Stock Explosiveness and Silent "Squeezes"

Updated Regularly: Latest Version

Yury Olshanskiy[1]

February 6, 2024

Abstract

This paper investigates abnormal behavior in individual stocks using two decades of high-frequency U.S. stock market data. It identifies hundreds of thousands of short episodes where stocks exhibit "explosive" behavior, deviating from the unit-root null hypothesis. These phenomena span multiple days, differ from typical return movements, and affect a wide range of stocks, including liquid and large-cap stocks. Explosive episodes account for a considerable portion of stocks' idiosyncratic variance. These are transitional episodes with partial reversal, providing predictable and tradable returns, setting them apart from large overnight and high-frequency jumps. I analyze stocks and their susceptibility to explosive behavior in connection with aggregate market fluctuations. While downward explosions tend to cluster among stocks and are more pro-cyclical, upward explosions appear as an idiosyncratic phenomenon. Explosive episodes involve significant buying and selling pressure along with trading volume. To explain explosive price movements, the paper introduces a model involving inelastic buyers, insiders, and competitive sellers. It emphasizes the role of explosions in the price discovery process and addresses the observed reversal. The frequency, severity, and reversal of explosiveness are explained by the expected size of inelastic demands, the knowledge possessed by a representative insider, and the frequency of seeing both in the market. Using short interest dissemination dates, empirical tests validate the model's predictions, indicating a higher likelihood of explosive behavior in stocks with substantial reported short interest.

[1] Yury Olshanskiy: MIT Sloan School of Managements, email: ols@mit.edu. website: www.ols-y.com

I am grateful to my dissertation committee Hui Chen, Leonid Kogan, and Jiang Wang for their continuous support. I thank Allison Cole, Maryam Farboodi, Joanne Im, Pierre Jaffard, Tong Liu, Debbie Lucas, Egor Matveev, Anya Nakhmurina, Robert C. Merton, Lira Mota, Jonathan Parker, Anton Petukhov, Kirill Rudov, Antoinette Schoar, Lawrence Schmidt, Justin Rand Scott, Roman Sigalov, Tim de Silva, Albert Shin, Markus Schwedeler, Maria Turina, David Thesmar, Adrien Verdelhan, Emil Verner, Haoxiang Zhu, Jiaheng Yu, and seminar participants at MIT for their thoughtful comments and discussions. This research was supported with Stone Finance PhD Fund and Mark Kritzam and Elizabeth Gorman Research Fund.

1 Introduction

Motivation

The primary focus of the finance literature centers around two key aspects of asset price dynamics: risk and average return. The vast body of research studies anomalies by examining average returns and the associated risk compensation (e.g., Fama and French (1992), Lakonishok et al. (1994)). Another significant strand of literature conducts event studies, analyzing pre-trends and post-trends while focusing on average returns and volatility surrounding events (e.g., Shleifer (1986), Bernard and Thomas (1989), Bernanke and Kuttner (2005)). In these lines of research, models are proposed that primarily rely on and target fit the aggregated statistics, such as total returns, variance-covariance structures, and various other mainly "static" aggregate measures of returns, including higher moments. (e.g., Merton (1980), French et al. (1987), Harvey and Siddique (2000)). These realized statistics typically aggregate the objects under study and serve as inputs for subsequent regression analyses. However, the actual "transitional dynamics," which emphasizes the specific *price path* of the assets, usually does not play a significant role in these studies.

This paper aims to provide a novel contribution by highlighting the importance of considering the transitional dynamics, in addition to risk and average return, as it contains unique and valuable information. The research is a *broad study* of the *explosiveness* in individual stock prices over 20 years of observation. In this context, explosiveness refers to transitional price dynamics, which typically persist from half a trading day to multiple days or even weeks. It is characterized by accelerated price changes, where past prices appear to predict future price movements, creating persistent *short-living* momentum. This transitional dynamics are unlikely to be consistent with the Ito semimartingale dynamics and, by extension, with most asset pricing models. Yet it emerges as an exceptionally prevalent phenomenon, being an essential part of the price discovery process in the market.

More formally, the term "explosiveness," borrowed from the econometric literature (e.g., Diba and Grossman (1988), Phillips et al. (2011)) where it denotes processes outside the unit-root circle, is employed in this paper to describe stock price dynamics that, *ex-post*, *statistically* inconsistent with the independent and identically distributed (i.i.d.) types of price dynamics over a certain period, as represented by the equation:[1]

$$\Delta \log Price_t = \alpha(\log Price_{t-1} - \bar{p}) + \varepsilon_t,$$

with $\alpha > 0$. In essence, the paper investigates the transitional dynamics that features a *stochastic* form of *convexity* in price path. Here, "convexity" pertains to the *persistent* increase (or decrease) in average returns

[1] Section 3 intoduces the concept more formally.

within the price adjustment window.

A prime illustration of an explosive episode in individual stocks could be the GameStop (GME) price dynamics in January 2021. See Figure 1. The red line drawn over the GME's log-price stamps emphasizes how the price transitioned to its pick values. Despite all volatility, the log price path exhibits a distinct convex pattern, showcasing the growth in average returns as prices increase.

While the extremeness and absolute magnitudes of the price dynamics exhibited by GME might seem exaggerated, it is essential to recognize that this explosiveness is not an isolated case. In fact, explosive price movements are far more common in the data than expected under the assumption of an i.i.d. process or within the framework of standard asset pricing jump-diffusion models for returns. According to the procedure developed in the paper, the detection of the episodes occurs on approximately 2.4% to 5.2% of trading days for individual stocks. It is detected up to 6-7 times more frequently than one would anticipate under standard models.

However, it is essential to note that, unlike the GME case, many of these episodes do not resemble "bubble"-like events; rather, they exhibit price discovery dynamics. The paper documents the existence of a partial reversal following the explosive dynamics. Furthermore, this accelerated price movement can occur in both directions, either as an increase or a decrease in prices. To illustrate this concept, Figure 2 showcases several other typical examples of explosive episodes observed in the data and detected using the implemented mechanism.

The observation of unorthodox price dynamics across different types of stocks serves as a strong motivator for further exploration of the underlying market mechanisms and economics that drive it. While we have a well-established understanding of how jumps and volatility arise from the incorporation of new information, the phenomenon of explosive price discovery extends beyond the conventional models' scope. The absence of complete price reversals implies that explanations invoking "irrationally exuberant" behavior by market participants, as suggested by Miller (1977), may be incomplete. As such, explosiveness potentially conveys valuable information about the market's underlying primitives embedded within the transitional dynamics. Consequently, this paper is driven to propose and test a mechanism capable of generating explosive episodes and their associated effects, including partial reversals. The primary objective of this paper is to accomplish this while employing realistic components that align with observed data.

Findings

This paper is divided into two main parts: the empirical part and the theoretical part. The empirical part involves constructing and analyzing the explosiveness measure, as well as testing the theoretical mechanism,

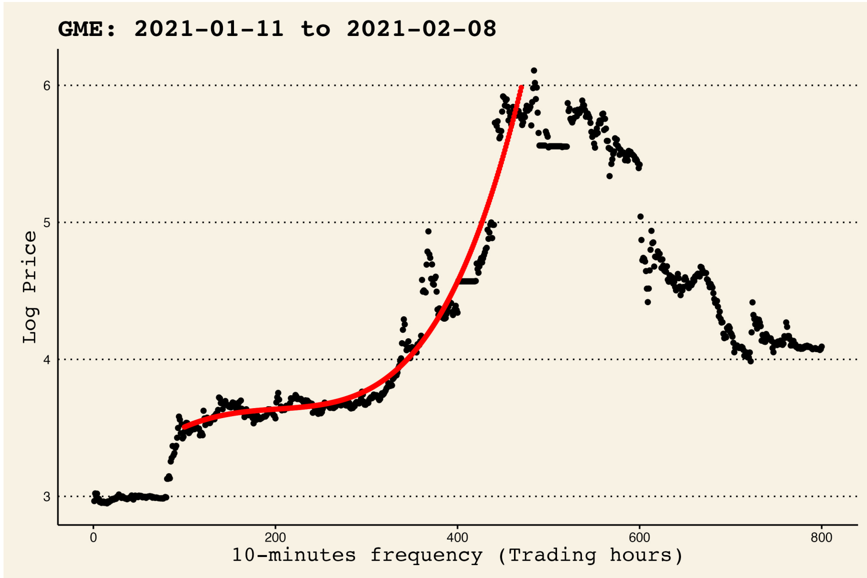

Figure 1: Game Stop, January 2021

relying on the measure as the primary dependent variable. The theoretical part comprises two modeling steps. In the first step, I demonstrate a mechanism that could generate explosiveness, and in the second step, I develop this mechanism to the point where it can generate price discovery.

The paper begins by providing a mathematical description of explosiveness and a brief explanation of the supremum Augmented Dickey-Fuller (SADF) procedure, initially introduced by Phillips et al. (2011), and adapted to detect explosive episodes in the high-frequency data of individual stocks. By applying this adapted procedure to 20 years of data in the U.S. common shares universe, encompassing over 22 million stock-date observations, the paper uncovers hundreds of thousands of explosive episodes in individual stocks. To my knowledge, this is the first paper that systematically analyzes explosiveness or any other high-frequency irregularities on such a large scale.

The precise count of explosive episodes varies based on the estimation method's assumptions. However, at a 1% significance level in the detection test, I identified six to seven times more explosive cases than would be expected under the null hypothesis with an auto-regressive process for returns. These explosive episodes can be classified into two distinct types: explosions up and down, depending on the price change direction. These two classes are almost evenly divided, with slightly more frequent cases of explosive up. The paper provides an overview of the basic average characteristics of these explosive episodes, revealing that the average magnitude of change prior to detection is approximately 10%.

To gain a deeper understanding of the potential drivers of these explosive episodes, I analyze explosiveness in the overall market and double-sorted portfolios that form the basis of standard long-short strategies.

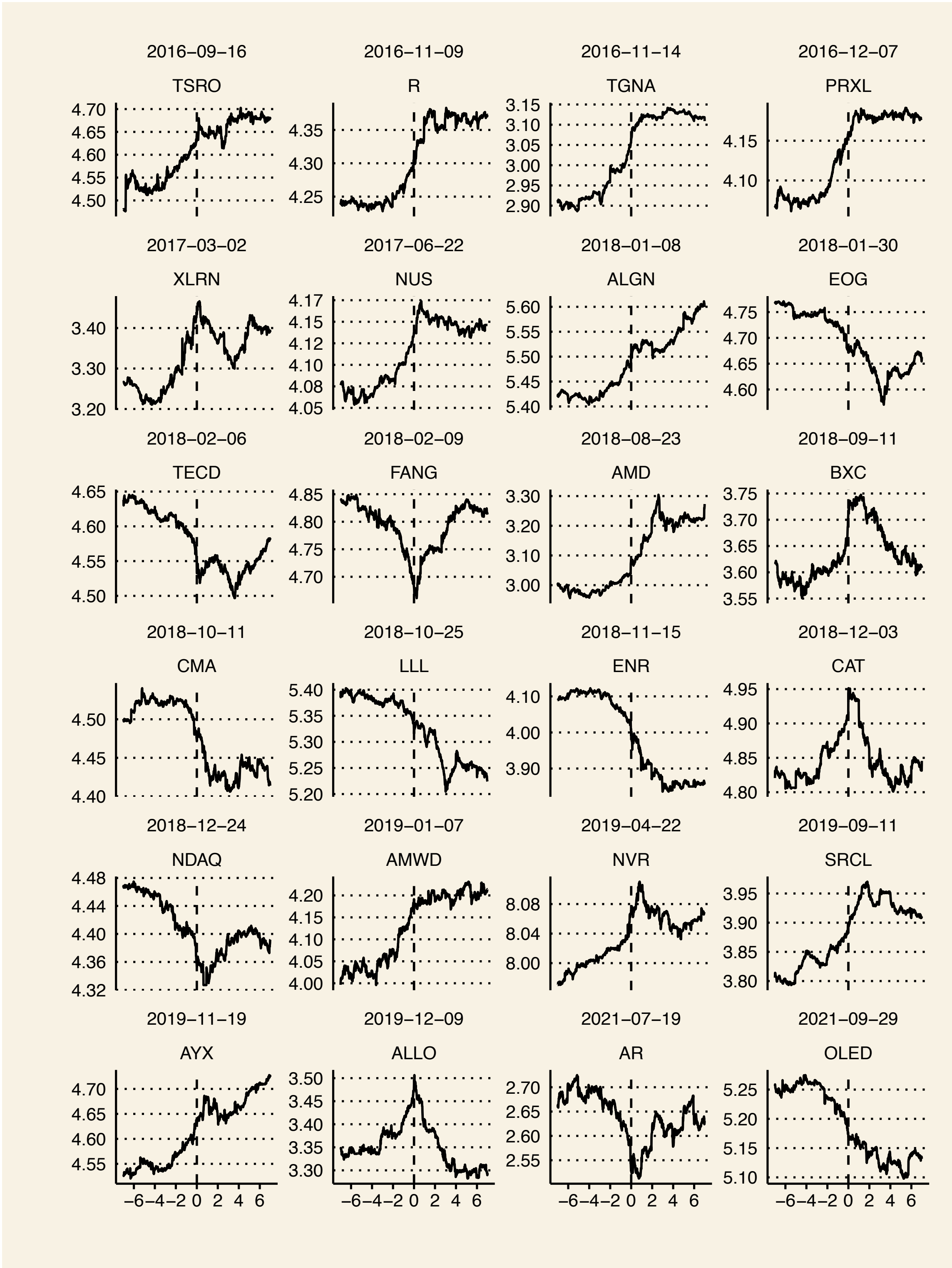

Figure 2: Examples of detected explosions

This figure provides visual examples of various detected explosive episodes within a window spanning 7 days before and 7 days after the detection timestamp. It displays log-price data for the respective stocks, with the tickers and detection dates indicated above each plot. The examples include explosive episodes with subsequent reversal, with partial reversal, with no reversal, and with continuing drift afterwards.

The results reveal that both the market and these portfolios exhibit a level of explosiveness that is lower than what individual stocks exhibit but still significantly higher than the null hypothesis predicts. Notably, the market itself is approximately four times more likely to display explosiveness compared to what the null hypothesis suggests. While market explosions are associated with individual stock explosions, they are insufficient to fully explain the emergence of explosive episodes in individual stocks, highlighting their unique, idiosyncratic nature. For instance, only 5.3% of detected upward explosive episodes in individual stocks coincide with dates featuring detected market explosions.

Interestingly, only explosiveness down is more frequent for the set of portfolios, while explosiveness up is detected consistent with the existence of false positive detections. Explosiveness down is much more clustered and indicates potential differences between these phenomena for individual stocks. Explosion down is a less idiosyncratic phenomenon. This hypothesis gains further confirmation when I perform risk adjustment for the raw stock returns and analyze the explosiveness of factors compared to the explosiveness of individual stocks.

