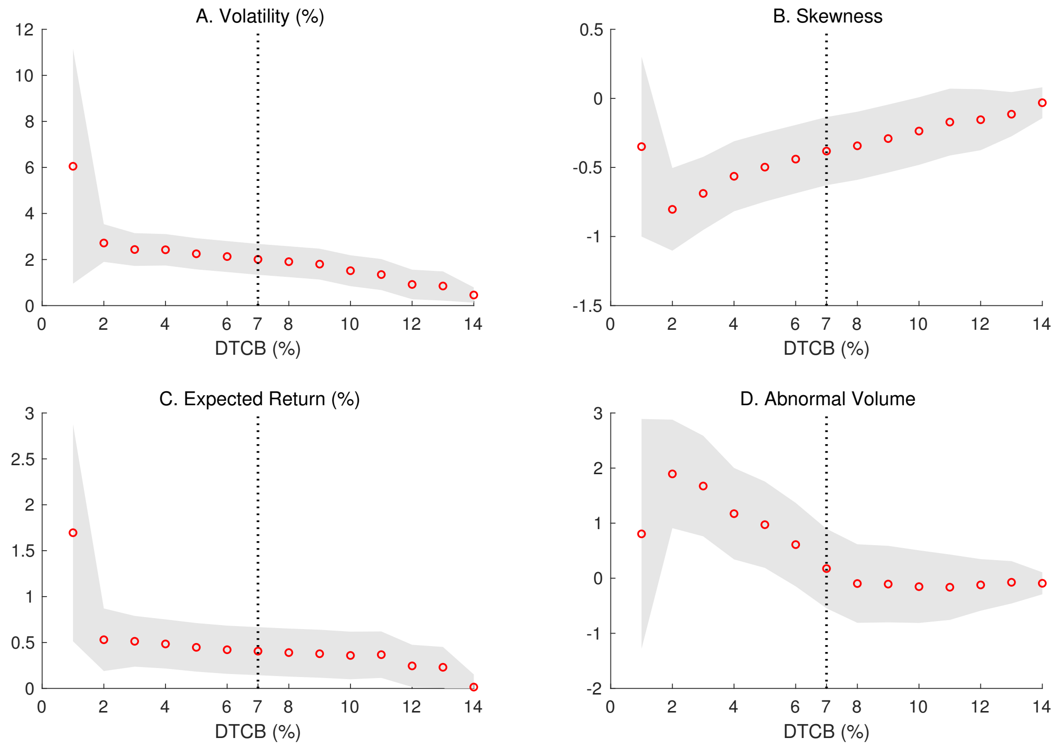

Figure OA.5: This figure shows the β regression coefficients of (40) for volatility (two scales realized volatility), skewness, expected return, and abnormal volume when 4 circuit breaker trigger days in March 2020 are excluded. $\text{Bin}_i = ((i-1)\%, i\%]$ for $i = 1, \ldots, 14$. The red dots indicate β coefficients and grey lines represent two standard deviation around.

In Figure OA.6, we report the volatility regression results when the minute-level volatility is calculated as the 1-minute integrated realized volatility of second-level returns or two scales realized volatility with 10-sec slow time scale. These results are consistent with the top left panel in Figure 6.

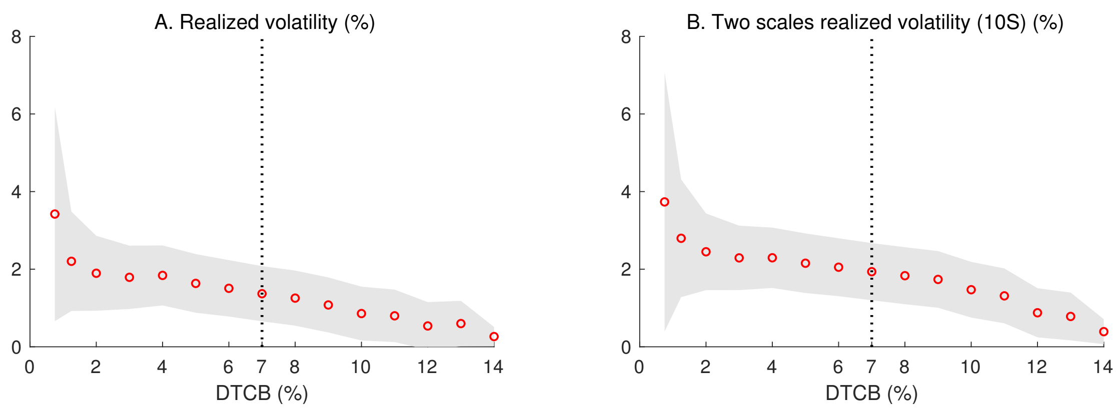

Figure OA.6: This figure shows the β regression coefficients of (40) for 1-minute integrated realized volatility of second-level returns and two scales realized volatility with 10 seconds sub-sampling. $\text{Bin}_1 = (0\%, 0.75\%]$, $\text{Bin}_2 = (0.75\%, 1.25\%]$, $\text{Bin}_3 = (1.25\%, 2\%]$, and $\text{Bin}_i = ((i-1)\%, i\%]$ for $i = 4, \ldots, 14$. The red dots indicate β coefficients and gray lines represent two standard deviation around.