The paper proceeds to investigate other properties of the explosive episodes. Firstly, it reveals that systematic reversals accompany explosive episodes following the moment of detection. These reversals, ranging from 10% to 15% of the change before detection, cannot be explained by bid-ask spreads and suggest the potential for a tradable strategy. In the sample of liquid stocks, it implies a persistent *tradable* daily alpha ranging from 0.6% to 1.2%.

Moreover, portfolios constructed based on the stocks following the detection of explosiveness generate similar alpha that cannot be explained by standard Fama-French factors, along with momentum and short-term reversal factors. Notably, in the same sample of liquid stocks, jumps of similar magnitudes are not followed by tradable reversals, emphasizing the role of the relatively smooth transition in the appearance of reversals.

Meanwhile, most explosive transitions persist in the following weeks, underscoring the role of explosion as a price discovery mechanism. The significance of explosiveness in the price dynamics is further confirmed by examining its contribution to the idiosyncratic volatility of individual stocks. On average, a stock experiences approximately 20.8% of its idiosyncratic daily variance on just 2.7% of explosive dates. This percentage, disproportionate to the number of daily events, rises to 35.2% when considering the days around the explosion detection. Despite a gradual decrease in the likelihood of explosiveness over the twenty years of observations, explosive episodes continue to play an essential role in volatility, particularly the explosive up episodes.

Finally, I examine explosiveness in relation to a comprehensive array of firm characteristics, utilizing variables commonly employed in cross-sectional asset pricing. Explosiveness, whether in an upward or downward direction, is consistently and visibly evident across diverse dimensions of firm fundamentals and past performance measures. I document both close-to-linear and non-monotonic relationships between firm characteristics and explosiveness. Through analyses utilizing non-penalized and penalized logistic regressions, it becomes evident that some dependencies are not consistently stable and are contingent on other firm characteristics. Nevertheless, explosiveness up seems to be influenced by size, momentum, and previous volume, suggesting that smaller firms and those with extreme returns over various horizons are more prone to exhibit explosive behavior. In contrast, explosiveness down is more closely associated with only variables describing momentum, such as the 12-to-2 return and closeness to the 52-week high.

Next, I demonstrate that explosive episodes are characterized by abnormal trading volume and buying pressure. Furthermore, by comparing these characteristics with jump episodes, I highlight that explosive episodes exhibit more pronounced abnormal buying pressure rather than just increased volume. This emphasizes the significance of directional trading typically associated with these events. Additionally, other liquidity measures based on bid-ask spread do not deteriorate to the same extent as they do for jump episodes. When measuring the time around explosive episodes in terms of trading volume, I demonstrate that these episodes continue to exhibit explosiveness in new time frames.

The paper transitions to proposing a model framework that can generate explosive episodes. First, I introduce a model with a minimal set of assumptions inspired by short-squeeze episodes. It includes an expected inelastic demand from an *impatient* inelastic buyer of an *unknown size*, while competitive sellers possess a limited amount of the asset. To simplify the model, it considers a two-stage framework in which sellers choose the price to sell, and buyers ultimately purchase the required amount of the asset at the best available price.

The model exhibits an asymmetric equilibrium where sellers propose different prices, and these prices increase explosively across the mass of sellers. This explosive price growth occurs because, to support the equilibrium, sellers submitting higher prices demand compensation for selling "later" with lower probability. As long as the size of the inelastic block order is distributed so that the additional buy amount is less likely than the previous volume, the price will grow with acceleration. Implementing this mechanism in an environment with a fixed amount of trading volume per unit of time would automatically lead to explosive price changes over time.

The model reflects the concept of market timing, where selling at a more favorable price comes with a higher risk of not successfully executing the trade. While this mechanism would give rise to temporary non-linear price impacts in the market, it does not align with the description of an explosion as part of the price discovery process. This motivates incorporating informed traders into the framework. I propose a dynamic model with unknown demand from either an insider or an impatient block buyer. Sellers dynamically choose when to sell their endowment, updating their beliefs about the type of buyer in the market.

Sellers with a remaining endowment of the asset face a decision whether to sell immediately or wait. The tradeoff, in comparison to the first model, is not just between selling and not selling but also involves dynamically changing adverse selection. Delaying the sale provides sellers with additional information to update their Bayesian probabilities about the buyer's identity and the potential loss to the buyer, conditional on selling. With these mechanisms, explosions represent not only temporary price impacts from inelastic liquidity traders but also the inflow of insider information into prices.

The model is parameterized by the *ex-ante* probability of insiders, the distributions of insiders' knowledge, the distribution of asset inelastic demand for an impatient buyer, and the mass and endowment of competitive sellers. It has a unique continuous solution over a wide range of parameters, effectively only restricted by trade conditions with the flavor of No-Trade Theorems (Milgrom and Stokey (1982), Tirole (1982)). If insiders possess too much risk for sellers, there is no trade in the model. I inspect a closed-form solution of the model under the assumption of exponential distributions for the inelastic demand size and the knowledge possessed by insiders, providing a tractable comparative static analysis.

The model highlights essential ingredients for the appearance of explosive episodes and their role in price discovery. First, inelastic demand in the market allows sellers to receive compensation for trading against insiders. Second, the limited seller supply of the asset, combined with a distribution of inelastic demand that allows extreme values, enables sellers to engage in a timing game. Third, the balance between the frequency of insiders and their average knowledge dictates the curvature of the explosion and the size of the relative reversal of the initial price change.

In the limiting case, I demonstrate that if informed traders mainly populate the model, the trade is either impossible or features an immediate jump and a linear price impact following it. This finding is intriguing as it highlights a different modeling structure for *linear* price impact compared to a substantial portion of market microstructure models, where price impact is typically defined by the strategy of hiding trading actions among noise traders (e.g., Kyle (1985)). As the share of inelastic buyers increases, the curvature of the potential price path increases, with the maximum curvature occurring when the number of insiders approaches zero. In this case, trading behavior resembles the first model described.

While the models discussed primarily focus on explaining explosive episodes in the upward direction, they can be adapted to explain downward explosions by considering the available endowment as cash for buying. In practice, raising liquidity is often easier than trying to borrow additional stocks on short notice during a market explosion. Therefore, the main focus of further testing the mechanism remains on explaining upward explosions.

The models generate an important prediction regarding the timing of explosive episodes. They suggest that explosions are more likely to occur in an environment with a higher expected inelastic demand and a lower probability of insiders. To test this prediction, I conduct a study of explosiveness in stock prices following the dissemination of short interest data. The rationale behind this test is that heavily shorted stocks are more likely to have inelastic demand from short sellers looking to cover their positions, and they are less likely to have insiders who may have taken positions in the short interest. Given the previously established predictive power of downward price drift following the dissemination of short interest data, a relevant test is to search for upward explosive episodes as a reaction to this dissemination.

To verify the connection between short interest and explosiveness up, I employ two measures: the short interest ratio (SIR) and partialled out SIR, which captures the abnormal part of SIR compared to similar industries and stocks with corresponding firm characteristics. Initially, utilizing these measures, I corroborate previous findings regarding the informational content of SIR in my sample. After the dissemination of the data, heavily shorted stocks exhibit significantly negative returns, ranging from -7% to -9% in annualized terms over the subsequent 10-day intervals.

Despite the initial expectation of a decrease in explosiveness, there is a statistically significant 10-20% relative increase compared to the default detection rate. These results are obtained using fixed-effect regressions. The regression specification accounts for specific dates, firms, and the standard firm characteristics often considered in short interest and event study analysis to mitigate the influence of other potential explanatory factors for the variance in explosiveness. The findings maintain their robustness across various specifications, including those involving risk-adjusted prices for identifying explosive episodes.

To scrutinize the informational content following the dissemination more closely, I perform an event-study-type analysis focusing on explosiveness immediately before and after the data's release. This approach enables a comparison of identical stocks with essentially the same firm characteristics. By employing analogous fixed-effect regressions to explain the change in explosiveness through the alteration in *disclosed* SIR, we discover that explosiveness does react to the *information* regarding (past) short interest. Furthermore, the indicator of being in the top decile by SIR change emerges as the sole significant explanatory variable, highlighting the significance of the informational content in influencing the likelihood of an explosive event upward.

As an additional test, I examine other static measures following the dissemination of short interest. In contrast to explosiveness, standard risk measures like within-period volatility, empirical quantiles of return distributions, and higher-order moments show minimal economic changes that could explain the increase in explosiveness. The primary alteration observed is the direction of further movement, which tends to be negative.

The final section of the paper discusses the robustness of the explosiveness detection mechanism within the context of high-frequency stock data. The SADF test was initially developed for different data using longer and less frequent time series. Although, in theory, under mild assumptions, its distribution should converge to the theoretical function derived from Brownian motion integrals, it is an open question whether it reasonably suits the highly fat-tailed distribution of individual stock returns in finite samples.

First, I demonstrate that simulating explosiveness under static returns while fitting only the first two moments throughout estimation results in a slightly lower detection rate than the selected significance level. I further bootstrap the actual returns from the high-frequency sample and show that the fat-tailed distribution of returns increases the false detection rate. Then, I estimate the auto-correlation of returns from the sample and simulate data based on randomly selected coefficients. The false detection rate increases to its maximum values, which I can produce, reaching around 2.7% at the 1% significance level. This emphasizes that the SADF procedure has natural limits and may fail to distinguish periods of high autocorrelation in auto-regression. Despite this limitation, the detection rate of explosive episodes remains significantly higher than the one suggested by the simulations.

Additional robustness tests are provided, including bootstrap simulations, stochastic volatility simulations based on the Heston model, and simulations with scheduled overnight returns. None of these specifications produce a false positive detection rate that describes the frequency at which we detect explosive episodes in real data. This underscores the fundamental role of accelerated time-series dynamics in generating explosive episodes captured by the SADF-based approach.

Related Literature and Contribution

This study builds upon and contributes to the existing literature on empirical and theoretical research in asset pricing, market microstructure, and financial econometrics. To maintain the structure of the section, I will classify the related literature into different sections that correspond to each of these areas, namely the High-Frequency Literature, Bubble Literature, Insider Trading Literature, Short Interest Literature, and Other Literature.

High Frequency Literature.[2] This paper contributes to the developing literature on high-frequency asset pricing (e.g., Bollerslev et al. (2016), Li et al. (2017), Bollerslev (2022)). Analyzing TAQ data across

[2] The high frequency market microstructure literature will be discussed in Insider Trading Literature.

the universe of U.S. common share stocks, it adds to our understanding of stock market behavior, identifying a new common feature at the 5-20 day horizons, referred to as explosiveness. In doing so, the paper draws upon the extensive literature on processing high-frequency data (e.g., Andersen et al. (2001), Barndorff-Nielsen et al. (2009)) and estimating high-frequency factors and risk exposures (e.g., Aït-Sahalia et al. (2020)).

The revealed transitory dynamics deviate from the conventional jump-diffusion model (e.g., Merton (1980)) and its widely studied modifications (e.g., Andersen et al. (2007), Todorov and Bollerslev (2010), Bollerslev and Todorov (2011a), Bollerslev et al. (2013)), which are primarily used to explain price movements of individual stocks and underlying factors. The section on the detection of explosions contributes to the emerging literature on studying violations of the Ito process in high-frequency data (e.g., Christensen et al. (2022)), as well as the more general literature on violations of the i.i.d. assumption (e.g., Jacod et al. (2017), Li et al. (2020), Li and Linton (2022)).

The findings of this paper are consistent with both hypotheses being statistically rejected in a variety of stock types. However, several key differences exist between this paper and the literature on violations. First, this paper considers lower frequency (5-10 minutes vs. 1 second) and longer horizons (10 days vs. intraday). Second, the paper includes a diverse set of stocks, many of which are less liquid than the assets typically considered in the literature. Third, most of the existing literature on Ito-violation does not delve into the market mechanisms behind them, leaving room for interpretation. In contrast, this paper primarily links these violations to the market environment. Here, it is closer to literature that studies specific events such as flash crashes (e.g., Easley et al. (2011), Kirilenko et al. (2017)). The mechanism of dried liquidity, which is implicitly the underlying mechanism in this paper, may share some similarities with these events. However, there are notable differences. This paper primarily focuses on understanding price discovery, while events like flash crashes involved temporary deviations that reversed within 10-60 minutes.

The dynamics observed in this paper bear some similarity to the concept of “gradual jumps” proposed by Barndorff-Nielsen et al. (2009). They suggest that what is observed as a jump at lower frequencies might be a series of jumps when viewed at a higher frequency[3], gradually zooming in on the interval. Applying the concept of gradual jumps to our lower frequency might lead to the ex-post detection of an explosion. Mathematically separating these two phenomena could be challenging. Conceptually, the gradual jumps could also be considered part of the explosive dynamic. However, most explosive episodes in this paper continue over a more extended series of times, requiring an alternative approach for modeling.

The concept of "drift burst," as discussed in Christensen et al. (2022), and the more general concept

[3]The 5-minute interval considered in the paper is the low-frequency interval in Barndorff-Nielsen et al. (2009).

of "persistent noise," as presented in Andersen et al. (2021), are somewhat related to the idea of explosive episodes. These concepts revolve around the idea that there are times when the drift term in asset prices is not negligible compared to the volatility terms. They recognize that there are episodes where drift can change significantly, challenging the assumptions of standard continuous-time arbitrage-free price processes. While these concepts share some common ground with explosive episodes, the structure and focus of explosions are different. Explosions aim to capture the predictive power inherent in the current levels of prices and do not assume that drift significantly exceeds the volatility.

Bubble Literature. While the main narrative of this paper is to demonstrate that explosive episodes are typically not "mini-bubbles," it shares close relationships with studies on market mechanisms and detection methods for bubbles. The primary tool used for detecting explosive episodes, the SADF procedure introduced by Phillips et al. (2011), has its roots in the literature dedicated to identifying bubbles (Phillips et al. (2015a), Homm and Breitung (2012), Phillips et al. (2015b), Phillips and Shi (2018)). This body of work aims to locate instances of explosive behavior (Diba and Grossman (1988)) followed by a sudden burst in long macroeconomic and financial time series. These tests have practical applications in identifying bubbles in various markets, including significant stock indices such as NASDAQ and the S&P 500 (Phillips et al. (2011); Homm and Breitung (2012), Phillips and Shi (2018)), cryptocurrencies (Cheung et al. (2015); Corbet et al. (2018)), and in select studies related to housing (Coskun et al. (2020)), oil (Hau et al. (2020)), and other markets. It is worth noting that these tests have yet to be widely applied to individual stocks thus far.

The paper's motivation for employing the detection mechanism with a "double-rolling" procedure, which involves one rolling of the original SADF procedure to identify explosions within a finite window and then rolling that window, is twofold. Firstly, this approach is computationally efficient. Secondly, it successfully captures the required phenomenon of acceleration in price changes.

The modeling of explosions in this paper shares similarities with the literature on bubbles (e.g., Harrison and Kreps (1978), Barberis et al. (1998), Allen and Gale (2000)), which often focuses on the constraints on short selling and the divergence of opinion, as seen in references such as Miller (1977), Shleifer and Vishny (1997), and Abreu and Brunnermeier (2003). The presence of these factors, along with the existence of informed traders, contributes to the generation of explosiveness in this paper, providing a link between the literature and price-discovery literature.

Insider and Liquidation Trading. The extensive field of market microstructure studies the price discovery process initiated by insiders (e.g., Kyle (1985), Huberman and Stanzl (2005)) and other liquidity trading (e.g., Bertsimas and Lo (1998), Huberman and Stanzl (2005), Obizhaeva and Wang (2013)). The prevailing mechanism is built on the assumption that informed traders disguise themselves as noise traders to extract knowledge. At the same time, sellers (or buyers) are compensated for their adverse selection, as described in the seminal works Kyle (1985) and Kyle (1989). In this paper, a similar concept is employed, but insiders do not hide behind noise traders; instead, they imitate big liquidity traders, allowing sellers (market makers) to detect that they are dealing with consistent buying pressure.

The conventional Kyle model and its variations are not customized for generating an explosive pattern since they typically result in insider trades having a constant and linear price impact. Models with multiple insiders (e.g., Holden and Subrahmanyam (1992), Back et al. (2000)) and liquidity traders (e.g., Huberman and Stanzl (2005)) generally lead to a different phenomenon where information first enters the market rapidly and then continues to drift slowly. While Kyle et al. (2017) does produce positive time-series momentum within the context of trading by overconfident oligopolistic investors, their model is more suited to long-term dynamics with moderate momentum rather than for capturing explosive episodes.

The findings of the explosiveness and buying pressure association in the paper are consistent with the previous literature, which showed that price changes are predictable by order flow (e.g., Hopman (2007), Bouchaud et al. (2009), Cont et al. (2014)). More recent literature (e.g., Korajczyk and Murphy (2019), Hirschey (2021)) confirms the assumption of a regular presence of big institutional orders and the role of HFT traders that "lean with the wind" as formulated by Van Kervel and Menkveld (2019), effectively amplifying the buying (selling) pressure from institutions. Nevertheless, to the best of my knowledge, the explosive-type dynamics did not appear in the literature.

Short Interest Literature. A substantial body of literature has highlighted the informational content carried in short interest data at both the individual stock level (e.g., Asquith et al. (2005), Boehmer et al. (2008), Hong et al. (2015)) and aggregate level (e.g., Lynch et al. (2014), Rapach et al. (2016)). The primary mechanism discussed in this literature is that since shorting is costly, mainly informed traders decide to short-sell a stock, leading to heavily shorted stocks being predictors of subsequent negative returns. Conversely, stocks with light short interest predict abnormal positive returns (e.g., Boehmer et al. (2010)). In line with this hypothesis, stocks with heavy short interest have more negative returns when institutional ownership is higher (e.g., Asquith et al. (2005)) and when shorting fees are greater (e.g., Drechsler and Drechsler (2014)). However, this literature has not delved into the specific mechanisms of how price discovery occurs following short interest reporting; it primarily focuses on abnormal risk-adjusted returns.

This paper contributes new insights into the dynamics following short interest dissemination. It reveals that changes in short interest not only predict directional shifts but also the nature of the subsequent transitional dynamics. This finding aligns with the discoveries in Callen and Fang (2015) regarding changes in crash risk after short interest dissemination. The novel evidence here is that these changes also apply to the positive side of the distribution, forecasting explosiveness upward. Furthermore, the distribution of high-frequency returns remains relatively stable, underscoring the importance of time-series dynamics over alterations in static return properties.

The predictive power of changes in short interest ratios is somewhat controversial, particularly regarding subsequent negative returns (see, e.g., Boehmer et al. (2010)). However, in this paper, changes in short interest ratios prove to be a reliable predictor of subsequent explosive episodes, highlighting the informational content embedded in announcements about market fundamentals beyond firm-specific information.

Other Literature. Market phenomena like overreaction (Jegadeesh and Titman (1993), Lehmann (1990)), underreaction (Daniel et al. (1998), Barberis et al. (1998)), momentum (Jegadeesh and Titman (1995)), and reversal (Pelger (2020)) are well-documented and widely acknowledged in both market indices and individual assets. While these market phenomena typically entail particular types of time-series dynamics in prices, empirical studies often adopt a narrow approach, primarily focusing on *alpha*, which represents the excess return over a suitable risk-adjustment benchmark. This paper contributes by uncovering specific transitional dynamics underlying these phenomena.

In particular, the paper reveals that short-term reversal is more closely associated with explosive episodes rather than high-frequency jumps. Explosive episodes can be linked to under- and overreaction, although the empirical evidence suggests that these explosions are less likely to follow earnings announcements, which are a primary focus in the existing literature. This finding implies the presence of different mechanisms at play in explaining explosive price dynamics.

The paper is also related to developing literature on demand system asset pricing (e.g., Koijen and Yogo (2019), Koijen and Yogo (2020), Gabaix and Koijen (2021)) and older literature on fund flows (e.g., Lou (2012), Vayanos and Woolley (2013)) that emphasize on the role of inelasticity and impact of institutional trades for price formations. Though this literature targets modeling aggregate and longer-lasting fluctuations, while my paper generally has more micro focus, the paper's mechanism significantly relies on traders' knowledge of the presence of those significant persistent flows. Moreover, explosiveness can be used as an additional tool on top of the standardly considered returns for studying demand systems.

The remainder of the paper is organized as follows. Section 2 briefly describes the data, cleaning procedures, and notations used for the paper. Section 3 provides definitions and the method for detecting explosive episodes. It also presents a particular case study of using it. Section 4 describes explosiveness in the context of aggregate market fluctuations. Section 5 characterizes individual stock explosive episodes. It examines analyzes partial reversal, studies links to buying pressure and trading volume, describes their contribution to the idiosyncratic variance of stocks, and links explosiveness to firm characteristics. Section 6 discusses explosiveness around short-squeeze events, suggesting the underlying mechanism. Section 7 presents the model generating explosions based on the expected inelastic demand assumption. Section 8 introduces insiders into the model, generating explosive price-discovery processes with partial reversal. Section 9 tests the model's prediction around the short interest dissemination dates. Section 10 provides a robustness check employing Monte Carlo simulations. Section 11 concludes.

2 Data and Variables

This section covers the data sources and key variables. Across the empirical sections of the paper, s subscript stands for a particular stock or a portfolio of stocks, i stands for industry, d stands for a day or the first day of the considered period, and t for a specific (high-frequency) moment in time. The $\{d,0\}$ and $\{d,t\}$ subscripts stand for the first and $t+1$-th observation in the given time period that starts in the morning of day d.

TAQ

The primary source of the high-frequency data is the New York Stock Exchange Trade and Quotes (TAQ) which covers all trading dates from 2003-09-10 to 2022-12-31. I access quotes data via Wharton Research Data Services (WRDS). The cleaning procedure that I use is highly adapted to the needs of building reliable explosiveness measures. It is described in the Appendix Section A1. The stock that did not have clean prices within the window of interest was excluded from further analysis both for the estimation of explosiveness and building high-frequency factor portfolios.

The high-frequency factor portfolios are constructed daily over the window of interest to match individual stock price paths. The weight assignment and assignment to a portfolio happens yearly or monthly according to standard portfolio sorting procedures (e.g., Fama and French (2015)). Using factor portfolios, I calculate high-frequency factor returns from the beginning of every trading date. The high-frequency factor exposures of individual stocks were derived from rolling 50-day window regression based on intraday returns only. The log-price fluctuations are adjusted by high-frequency factors according to estimated betas.

Finally, I utilize the Millisecond Trade and Quote product by WRDS to obtain measures on buying and selling pressure, trading volume, and other liquidity indicators.

CRSP

The daily stock data, such as returns, prices, share outstanding, industry identifiers, and trading volume are obtained from the Center for Research on Security Prices (CRSP). The daily data is used to verify high-frequency TAQ data and make adjustments for overnight stock splits and distributions. Using the daily factor returns reported on Keneth French's website, I calculate daily factor exposures, idiosyncratic volatility, and daily jump measure in line with Kapadia and Zekhnini (2019).

Compustat, I/B/E/S, Thomson Reuters (13F)

The quarterly accounting data from Compustat is used to classify stocks into high-frequency factor portfolios and to control for firm characteristics in our empirical analysis. I use I/B/E/S to identify earning announcement dates and surprises. The quarterly institutional holdings are built based on the s12 file of Thomson Reuters (13F). The short interest data with bimonthly FINRA releases is taken from Compustat and discussed in Section 9.

3 Explosiveness

This section discusses the detection of explosive episodes in stock price data. Our goal is to identify temporary periods of observations with a particular pattern in the data, where (1) previous prices appear to predict subsequent returns positively, and (2) there is evidence of acceleration in the underlying dynamics. I start by discussing the definition of the objects suitable for detection and formally targeted by the test. Then, I provide an example of a data generating process that may appear identical to the theoretical explosion *ex-post* but is not inherently explosive, emphasizing that this paper's application is focused on the ex-post classification of transitional dynamics. Afterward, we will discuss the empirical identification of explosive episodes based on high-frequency data and provide a case study to illustrate the detection process.

3.1 Definition for empirical test

Consider a time series of stock prices $P_{s,t}$, $t = t_0, \dots, T$, defined by some data generating process (DGP), such that $(\log P_{s,t_0}, \dots, \log P_{s,T})$ is an element of L^2 space in some probability space $(\Omega, \mathcal{F}, P)$, with a filtration $\mathcal{F}_\bullet = (\mathcal{F}_t \subseteq \mathcal{F} : t \in \mathcal{T})$, $\mathcal{T} = t_0 - 1, \dots, T$. Define $\widehat{\Delta \log P_{s,t}}$ as the projection of stock log-return $\Delta \log P_{s,t} = \log(\frac{P_{s,t}}{P_{s,t-1}})$ into the space spanned by k previous returns and a constant, $\{1, \Delta \log P_{s,t-1}, \Delta \log P_{s,t-2}, \dots, \Delta \log P_{s,t-k}\}$. Then, call a period $[t_0, T]$ an **explosive episode** for a stock s if there exists $\bar{t} \in (t_0, T)$ such that the residual of projection is positively correlated to the previously observed log-price, $\log P_{s,t-1}$:

$$\mathbf{cov}_{t_0-1}\left(\log P_{s,t-1}, \Delta \log P_{s,t} - \widehat{\Delta \log P_{s,t}}\right) > 0, \qquad t = t_0 + k + 1, \dots \bar{t}.$$

If the *realized* return over the explosive episode, $\log P_{s,\bar{t}} - \log P_{s,t_0}$, is positive (negative), then I call it the **explosion up (down)**. Note that the *up/down classification* is not a characterization of DGP but rather an *ex-post* attribute of observed data points. Explosiveness refers to a specific case in which a non-stationary episode of returns occurs, where a previous high (low) stock price predicts future price movements in the same direction, even when controlling for return autocorrelation. Importantly, explosiveness does not impose restrictions on the subsequent behavior of stocks when the positive covariance ends. This allows explosiveness to represent both transitional dynamics leading to new prices and dynamics with subsequent reversals.

The simplest example of explosiveness is when the stock price $P_{s,t}$ obeys[4]

$$\Delta \log P_{s,t} = \alpha + c_{\bar{t}} \log P_{s,t-1} + \sum_{i=j}^{k} \phi_j \Delta \log P_{s,t-j} + \varepsilon_{s,t}, \quad t \in \{t_0, \dots, \bar{t}\}, \tag{1}$$

where $c_{\bar{t}} > 0$ and ε-s are uncorrelated and zero-mean error terms for some lag order k and lag parameters $\{\phi_j\}_{j=1}^{k}$. Here, explosion is a time series property that violates autoregressive model $AR(k)$ for the stock returns. For example, assuming no autocorrelation in (1), i.e., $\phi_j \equiv 0$, explosion implies expected exponential change of price in time:

$$\mathbb{E}_{t_0}[\log P_{s,t}] = -\alpha c^{-1} + (1+c)^{t-t_0} \left(\log P_{s,t_0} + \alpha c^{-1}\right)$$

If $c \log P_{s,t_0} > \alpha$, *ex-post* it is more likely to expect explosion up. If $c \log P_{s,t_0} < \alpha$ then explosion down is more likely to be observed.

3.2 Alternative DGP ex-post explosive

With the object defined above, it is important to understand that the realized data might exhibit an alternative data generating process that appears identical in hindsight to the observer. Consider the example of the DGP,

$$\Delta \log P_{s,t} = \begin{cases} c \log P_{s,t-1} + \varepsilon_{s,t}, & t < \tilde{\tau}, \\ -c \log P_{s,t-1} + \varepsilon_{s,t}, & t = \tilde{\tau} \qquad t = t_0, \dots, T \\ \varepsilon_{s,t}, & t > \tilde{\tau} \end{cases} \tag{2}$$

where $\tilde{\tau}$ represents a stopping time, which is a measurable random variable within the probability space $(\Omega, \mathcal{F}, P)$. Its support is given by $\text{supp}(\tilde{\tau}) = t_0, \dots, T$, and it satisfies the property $\tau \leqslant t \in \mathcal{F}_t$ for all $t \in \mathcal{T}$.[5] Additionally, it is specified that the conditional probability of $\tilde{\tau}$ being equal to $t'+1$ given that $\tilde{\tau} > t'$ is the same as the probability of $\tilde{\tau}$ being greater than $t'+1$ given that $\tilde{\tau} > t'$, and both are equal to 0.5.

$$Pr(\tilde{\tau} = t'+1 | \tilde{\tau} > t') = Pr(\tilde{\tau} > t'+1 | \tilde{\tau} = xt') = 0.5$$

[4]Here, assume no dividends and stock splits over the period.

[5]This property ensures that the event of τ being less than or equal to a specific time t is measurable with respect to the sigma-algebra $\mathcal{F}_t$.

In other words, before the stopping time, at each time t, there is an equal probability that the next period will either be the stopping period (where prices reverse on average) or continue its growing path. It is trivial to show that at every point in time, $\Delta \log P_{s,t}$ is a martingale by construction:

$$\mathbb{E}[\Delta \log P_{s,t+1} | \mathcal{F}_t] = 0,$$

and the true data generating process is not explosive:

$$\mathbf{cov}_{t'}\left(\Delta \log P_{s,t+1}, P_{s,t}\right) = 0. \qquad t' < t$$

Nevertheless, *ex-post* observations for the processes must be indistinguishable from the explosive process with some unknown $\bar{t}$ and

$$\Delta \log P_{s,t} = \begin{cases} c \log P_{s,t-1} + \varepsilon_{s,t}, & t < \bar{t}, \\ \varepsilon_{s,t}, & t > \bar{t} \end{cases} \qquad t = t_0, \ldots, T. \tag{3}$$

Hence, a detection mechanism aimed at capturing explosions might also capture processes with similar characteristics that are not necessarily martingale violators. The provided example emphasizes that the empirical detection mechanism used throughout this paper is designed to identify episodes with prices exhibiting a stochastic version of convexity. It does not claim to identify actual martingale violations but instead, episodes when realized prices exhibit characteristics similar to unit-root violations.

3.3 Empirical identification

To capture explosiveness in individual stocks and stock portfolios, it is necessary to identify time-series episodes where the covariance is positive. To accomplish this, I adapt a rolling window procedure from Phillips, Wu, and Yu (2011), abbreviated as PWY (2011), applied to rolling windows ranging from 5 days to monthly intervals of high-frequency stock price observations. Effectively, that employs a double rolling procedure. This method is well-suited for dealing with the temporary nature of explosive episodes. It remains robust even when the positive covariance disappears and subsequent reversals occur in the data. The procedure is described below.

The objective is to construct a daily explosiveness measure, denoted as $p^{l,f}_{d,s}$, based on TAQ high-frequency data for each stock s. The variable d stands for the first day of the considered period, l takes values from the set {5 days, 10 days, 20 days}, representing the length of the period over which explosive episodes are identified. The variable f indicates the frequency at which TAQ-based high-frequency prices are utilized. The analysis is restricted to business days and trading hours, starting from 9:40 AM[6]. For example, if f is set to 5 minutes[7] and l is set to 10 days, this would result in a total of 770 observations within that specified period. Subsequently, a set of regressions is estimated:

$$\Delta \log P_{s,t} = \alpha + c_{T'} \log P_{s,t-1} + \sum_{i=j}^{k} \phi_j \Delta \log P_{s,t-j} + \varepsilon_{s,t}, \quad t \in \{t_0, \ldots, T'\}. \tag{4}$$

for $T' \in [T_0, T]$ and pre-selected lag order k, collecting t-statistics on $c_{T'}$, known as Augmented-Dickey Fuller statistics, ADF_T'. After collecting them, I identify supremum-ADF statistic, as the maximum of the set,

$$SADF_{d,s}^{l,f} = sup_{T'} ADF_{T'}.$$

As shown in PWY (2011) under mild conditions, if the null hypothesis is true, i.e., stock returns follow an autoregression process:

$$\Delta \log P_{s,t} = \alpha + \sum_{i=j}^{k} \phi_j \Delta \log P_{s,t-j} + \varepsilon_{s,t}, \quad t \in \{t_0, \ldots, T\}, \tag{5}$$

the estimated measure has the following asymptotic invariant distribution:

$$D(SADF)_{r_0} \sim \sup_{r \in [r_0, 1]} \left\{ \frac{\frac{1}{2} r \left[W(r)^2 - r \right] - \int_0^r W(s) ds W(r)}{r^{1/2} \left\{ r \int_0^r W(s)^2 ds - \left[\int_0^r W(s) ds \right]^2 \right\}^{1/2}} \right\}. \tag{6}$$

The right-tail quantiles of the distribution can be used as critical values for testing. For example, critical values at α =1%, 5%, 10% for $r_0 = 0.2$ are 1.997, 1.421, and 1.125, respectively. These values are derived from 100,000 Monte-Carlo simulations.[8] Since the distribution in the finite sample can be affected by the lower frequency of price observations and by the fat tails of the observed stock returns, I provide a battery of Monte-Carlo simulations to identify the accuracy of the test in the environment of high-frequency stock data in Section 10.

From the perspective of a market participant who runs the rolling ADF procedure in real-time, the first instance $\widetilde{T}'$ at which $ADF_{\widetilde{T}'}$ exceeds a given critical value CV_α can be considered as the detection of an explosion:

$$DE_{d,s}^{l,f,\alpha} = \arg \inf_{\widetilde{T}'} \left(T' : ADF_{T'} \geqslant CV_\alpha \right). \tag{7}$$

[6]Note that the actual trading hours are from 9:30 AM to 4:00 PM, but we observe larger spreads and extra volatility in the mid-quote prices during the first minutes of a trading day (see e.g. Lee and Mykland (2012)) , which could potentially disrupt the results.

[7]The 5-minute frequency is the most common in volatility forecasting.

[8]I simulate the standard Brownian motions with 100,000 steps over a unit interval.

One possible definition of the end of an explosion is the first moment afterward when $ADF_{T'}$ falls below the critical value:

$$EE_{d,s}^{l,f,\alpha} = \arg\inf_{\widetilde{T}'} \left(T' > DE_{d,s}^{l,f,\alpha} : ADF_{T'} < CV_{\alpha} \right). \tag{8}$$

If an explosion is detected, its length can be calculated as

$$LE_{d,s}^{l,f,\alpha} = EE_{d,s}^{l,f,\alpha} - DE_{d,s}^{l,f,\alpha}. \tag{9}$$

In practice, when dealing with individual stocks, this definition will likely underestimate the length of an episode that a researcher might want to analyze as an explosion. This is because the volatility of returns can lead to multiple crossings of the critical value level by $ADF_{\widetilde{T}'}$.

The previous literature that utilized the SADF test primarily focused on accurately classifying specific episodes, often spanning several years and labeled as 'bubbles.' In these cases, researchers commonly employed model selection criteria to determine the optimal number of lags. However, it is worth noting that the inclusion of these selection criteria significantly increases the computational cost of SADF, providing only marginal improvements in classifying periods as either explosions or not. Therefore, in this paper, I predominantly rely on classification based on a fixed lag order[9] or BIC-selected lag order from zero to four lags. The details will be discussed in Section 5.

3.4 A Case Study of Lumen: detection of explosiveness, January 2021

Let us show how the testing procedure detects an explosion for Lumen stock in January 2021. Figure 3 shows the price dynamics of Lumen stock over a 10-business-day interval, starting in the morning of January 20, 2021. The time series consists of seven hundred and seventy observations that come with 5-minute frequency over the trading hours.[10] To tag the stock as an explosive, we do the following. First, we set the minimum window size $r_0 = 0.2$ and lag-order parameter $k = 2$ to estimate (1) $ADF_{T'}$ starting from $T' = r_0 \times 400 = 80$. The maximum ADF is achieved when $T' = 377$ that I record as the *time of explosion*, which corresponds to observation on January 26, 2021, at 3:20 pm. $ADF_{T'} = 6.58$. Comparing the value with simulated critical values from (6), I conclude that the null can be rejected almost certainly since p-value for $SADF = 6.58$ is zero[11]. Looking at the size of the estimated coefficient, $c_{T'} = 0.083$, and the initial log price, one can assess how much drift in the price of the stock changed within the explosive episode: $c \times (2.8 - 2.4)$. Assuming

[9] Throughout the paper, I primarily utilize $k = 1$ as the most conservative version, accounting for the first-order auto-covariance term. Relaxing the value of k typically leads to an increase in the detection rate, but it also appears to introduce more noise into the estimation.

[10] That is seventy-seven observations per day, that excludes 9:30 and 9:35 timestamps

[11] Since I do only a finite number of simulations, $N = 100,000$, I actually do not get simulations with higher values.

that over the initial interval the drift in the log price was close to zero, the movement of the prices shifted the drift to the level of around 3% per 10 minute, which is abnormally high for individual stocks and barely matches any risk-based explanation.

In the specific case of the Lumen stock explosion, the price experienced an almost complete reversal, returning to values around 2.45. Despite the stock still showing substantial appreciation during this period, roughly around 5%, this appreciation is significantly smaller when compared to the 40% peak it reached. It is important to note that the moment of explosion detection, denoted as $DE^{10,5,5\%}_{\text{Jan 20,LUMN}}$ and illustrated by a blue dashed line, occurred just 30 minutes before reaching the peak, which may seem visually late for this particular case. However, this delay can be attributed to the highly volatile and reversal-prone nature of stocks, making it challenging to obtain a high ADF. At this moment, the stock has appreciated by 7% since the beginning of the interval. Further, 46 periods later, at 12:20 on the following day, the end of the explosion is defined according to (9).

It is essential to emphasize that relying on right-tail ADF testing over the observed period to detect the anomaly, as observed here, would lead to failure. This is because the ADF t-statistic of -1.57 does not indicate any explosiveness (with a p-value of the test being 0.504). This phenomenon, observed from a market perspective, is relatively short-lived yet carries significant implications for a broad spectrum of investors, contributing substantially to overall market volatility. Moreover, the atypical stock dynamics pose difficulties in smoothly fitting into a standard jump-diffusion framework, even with the incorporation of stochastic volatility instruments. This paper aims to address the central question of whether such unorthodox stock price dynamics are a common occurrence, and whether it resembles the temporary price deviation seen in the case of Lumen or result in more persistent changes.

3.5 Basic Descriptives

I apply a similar procedure to the one discussed for Lumen stock for every stock on a daily basis. The collected p-values are summarized in the empirical probability distribution function (PDF) shown in Figure 5. The red dash-dotted line corresponds to the theoretical PDF of p-values under the null hypothesis. The heavy tails of the empirical PDF around the edges indicate that the null hypothesis must be rejected for a substantial portion of our observations. The right tail, with p-values close to one, suggests that the data regularly experiences a strong reversal that cannot be accounted for by autocorrelation lags in (4). Specifically, when stock prices reach high values, they systematically bounce back. As I demonstrate in the appendix, this effect is more pronounced when one does not consider the autocorrelation of returns by setting the lag order, denoted as k, to be small.

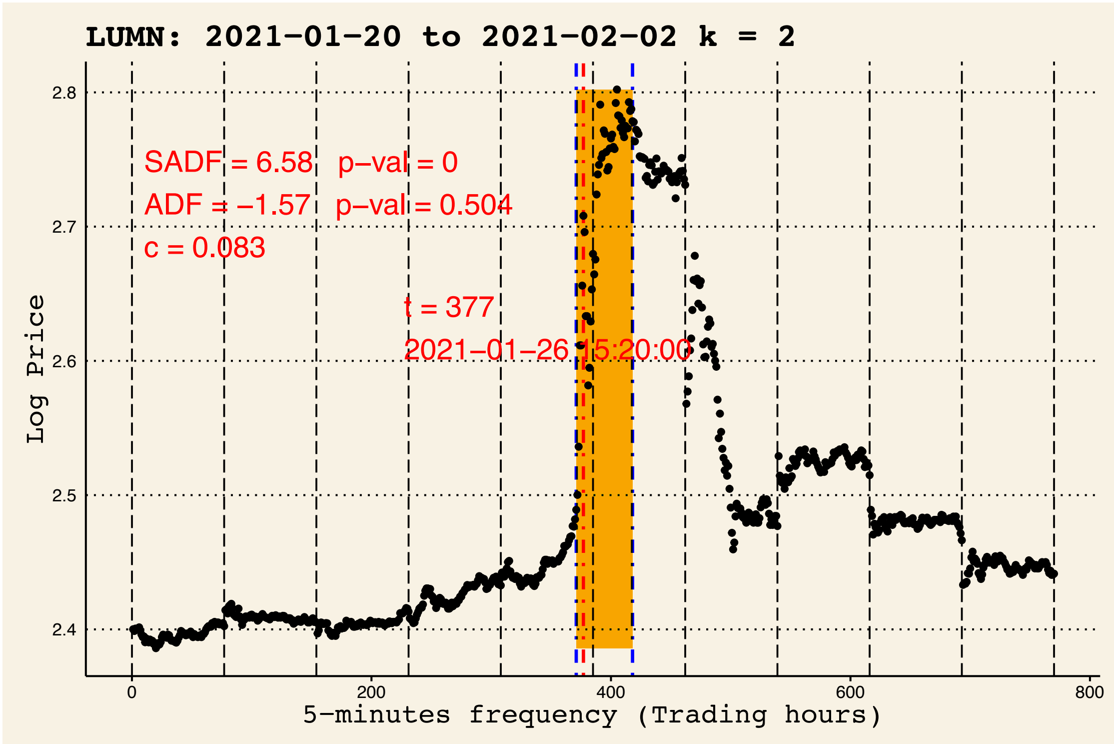

Figure 3: Explosion of Lumen Stock

The primary focus of this paper lies in examining the left tail, particularly when the *p*-value is small. Approximately 6.2% of observations are identified as explosive at the 1% significance level. The vertical red lines aid in assessing the proportion of observations that would be classified as explosive using the standard 1%, 5%, and 10% significance levels. Integrating over the density to the left of these cutoff levels yields percentages of 11% and 15% for the latter two levels. It is worth noting that the exact figures are dependent on the specifications chosen, including the lag order and the value of r_0, which can impact the final test results. However, as demonstrated in the appendix, the observation that explosive episodes are significantly more common than predicted by the null hypothesis remains robust.

It is important to note that the frequency of explosiveness detections within a given 10-day interval does not directly correspond to the number of explosive events when considering a longer timeframe. To illustrate, in simulated martingale data with 100,000 observations, only 0.4% of days feature a peak of an explosion under a 1% significance level. However, randomly selected intervals are detected as explosive with close to a 1% probability because the intervals often overlap and share the same peak of explosions. In our actual data, when analyzing explosiveness detections, we found that 6.2% of these detections actually result in 2.4% separate events when overlap is taken into account. To address this issue, when constructing summaries and analyzing explosive episodes, we primarily work with explosive events by removing overlapping periods to avoid double counting.

As the definition of explosive episodes is based on the convexity of price movements rather than their direction, I categorize explosiveness into two distinct groups for events detected at a significance level of at least 10%. One group relates to upward explosiveness, while the other pertains to downward explosiveness.[12] To illustrate the accuracy of this classification, I plot the distribution of returns conditional on the detection of explosiveness in Figure 4. The return is measured from the starting point of estimation to the moment of explosion detection. As we can see, the distribution is bimodal, with almost no mass around zero return. Explosions up feature a fatter tail, though the total mass of explosiveness down is slightly higher.

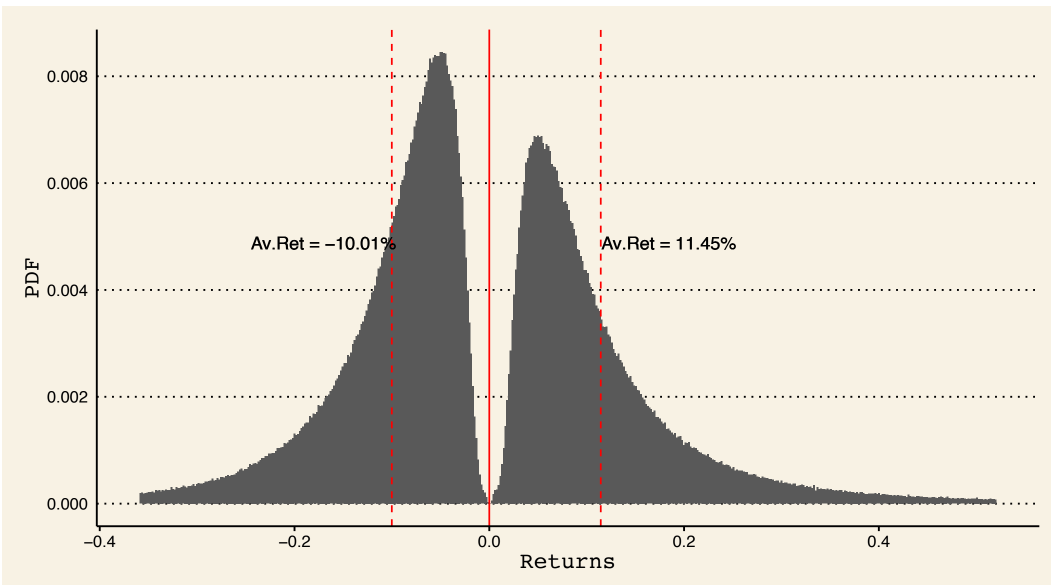

Figure 4: Return distribution

The difference between the two directions of explosions is represented by the varying colors of the EPDF in Figure 5. The density in the explosive region is divided into red and blue segments, representing explosiveness episodes classified as up and down, respectively. Although the occurrence of downward explosiveness is slightly more frequent than upward explosiveness for individual stocks, it is important to note that explosive bursts and booms, in general, occur much more frequently than what a null hypothesis would predict. This null hypothesis would anticipate roughly an equal distribution of these events, with approximately 0.5% of cases being explosive up or down at a 1% significance level.

[12]The classification is akin to the previously studied split into downward/upward variance/jumps decomposition (e.g., Barndorff-Nielsen et al. (2008), Kilic and Shaliastovich (2019)).

				ST Returns				
Expl	N obs.	$\bar{r}_{t,t+l}$	r^{Expl}	Max	Min	s.d.	Skew	Kurt
No	18527676	0.27		3.09	-2.97	0.37	0.15	37.80
Up	1140362	12.57	16.18	4.51	-3.08	0.44	1.71	52.69
Down	1250688	-10.02	-12.50	3.20	-4.40	0.44	-1.44	53.08

Table 1: Return characteristics in explosion window

The first two columns display the average percentage return over a 10-day period and the return at the point of explosion. The remaining columns provide information on extreme short-term (5-minute or overnight) returns and the statistical moments (standard deviation, skewness, kurtosis) of their distribution. The initial row represents non-explosive episodes, determined at a 5% significance level.

The raw properties of return over all observations with clean data in the randomly chosen period of explosion at 5% significance level are summarized in Table 1. On average, the explosive periods are associated with the price shifts in certain direction. Stocks are more volatile and have higher extreme return values when explosions are detected. In section 5 we will do further refinement.

3.6 Short Discussion

Before delving further into the analysis, it is essential to discuss the pros and cons of the detection mechanism employed in this study. The way we have formulated the phenomenon of persistent accelerated returns allows for various methods to detect it. However, any standardized method for uncovering such phenomena will inherently come with a false detection rate, both empirical and conceptual.

The empirical false detection rate arises because any statistical test has imperfect power, meaning that it might not correctly identify every true episode of the phenomenon, leading to both false positives and false negatives. That can be overcome by using the large data sample and considering the false detection as an additional noise to our measure.

The conceptual false detection rate is related to the idea that identified episodes may not perfectly fit the true desired dynamics. In finite samples, numerous different processes can produce similar realizations in data, as demonstrated in our previous example. Additionally, in simulations, explosiveness might resemble high autocorrelation. This is because when allowing autocorrelation to change, processes can become mathematically indistinguishable even though they carry slightly different economic meanings. Distinguishing between these cases for a specific situation can be as challenging as distinguishing between jumps and

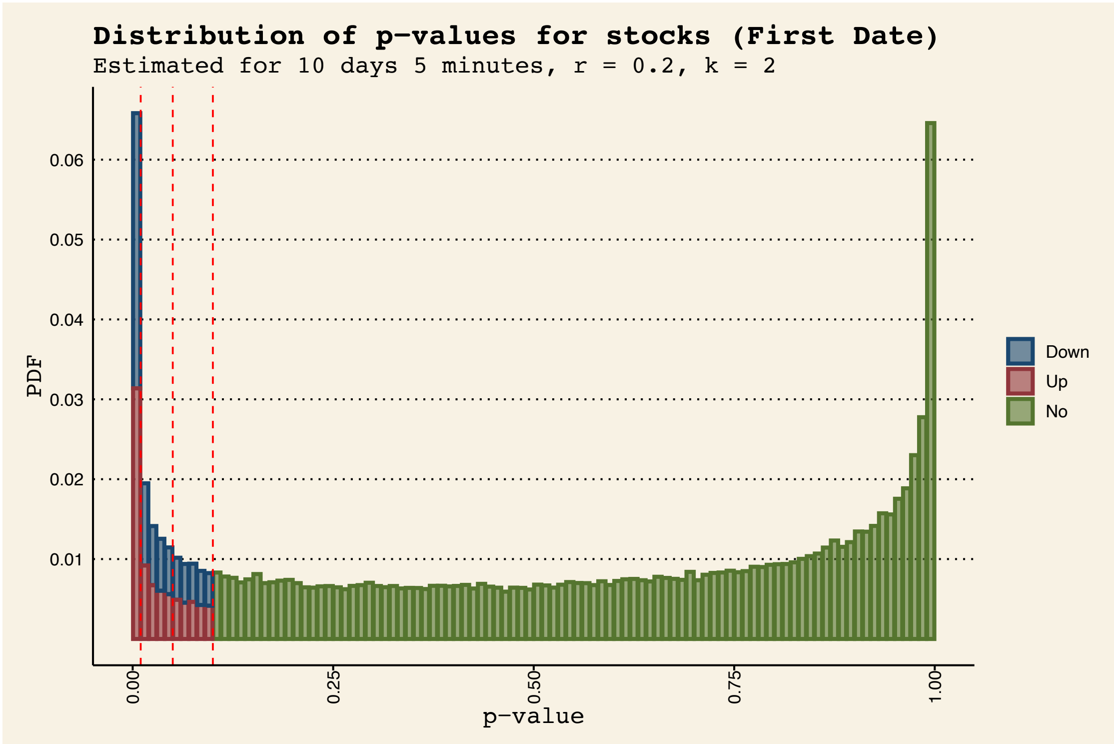

Figure 5: Empirical probability distribution function of p-values of explosiveness

volatility with only a few data points. These challenges are akin to issues of weak identification in econometrics, where it is either impossible or requires more sophisticated techniques to separate the null and alternative models (e.g., Staiger and Stock (1997), Stock and Wright (2000), Mikusheva (2012)).

With these considerations in mind, the chosen detection procedure needed to satisfy several criteria. First and foremost, it had to possess economic meaning and capture episodes that met the required properties: specifically, accelerating returns. Visual testing and summaries based on aggregations, which will be discussed later, indicated that this procedure aligns with these needs. Second, it had to be implementable on a large scale, considering the vast universe of U.S. stocks with high-frequency data. The computational demands of aggregating statistics more complex than simple averages and second moments can be prohibitive. In this paper, we deal with over 22 million observation periods over which we identify explosiveness. The SADF procedure, initiated on a daily basis, strikes a reasonable balance between complexity and computational feasibility.[13] Third, the chosen procedure had to be capable of accommodating bursts of explosive

[13]For context, computing SADF estimates for a specification with a 10-day length and 5-minute frequency takes around a month using standard efficient packages with regressions. However, I have optimized the process writing efficient c++ code to get the estimates within a 7-10 day period using a modern 16-core computer.

episodes and had to run frequently enough to capture the phenomenon effectively. The SADF test is designed specifically to meet these criteria. By employing SADF and rolling its starting point, I aimed to capture the majority of explosive episodes that are not too short-lived. Fourth, it had to provide a clear and testable prediction under mild assumptions, enabling model-free classification of explosiveness in line with theoretical foundations. SADF-based measure does fit the condition since it is effectively a p-value of the test requiring minimum assumptions on the distributions of returns.

The SADF test, while effective for detecting explosiveness in stock prices, does have its limitations. Complex price dynamics, where prices temporarily jump up and down, can pose challenges for the test, potentially leading to false negatives when the economic meaning of the episode is an explosion. For instance, even in a case like GameStop (GME) squeeze, which could serve as a textbook example of explosiveness, the SADF test might not capture it at the most stringent significance level due to multiple jumps and fluctuations that reduce the SADF measure's power.

Alternative detection mechanisms could involve refinements using machine learning techniques or more advanced versions of ADF-based measures (e.g., Phillips et al. (2015b), Phillips et al. (2015a)). These methods might offer more precise results for particular episodes, but they often come with increased computational demands. It is worth emphasizing that the primary goal of this paper is not to accurately classify each and every explosion among the more than 20 million data periods. Instead, the focus is on capturing a more general phenomenon, identifying its prevalence beyond what could be attributed to noise, and exploring its relationship with other stock characteristics and liquidity measures.

4 Market Explosiveness

To isolate the unique explosiveness of individual stocks amidst the broader fluctuations in the market, it is crucial to first study explosiveness of the factors driving the common stock market variance structure. In this section, I will assess the explosiveness of the standard long portfolios commonly used in existing literature to construct factors. I will also draw connections between this portfolios' explosiveness and other return metrics while examining its relevance to the broader market's explosiveness.

In their study, PSY (2011) scrutinized the explosiveness of the market measured by S&P 500 index as a long-term phenomenon in order to empirically identify bubbles. In contrast, analysis of this paper closely examines high-frequency data.

Following the procedures outlined by Fama-French, I create 24 portfolios and a market portfolio using a value-weighted approach that encompasses common shares traded on the NYSE, NASDAQ, or AMEQ.[14] These 24 portfolios result from a combination of four 2x3 portfolios categorized by size and firm characteristics such as Book-to-Market, Investment, Operating Profitability, and Momentum (prior 12-to-2 return). Therefore, I consider portfolios central to Fama-French's five factors, including the momentum factor. Utilizing data matching between TAQ and CRSP datasets, I calculate values at a 5-minute frequency for each portfolio over specific intervals and estimate the explosiveness of the logarithmic values. The observations from September 2003 to December 2022 form 4,853 overlapping ten-day intervals under study.

The market portfolio occasionally exhibits explosive characteristics, although these instances are less frequent than for individual stocks. Figure 7 illustrates the empirical cumulative distribution function of intervals with a specific p-value of explosiveness. Each column in the plot represents the probability of an interval being detected as explosive at a given significance level, ranging from 0.01 to 0.05 with a 1% increment. As before, the columns are divided into red and blue segments, representing explosiveness episodes classified as up and down, respectively.

On average, we observe more explosive episodes than one would anticipate if the market returns were following an autoregression process. Despite approximately 60% of the ten-day periods showing a positive cumulative return, only .54%, 1.07%, 1.63%, and 2.91% of the observations can be classified as experiencing explosive upward movements at the 1%, 2.5%, 5%, and 10% significance levels, respectively. These percentages are notably lower than one would expect under the hypothesis of autoregression, which suggests a stronger market reversal as the market consistently rises. However, it is not uncommon for markets to exhibit sudden declines with explosive characteristics, leading to measured explosiveness probabilities of 3.48%, 5.09%, 6.86%, and 9.99% at the same significance levels.

As discussed in Subsection 3.5, it is important to note that the frequency of detection is not equivalent to the number of occasions when explosive episodes can be detected. Specifically, explosive episodes were detected on 176 days at the 5% significance level, with 89 of these days featuring detection at the 1% significance level. Out of the last 89 days, 72[15] can be attributed to explosions down, while the remaining 17 are associated with explosions up.

Similarly, the other core portfolios exhibit more pronounced downward explosiveness when compared to their upward movements. The summary for the 24 portfolios is captured by the left-tail of the empirical cumulative distribution function of their p-values in Figure 7. The rejection of the non-explosive null

[14]For additional portfolio details, please refer to the Appendix.

[15]Out of the 176 detected explosion dates at the 5% significance level 132 (44) are explosive down (up). Out of the 135 detected explosion dates at 2.5% significance level 105 (30) are explosive down (up).

hypothesis occurs more frequently than expected when considering an autoregressive model for portfolio returns. This phenomenon is primarily driven by downward explosions, which are highly concentrated on certain days. To illustrate, out of 116,472 portfolio-day observations, 8,654 (7.43%) are identified as downward explosions, while 2,897 (2.49%) are labeled as upward explosions at 5% significance level.[16] If we focus on identifying days when at least 20 (or 16) out of the 24 portfolios exhibit explosiveness, we find that these days account for 2,797 (or 4,073) observations in the case of downward explosions and 371 (or 673) observations in the case of upward explosions. This demonstrates a significant clustering of downward explosions within the portfolios, possibly reflecting common downshifts in the market.

The distribution properties of the returns are presented in Table 4. When compared to individual stocks, explosiveness can be achieved with relatively modest returns of 3.44% upward and -5.06% downward. However, it is important to note that within a 10-day interval period, these returns are considered substantial. Furthermore, the overall return during these periods is typically significant, indicating that portfolio explosions often have a transitional nature. Though, it partially reverse on average within the estimation window since $|\bar{r}_{t,t+l}| < |\bar{r}^{Expl}|$. Overall, there is no big difference between the returns in terms of second and fourth moments for explosive episodes of the portfolios. The upward movements are slightly less volatile while downward movements are more volatile. However, the bias toward larger positive (negative) returns during upward (downward) explosions is reflected in the higher (lower) average returns and skewness values.

Lastly, it is important to note that explosiveness is not uniformly distributed across all 24 portfolios. Portfolios composed of small stocks are more likely to be detected as explosive, particularly in the context of explosive downturns. Additionally, portfolios characterized by high 12-to-2 returns and a high book-to-market ratio exhibit a notably higher propensity for explosiveness. This does not necessarily imply that the individual stock constituents within these portfolios are also explosive, but it does suggest the importance of controlling for market fluctuations in these dimensions when estimating stock explosiveness.

4.1 Risk-adjusted explosions

Since explosiveness can also be observed in the fluctuations of market portfolios, it is essential to disentangle the phenomenon and determine if the explosions in individual stocks are solely driven by them. To address this concern, I estimate the explosion of individual stocks based on risk-adjusted prices. Specifically, to

[16] Out of 116,472 portfolio-day observations, 4,011 (3.44%) are identified as downward explosions, while 978 (0.84%) are labeled as upward explosions at 1% significance level.

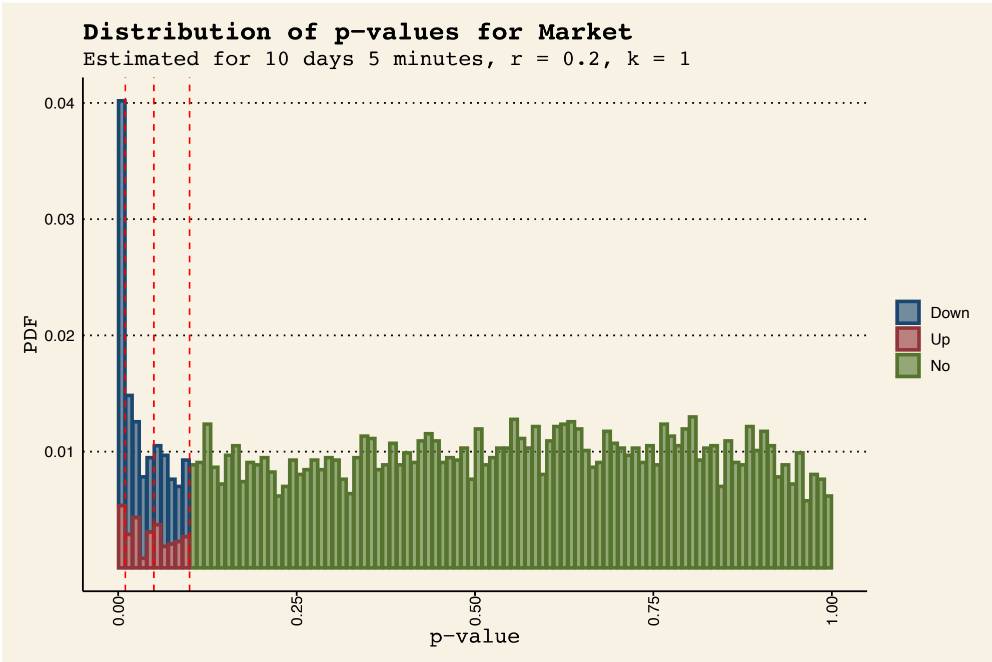

Figure 6: Empirical probability distribution function of p-values of market explosiveness

The plot illustrates the empirical probability distribution function of SADF *p*-values, which were estimated using the market portfolio data over ten-day intervals spanning from September 2003 to December 2022. The sample consists of 4,853 ten-day intervals, and the intraday data is timestamped at five-minute intervals. The parameters used for the estimations are set at $k = 1$ for the lag order and $r = 0.2$, indicating a pre-estimation period of two days. For alternative parameter settings, please refer to the appendix.

isolate idiosyncratic fluctuations, I estimate the explosion based on the price:

$$P_{s,t}^{Adj} = \frac{P_{s,t}}{1 + Adj_{s,t}}.$$

Here, $Adj_{s,t}$, represents the systematic return in the stock according to one of three models: JM (just market), CAPM, or FF3 (Fama-French 3 factors). The systematic return is defined based on realized factors over the same periods and pre-estimated risk exposures[17] to the factors of the individual stocks. For example,

$$Adj_{s,t}^{FF3} = \beta_{s,t}^{Mkt,d} Mkt_{s,t} + \beta_{s,t}^{SMB,d} SMB_{s,t} + \beta_{s,t}^{HML,d} HML_{s,t},$$

where $HML_{s,t}$, $SMB_{s,t}$, $Mkt_{s,t}$ represent returns over the respective portfolios if invested at time $t = 0$. The JM specification ignores the estimated factor exposures and assigns a value of 1 to the market beta while setting all other risk exposures to zero.

[17]The risk exposures are estimated based on high-frequency data as well, see Appendix for details.

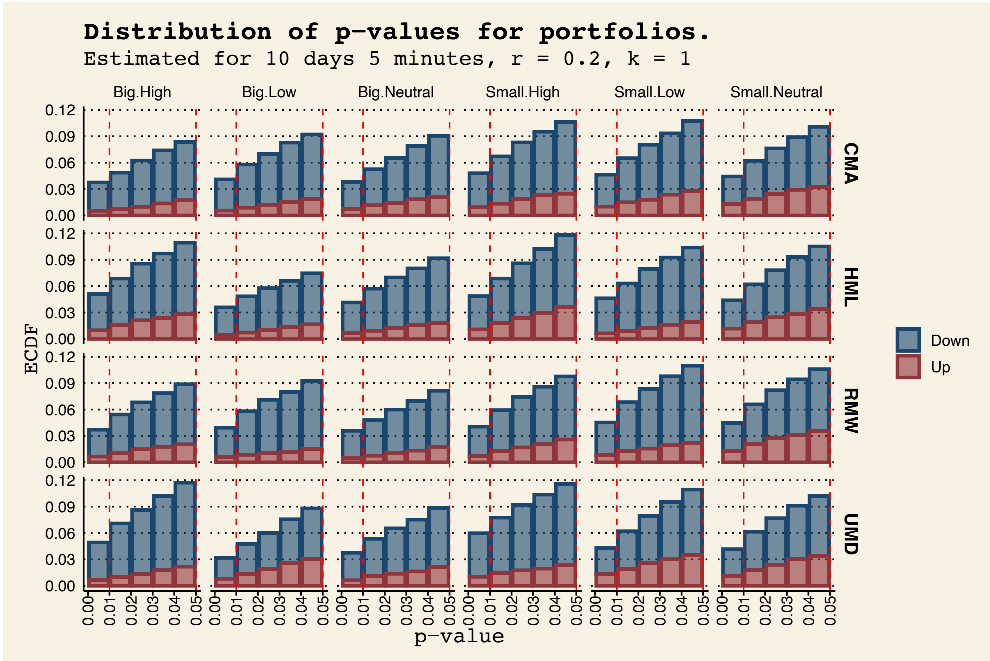

Figure 7: Empirical cumulative distribution functions of p-values for FF-portfolios

The plot illustrates the empirical cumulative distribution function of SADF p-values, estimated using 24 long portfolios underlying SMB, HML, RMW, CMA, and UMD long-short factors. The estimate produced over 4,853 ten-day intervals from September 2003 to December 2022. The intraday data is timestamped at five-minute frequency. The parameters used for the estimations are set at $k = 1$ for the lag order and $r = 0.2$, indicating a pre-estimation period of two days.

Table 5 presents the frequency of breaking through the critical values at 1%, 2.5%, 5%, and 10% significance levels for each of the specifications. It can be observed that the frequency of explosion detection slightly decreases, moving from 6.2% to values around 5.6% at the 1% significance level. While this may be partly due to the introduction of some noise into individual stock returns, it is noteworthy that the adjustment primarily affects explosion down, whereas explosion up remains largely unaffected. This observation aligns with our previous findings that the market as a whole tends to exhibit explosiveness in the downward direction, while there are fewer instances in the upward direction.

The estimated explosions, under various adjustment methods, demonstrate a reasonable level of stability. The conditional probability of detecting an explosion according to one specification relative to another, at the 1% significance level, ranges from 0.64 to 0.82. Lower probabilities of overlapped detection are observed when a no adjustment specification is considered, while larger probabilities are evident when comparing two models with some adjustments. These probabilities increase when we relax the significance level to 5%. For further details, refer to Table 6.

The subsequent findings in the paper about individual stocks maintain their robustness even when we eliminate the effects of systematic risk fluctuations by applying the adjustment method as described in the section. Nevertheless, the specification without adjustments serves as the one that will be mainly used in the paper due to its ease of interpretation and its adherence to the conventional approach found in the literature that analyzes high-frequency stock prices.

4.2 Association of Explosive Episodes in the Market and Individual Stocks

Let us explore the relationship between the detection of explosiveness in individual stocks and the market portfolio. Table 8 provides a summary of this relationship. First, without any adjustment, the detection of explosiveness upward at the 1% significance level increases slightly by seven basis points. The detection of explosiveness even decreases at the 5% and 10% levels. In contrast, when the market is explosive, 8.84% of stocks exhibit explosiveness downward. This proportion increases as the significance level for detecting individual stock explosiveness becomes more relaxed. However, for the sample with no market explosion, the probabilities of upward and downward explosions in individual stocks are almost identical, with both being around 2.83%$\pm$ 0.01%.

However, if we classify market explosions into "up" and "down" groups, the corresponding up and down individual stock explosions increase. This can be observed in the second and fourth sections of Table 8. Therefore, it appears that individual stock explosions are procyclical. Explosions in the market, whether up or down, tend to predict both upward and downward explosions in individual stocks. However, due to the rarity of these market events, especially explosions, it would be inaccurate to claim that they are the sole explanation for why individual stocks become explosive.

Sections 5 and 6 of Table 8 provide information on the percentage of dates when individual stock explosions are detected and whether they coincide with a market explosion on the same date. These statistics reveal that only 5.32% of upward explosions in individual stocks occur on the same date as a market explosion, while 19.39% of downward explosions align with market explosions. This suggests that individual stock explosions are primarily an idiosyncratic phenomenon, even though they exhibit some connection to market fluctuations.

By introducing adjustments by the market or by standard factor adjustment to stock prices before SADF-estimation, we might partly mitigate the driving force of the market behind explosive episodes. In the second and fourth sections of Table 8, we observe that when the market is classified as explosively up, the detection of individual stock explosiveness decreases from 7.50% to 3.58%, 3.21%, and 2.97% with respect to JM, CAPM, and FF3 adjustments. Similarly, when the market is explosively down, the detection of individual stock explosiveness decreases from 10.41% to 3.94%, 3.90%, and 3.73% with respect to JM, CAPM, and FF3 adjustments.

These adjustments, however, introduce some noise in the estimation, which can be seen when examining the reaction of upward explosions to adjustments, as shown in the first section of the same table. Upward explosions increase in response to each of the adjustments, particularly for the Just Market specification. The main reason for this behavior is that by subtracting a downward explosive component from the price, we effectively incorporate an explosive element in the opposite direction. Given the tradeoff between employing the most sophisticated adjustment scheme and introducing additional noise, this paper ultimately considers the relatively simple CAPM and FF3 adjustments, which hopefully remove the majority of market fluctuations. The JM specification, which exhibits the most extreme "poisoning" effect, will be excluded from further analysis.

5 Characterization of explosiveness in stocks

Once we have established the prevalence of explosive behavior among individual stocks, the central questions that arise are what precedes and follows such behavior for these stocks. Can investors generate alpha from these stocks, and what types of risks are associated with doing so?

To ensure that the subsequent analysis is not affected by discussions concerning low-priced stocks, I implement additional filters by selecting only stocks with a price of at least $5 measured 30 days prior to the explosive period. To avoid studying overlapping events, I apply a procedure[18] that ensures a minimum of 10 days' difference between analyzed explosions for the same stock. This effectively reduces the number of studied events but guarantees studying unique market episodes. A few additional procedures discussed in the Appendix are incorporated to remove likely temporary jumps that would cause in extra reversal but are most likely associated with the microstructure noise not removed at pre-cleaning steps. The results do not change qualitatively if the filters are omitted.

Table 3 provides a summary of the distribution of total returns and high-frequency returns over three periods: the 10 days before, during, and after the identified explosion events. Explosions are initially identified on the raw price data, but post-explosion analysis is conducted using Fama-French 5-factor adjustment. Alternative combinations of estimation and adjustment are presented in the appendix.

[18] Refer to the Appendix for additional details on the selection procedure. The total sample results in 14,201,205 firm-dates when an explosion *can be* detected, assuming no other filters are applied (e.g., requirements for identified other measures such as firm characteristics).

The summary of 10-day intervals reveals a slight pre-trend for explosions up before adjustment, which is completely removed by risk-adjustment. Post-explosion 10-day intervals show no evidence of abnormal returns. However, during the ten-day window surrounding the explosion, a substantial change in the direction of the explosion is observed. This change amounts to 7.54% for explosions up and -6.65% for explosions down. The FF5 adjustments mitigate this effect, but the aggregate change remains economically significant, with abnormal returns of 5.46% for explosions up and -4.68% for explosions down over the window containing the explosion. This underscores that explosions are part of transitional dynamics related to price discovery and play a role in creating persistent changes in prices.

Figure 8 depicts the average price dynamics of these stocks after the day when an explosion is detected, with events centered around the detection time. The graph reveals that, conditional on finding oneself in the midst of an explosion, the optimal strategy, on average, is to divest from the asset. Even selling at the bid and later buying at the ask yields, on average, more than a 1% alpha, a statement that we will subject to more rigorous testing later on. However, it is essential to note that this graph does not capture the associated risks associated with this trading strategy.

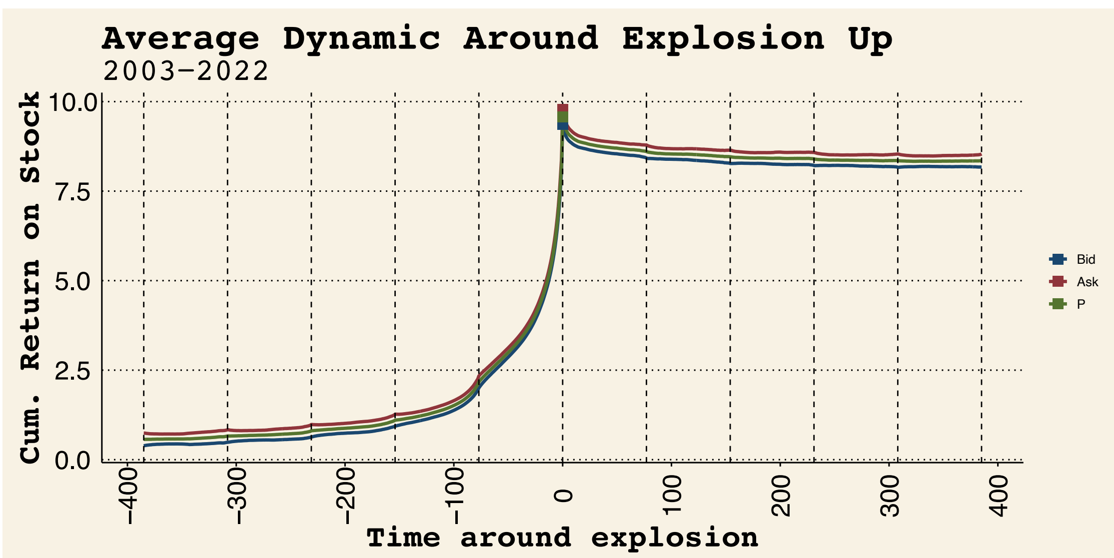

Figure 8: Average cumulative return on explosive stock

The average cumulative return on an investment in an explosive asset over the 5 days leading up to the explosive detection period is shown in the graph. The green, red, and blue lines represent the total return when considering the ask, bid, or midquote price as the asset's value. This graph encompasses the 5 days before and 5 days after the explosive detection timestamp. The investment happens 10 days prior the detection time.

It is worth mentioning that although the average explosion might, at first glance, resemble a price jump, the detection of explosions is not actually driven by jumps. To demonstrate this, let us examine the distribution of returns immediately before detection. Table 7 summarizes the distribution of returns (r_{DE}) just before an explosion, grouped by the direction of the explosion. As we can observe, the average returns are only a few times larger than the standard deviation of returns, which is not high for high-frequency returns and significantly smaller than overnight returns.

Nevertheless, it is true that following an extreme return, the detection of an explosion is mathematically more likely. This is because the rolling supremum-ADF measure is more likely to reach the threshold specifically at the moment. This can be observed by analyzing the times of the day when explosions are usually detected. Explosions are more concentrated at the market opening since this is when extreme (overnight) returns are typically realized. Conditional on at least 5% (-5%) jump[19] in the stock price, 16.6% (21.9%) of the days happen to be identified as days with detected explosion upward (downward) at 5% significance level[20].

5.1 Alpha in buy-sell and sell-buy strategies

The open question remains whether anyone can actually generate alpha, accounting for market frictions when they find themselves in an explosive episode. Figure 9 demonstrates that even after risk-adjustment, performed using the Fama-French five-factor model, explosive episodes tend to exhibit partial reversals on average, both for upward and downward explosions. The red and blue lines in the graph represent the average bid and ask prices, respectively, converted into return terms. Notably, the fact that the blue (bid) curve eventually rises above the red (ask) curve suggests the possibility of a profitable Sell-Buy strategy. This finding stands in stark contrast to a similar picture based on high-frequency jumps in prices, where some minor reversals might occur but cannot be profitably confirmed based on available bid and offer prices.

Table 9 presents the systematic risk exposures of the daily trading strategy involving investment in explosive stocks, utilizing mid-quote prices.[21] These portfolios are mandated to include a minimum of 10

[19] That means that the absolute value of the return over a 5 minute interval or the overnight return is greater than 5%. The summary is based on stocks that cost at least \$5 30-days before the date. Vice versa, only 8.6% and 9.0% of the detected explosions up and down respectively experience the jump of the magnitude. See the Appendix, for alternative definitions (i.e., different thresholds) of jumps.

[20] Here, the lag-order is selected by BIC criterium.

[21] I must mention that the provided daily portfolio adjustment is imperfect since it does a daily adjustment while investing into different stocks happen intraday at the time specific for given stock. In the essence the portfolio is dynamic since it includes and excludes stocks on rolling basis, so ideally one need to derive factor exposures of the dynamic portfolio and make the adjustment accordingly. I did not have time to implement the procedure and plan to add it later.

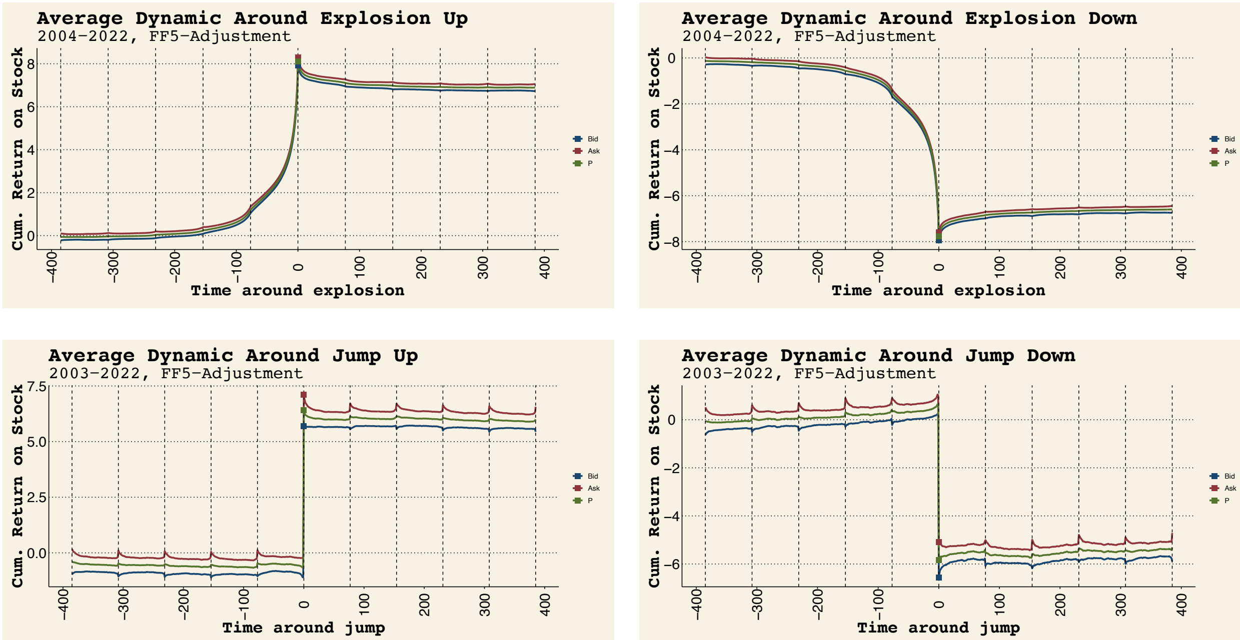

Figure 9: Around Explosions and Jumps with risk-adjustment

The explosive intervals are selected using a non-overlapping procedure. Additionally, the five-day intervals before and after the explosive interval are chosen for the same stocks. For a stock to be considered, it must have reliable price data for a thirty-day interval and an initial price of at least \$5. Explosiveness is detected on FF3-adjusted prices at 5 minute frequency with $k = 1$.

stocks.[22]. In comparison to portfolios based on explosive up stocks, portfolios comprising explosive down stocks tend to be more susceptible to aggregate market movements. They exhibit a market beta closer to one, larger SMB and momentum betas, and a higher r-square in the sample. Despite that they also generate slightly higher alpha. As observed, although the strategies exhibit a procyclical nature, the risk-adjusted alpha of the portfolios stands at 1% (1.3%) for short (long) position in the portfolio following upward (downward) explosions.

To test the feasibility of this strategy, let us calculate returns while considering the spread. Table 10 illustrates that although the generated alpha is naturally lower for these strategies, it remains statistically significant. For calculating returns, I use the actual bid and offer prices. A buy-sell strategy is employed following an upward explosion, where we use the offer price as the numerator and the bid price as the denominator. Conversely, a sell-buy strategy is used following a downward explosion, with the bid price as the numerator and the offer price as the denominator. The strategies yield daily alphas of 0.6% and 0.9%.

[22]It is worth noting that the results become even more significant without this condition, although there is a room for debate on whether it can truly be classified as a portfolio.

5.2 Role of liquidity

The model mechanisms developed in Sections 7 and 8 hinge on the ultimate connection of explosive mechanisms with both *expected* and *realized* directional trading. The former will be discussed in Section 9. Now let us focus on which *realized* liquidity changes accompany the explosive episodes. To underscore significance of the explosive transitional dynamics, I continue examining the dates with detected explosions and jumps collectively, considering major liquidity measures such as trading volume around detected explosions, buying pressure (order imbalance), and effective spread.

The daily liquidity measures are derived from data sourced from Intraday Indicators by WRDS. Daily buying pressure is defined as the net order imbalance, representing the difference between the share of trade volume assigned to buyer-initiated and seller-initiated trades. This classification is based on Lee and Ready's (1991) algorithm. Abnormal characteristics refer to the difference between the daily measure and the rolling average of the characteristic from the previous 30 days, normalized by the average. The reported characteristics represent the two-day average of these abnormal liquidity measures on the day of event detection and the preceding date. Including the preceding date enhances accuracy, as explosions must be preceded by buying pressure, often occurring at the beginning of the trading day. All variables are winsorized at the 2.5% level. To maintain consistency with the previous analysis, I concentrate on a subset of non-overlapping explosive episodes to prevent double-counting similar events. However, the results of this section are even more pronounced when considering all detected dates.

Figure 10 provides an unconditional summary of daily liquidity measures. For clarity, cases with relatively few observations, such as explosions up with jumps down on the same day, are excluded (refer to Table 11 for specific numbers, including rare cases). In the first panel, the figure shows that both explosion and jump dates feature abnormal trading volumes. The latter increases by 65% and 69% for jumps up and down, respectively, with no explosion. Trading volume for explosions grows more moderately, by 44.7% and 35.6% for up and down explosions, respectively. The effect is even more pronounced when dealing with the interaction of jumps and explosions, especially for upward movements. Explosive episodes featuring a jump have a trading volume almost three times larger than the average trading volume in the previous 30 days. All deviations are greater than or close to one standard deviation of 48.2% of the abnormal trading volume.

The next panel showcases abnormal buying pressure for the same groups. Despite abnormal trading volume for jump days, buying pressure is relatively small for both jumps up and down. This could be partially attributed to the known noisy nature of directional classification (see, e.g., Chakrabarty et al. (2015)), but notably, abnormal buying pressure is strikingly different for explosive episodes. For explosions up with and without jumps, it forms an economically meaningful 9.3% and 9.6% increase, around 70% of the standard deviation in value (13.6%). The difference does not vary much based on whether jumps happen on the same date. Explosiveness down is associated with negative buying pressure, indicating that most trades were classified as sales. The magnitude is smaller, especially for interaction with jumps down. This is an additional argument pointing out that explosive down episodes are different by nature from explosions up.

Finally, the last panel emphasizes that the bid-ask spread typically widens for all these events, but the growth is modest for explosive episodes compared to jumps. The column bars report abnormal bid-ask spread[23]. Explosions up without jumps almost do not feature a significant increase in spread, opposite to explosive down episodes that feature a 12.8% increase in abnormal effective spread.

The documented summaries cannot be explained solely by other firm characteristics. Table 12 reports the coefficients of two-way fixed-effect regression using the same liquidity measures as dependent variables and the indicators of the events as the primary explanatory variable. The fixed effects include firm and date. Additional firm-characteristics controls are reported in separate specifications. These controls include size, book-to-market, turnover, momentum, and other characteristics discussed in the table's footnote. After controlling for all the confounding factors, the coefficients confirm the discussed summaries in approximately the same magnitudes: detected explosion up (down) that do not feature jumps are associated with an 8.2% (-5.6%) change in abnormal buying pressure and 45.2% (44.1%) in abnormal trading volume. Notably, conditional on jumps in the same direction, explosive episodes feature significantly higher trading volume. At the same time, the effect is dampened for abnormal buying pressure by -1.5% (4.7%), emphasizing that jumps have a different nature. The latter effect is the opposite for observations that feature explosions up (down) and jumps down (up) on the same date: those explosions feature more imbalanced orders. This could be related to cases of extreme price impact that immediately reversed afterward, consistent with the notion of a bubble burst.

To summarize, explosive episodes are closely related to the presence of abnormal trading volume and abnormal buying pressure. While the former accompanies jumps as well, even with larger magnitudes, the buying pressure is more related to the more gradual explosive price changes. The episodes with explosions are associated with the widening spread, but the magnitude of the effect is modest, especially in comparison to jumps.

[23] The reported abnormal bid-ask spread is normalized by the preceding 30-day average of daily dollar volume effective spread

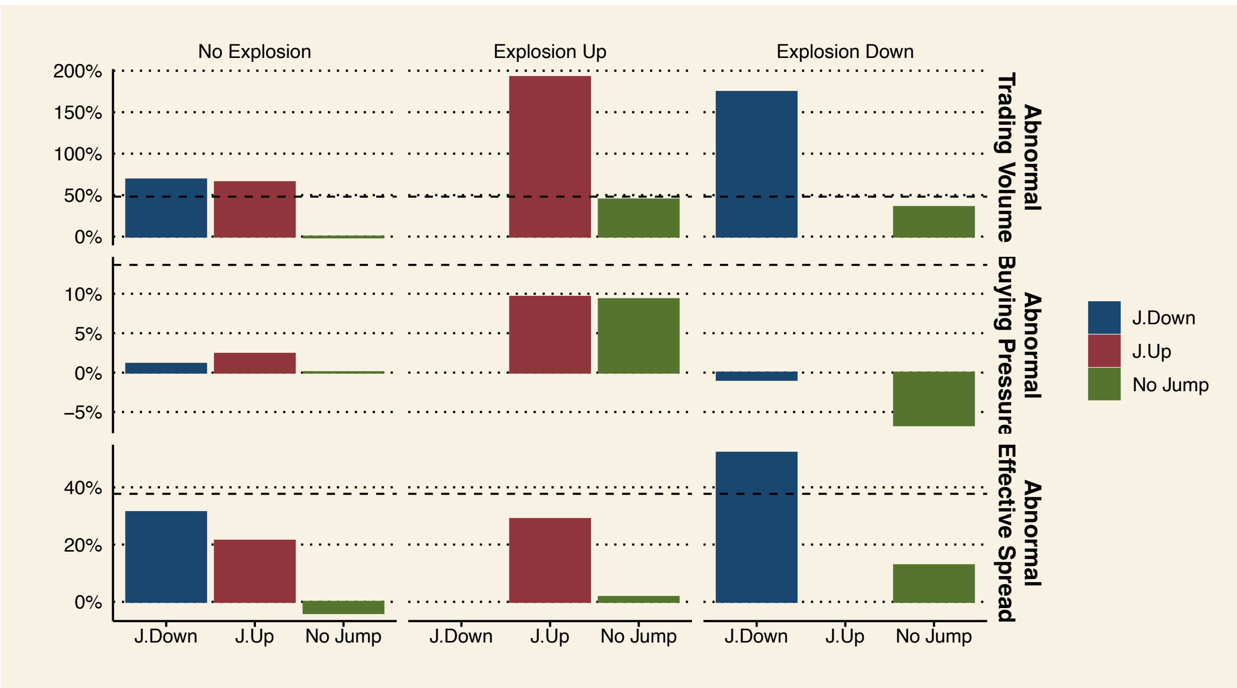

Figure 10: Average Liquidity Measures around Explosions and Jumps

This figure displays the average daily abnormal trading volume, abnormal buying pressure, and abnormal effective spread measured on the day of detection together with the preceding day for either explosions, jumps, or both. For the sake of clarity in presentation, cases with a relatively low number of observations are excluded, such as explosions up with jumps down on the same day. Refer to Table 11 for the specific numbers including the rare cases.

5.3 Explosiveness in low frequency volatility

Prior research has emphasized the importance of daily extreme returns, also called *price jumps* in the literature, in influencing both realized variance and the behavior of average stock prices (e.g., Savor (2012), Jiang and Zhu (2017), Kapadia and Zekhnini (2019)). When viewed from a broader perspective, many explosive episodes might initially appear as jumps due to their association with substantial price movements.[24] However, as we have observed, when we delve into the high-frequency price dynamics, the behavior of explosiveness differs from the theoretical concept of jumps. Explosive episode returns tend to display a relatively smooth pattern, characterized by partial reversals—a feature not typically observed with real high-frequency jumps. The absolute magnitudes of the explosions prior to detection already suggest that they must significantly contribute to the overall variance of stocks. This subsection reports the total share of the explosive episodes in the idiosyncratic variance.

I will focus on daily returns in stocks and adjust them based on six factors, five Fama-French factors together with momentum factor,[25] to get an idiosyncratic component of the stock return:

$$r_{s,d}^{adj} = r_{s,d} - \sum_{f \in F} \beta_{s,f}^{d-30,126} f_d.$$

Here, the factor exposures $\beta_{s,f}^{d-30,126}$ are daily pre-estimated betas derived from a 126-day rolling time-series regressions on the six factors.[26] $d-30$ superscript indicates that the 30-day lagged exposure is taken. The previous values of $r_{s,d}^{adj}$ are also used to build a rolling measure of idiosyncratic variance $\sigma_{s,d-30}^{adj}$ that uses an exponentially weighted moving average (EWMA) model to aggregate previously observed.[27]

To study the contribution of explosive episodes to idiosyncratic variance, let us decompose the adjusted return into three components. One is the return that occurs on the date when an explosion is detected. The other is the return on days when no explosions are detected but still feature a substantial price change, named a *daily frequency* or *low frequency* jump, and the residual return that can be attributed to other *diffusive* parts of the return. The jumps are different from the high-frequency jumps used for comparison in Sections 5.1 and 5.2; the days with those jumps potentially could be classified into every one of the components including the *diffusive* group. Additionally, split each of the explosions and jumps into components with upward and

[24]Similar spurious classifications of high-frequency "jumps" that are related to a local increase in volatility are discussed in Christensen et al. (2014).

[25]The data is obtained from French's website

[26]The rolling regression include an intercept.

[27]This approach follows Kapadia and Zekhnini (2019). $\sigma_{i,d} = \sqrt{(1-\lambda)\sum_{s=1}^{d-1} \lambda^s \left(r_{i,d-s}^{adj}\right)^2}$, where $\lambda = 0.94$. Hence, the greater weight is placed on the most recent realized abnormal returns.

downward movement on the day. More specifically, the abnormal return decomposition is

$$r_{s,d}^{adj} = \underbrace{r_{s,d}^{adj} \times \sum_{dir \in \text{up, down}} E_{s,d}^{expl,dir}}_{\text{Explosions}} + \underbrace{\sum_{dir \in \text{up, down}} r_{s,d}^{adj} \times J_{s,d}^{dir}}_{\text{Jumps}} + \underbrace{r_{s,d}^{adj} \times \left(1 - \sum_{dir \in \text{up, down}} E_{s,d}^{expl,dir} - \sum_{dir \in \text{up, down}} r_{s,d}^{adj} \times J_{s,d}^{dir}\right)}_{\text{Residual } diffusive \text{ component}}, \tag{10}$$

where $E_{s,d}^{expl,dir}$ is an indicator of a detection of an explosion on date d and $J_{s,d}^{dir}$ is an indicator of a 3 idiosyncratic volatility deviation on the day that is not an explosion.

$$J_{s,d}^{up} = (1 - E_{s,d}^{expl}) \times \mathbb{I}\{r_{s,d}^{adj} \geqslant 3 \times \sigma_{s,d-30}^{adj}\},$$

$$J_{s,d}^{down} = (1 - E_{s,d}^{expl}) \times \mathbb{I}\{r_{s,d}^{adj} \leqslant -3 \times \sigma_{s,d-30}^{adj}\}.$$

	All	Other Diffusive	*L.F. Jump* Up	Explos. Up	Explos. and *L.F.J.*Up	*L.F. Jump* Down	Explos. Down	Explos. and *L.F.J.*Down
Share obs.	100	95.33	1.09	1.01	0.36	0.87	1.02	0.32
Average	2.23	-4.74	22.34	8.09	9.2	-16.62	-7.91	-8.15
S.E.	0.45	0.46	1.92	0.16	0.21	1.66	0.17	0.38
t	4.95	-10.34	11.64	49.42	43.78	-9.99	-46.45	-21.23
VW Average	0.55	-0.86	12.82	4.4	3.66	-10.99	-4.52	-3.95
2-days Return	2.31	-4.52	16.73	10	8.58	-11.89	-9.35	-7.25

Table 2: Decomposition of daily return into explosions, jumps, and other *diffusive* component.

The table presents the average returns based on the decomposition of abnormal returns into explosive, jump, and other diffusive components. The components consist of (1) days with extremely positive (negative) returns, named low-frequency jumps up (down), that do not feature explosion detection, (2) days that feature explosion detection (either up or down) but not the extreme return, and (3) days that feature both of the phenomena and (4) all other daily returns, named "other diffusive." The extreme returns are defined by the return being three times higher than the estimated idiosyncratic volatility. Each component is assigned zero if another component is not zero. The first line of the table reports the share of daily observations attributed to one of the groups. The second line reports the average annualized return calculated based on the component (including zeros mainly assigned to non-diffusive components). Standard errors, clustered by calendar month and respective t-statistics, are reported in the following two lines. The fifth line reports the average calculated using a value-weighted scheme. The sixth column reports the annualized returns calculated over 2-day windows covering the event's day (either l.f. jump or explosion, or both) and the preceding day. Given the aggregation, the calculation uses twice less observation. Stocks with prices less than $5, observed within 30 trading days prior, and those lacking clean data in TAQ meeting the filtering conditions for SADF estimation are excluded. The table includes common stocks from September 2003 to December 2022, totaling 13,809,482 observations. All values are presented in percentage points. See summaries in Table 13.

Table 2 presents each component's annualized average contribution in percentage points to the daily idiosyncratic return. Note that since ill-defined time-series data, mainly from the small and least liquid stocks, are filtered out from the observations, the sample is biased towards medium and large stocks. See

Table 13 for sample details. Explosiveness is captured on Fama-French 3 adjusted prices with $k = 1$ and $\alpha = 0.025$[28]. The total *equally weighted average* of the annualized adjusted return is 2.23%, which is not zero due to (1) noisy estimates of factor exposures and their imprefect performance out of the sample, (2) the presence of short-term reversal.[29]

The jumps component, excluding explosions, constitutes approximately 1.96% of the sample. The low frequency jumps, known for bias toward upward movements, contribute approximately 22.34% to yearly upward movements and 16.62% to downward movements, with a net positive effect generally compensated by other diffusive parts of the return. Explosive episodes, with roughly equal probabilities in the sample, can be split into two groups: those featuring a three standard deviation price change on the detection date (0.36% and 0.32% for upward and downward movements, respectively) or not (1.09% and 1.02%). Explosions disproportionately contribute to the average changes annualizing into 8.1% (9.2% for the interaction) and -7.9% (-8.2%) for upward and downward movements, respectively. Hence, the dates of detected explosions on average form 77.4% (upward) and 96.6% (downward) of the other extreme daily returns when an explosion is not detected.

Similar to jumps, the magnitudes of the abnormal returns naturally decrease when using a value-weighted scheme. They are reported in the last line of Table 2. The largest-cap stocks forming substantial value in the sample have smaller idiosyncratic variance and generally better fit the rich factor structure we impose. Since a substantial part of the explosive path might occur on the previous days as well, I also split the sample into two-day intervals *requiring* the second day happening on the *event*, i.e., either detection of explosion or jump,[30] calculating two-day total return $r_{s,d}^{adj,2} = (1 + r_{s,d}^{adj,2}) \times (1 + r_{s,d-1}^{adj,2}) - 1$.

Similar to jumps, the magnitudes of the abnormal returns naturally decrease when using a value-weighted scheme. They are reported in the last line of Table 2. The largest-cap stocks, forming a substantial value in the sample, have lower idiosyncratic variance and generally better fit the rich factor structure we impose. Since a substantial part of the explosive path might have also occurred on the previous days, I also split the sample into two-day intervals, with the second day happening on the event, i.e., either the detection of an

[28]Results are similar for no adjustment, but using estimates on adjusted prices is a more accurate approach for focusing on idiosyncratic variance. The significance level $\alpha = 0.025$ is chosen to balance between the main specification $\alpha = 0.01$ and $\alpha = 0.05$.

[29]This value is slightly smaller than the average idiosyncratic return based on FFC adjustment over the longer samples. For example, see (daily) summary for 1926 to 2016 sample in Kapadia and Zekhnini (2019). However, the jumps have a larger order of magnitude in my sample.

[30]I start the rolling procedure from the first date, that splits the sample for a firm into two-day intervals. For the few cases that feature two event dates in the same 2-day window, I use only the second day for identification and classification. If there is an even number of days between two events, so that the next event happens on the first date, I will use the one day return prio the two-days window with event and one return after, for consistencu

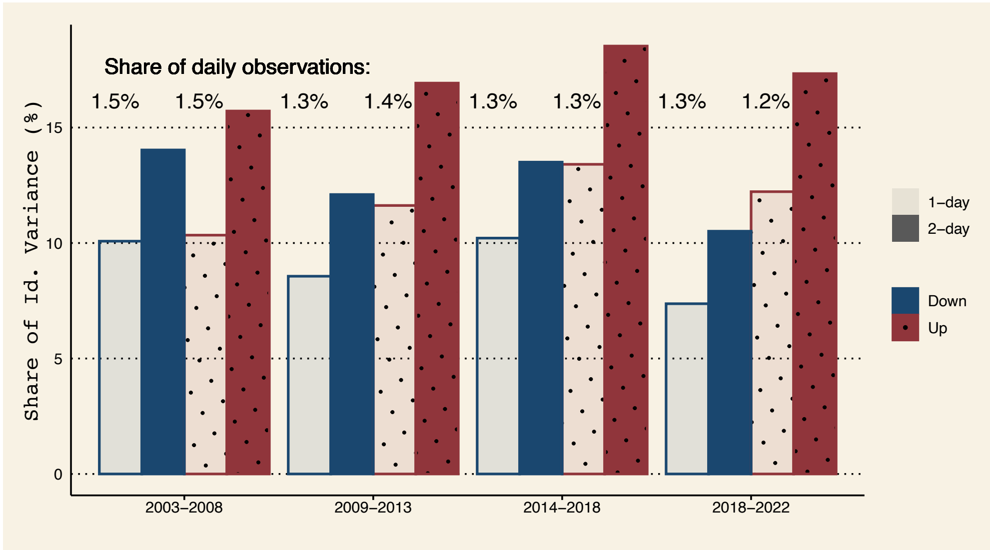

Figure 11: Idiosyncratic Variance Share by explosive events

explosion or jump,[31] calculating the two-day total return $r^{adj,2}_{s,d} = (1 + r^{adj,2}_{s,d}) \times (1 + r^{adj,2}_{s,d-1}) - 1$. Annualizing those returns the explosive two-day periods overpass contribution into the average return by two-day windows that feature the low-frequency jump being 18.58% (-16.6%) for jumps up and jumps down respectively.

The contribution of explosive dynamics to the idiosyncratic variance of stocks can be calculated based on the return decomposition of (10). The first-order effect comes from the squared sum of $r^{adj}_{s,d} \times \sum_{dir} E^{expl,dir}_{s,d}$.[32] Figure 11 reports the share of the component in the variance for four five-years periods in my sample. Specifically, I report idiosyncratic variance share upward and downward explosions defined as

$$\frac{\sum_d \sum_s (r^{adj}_{s,d})^2 \times E^{expl,dir}_{s,d}}{\sum_d \sum_s (r^{adj}_{s,d})^2}, \quad \text{for } dir \in \{\text{up, down.}\}$$

Over the years, explosive episodes have disproportionately contributed to the share of idiosyncratic variance. The total share of the variance accounted for by explosions is 20.8% (24.2% for a 5% significance level) and ranges from 19.6% to 23.6% in the four groups, with a greater share of 12.2% (13.7%) assigned to explosions up, ranging from 10.2% (11.9%) to 13.2% (15.9%). Including a day before the explosion

[31] I start the rolling procedure from the first date, which splits the sample into two-day intervals. I use only the second day to identify and classify the period for the few cases with two separate event dates in the same 2-day window. If there are even days between two events, the next event will happen on the first date of the respective second-day window. To deal with the problem, I use the one-day return before the two-day window with the event and one return after for consistency.

[32] Formally since explosive episodes are preceded by some momentum and followed by some reversal, those parts can be incorporated into the variance decomposition as well using the covariance terms. The numbers are of second-order importance after incorporating the day before and after the period into the gross-return

and considering a 2-day interval period, as we did for the analysis of average returns, increases the share of captured idiosyncratic variance to 29.5% (35.2%). It is worth noting that although the frequency of the detected episodes slowly decreases, the contribution to idiosyncratic variance does not. I must mention that the difference is more pronounced in the most recent years and when relaxing the significance level of detection, pointing to the less short nature and more difficult-to-detect nature of episodes in recent years.

The impact on idiosyncratic variance is notably more pronounced for small and medium stocks. As a consequence, the value-weighted share attributed to explosions reaches 15.2% (18.4%). Specifically, large stocks exhibit a lower susceptibility to upward explosions than small and medium stocks. In contrast, the incidence of downward explosions remains relatively consistent across the size groups and the years.[33] This observation aligns with the mechanism that will be discussed further with rooting in the concept of limited supply in the economy. Medium and smaller stocks typically entail higher shorting costs and lower overall supply in the market, fostering an environment conducive to explosive episodes. However, when considering downward explosions, the primary challenge lies in obtaining additional liquidity (cash), a hurdle that should not differ when trading medium or large stocks.

5.4 Other Descriptives of Explosiveness

The previous sections have detailed explosive events in stocks, emphasizing their connection to liquidity, time-series properties, average returns, and volatility. The pivotal question that remains is: Which firms are prone to being explosive? To address this, I will leverage a comprehensive set of 34 firm characteristics, including main variables from Freyberger et al. (2020).[34] This set is commonly employed in recent literature that explores the cross-section of returns in conjunction with firm characteristics (e.g., Kelly et al. (2017), Kozak et al. (2020)). The included variables encompass market capitalization (mktcap), book-to-market ratio (bm), factor exposures (betamkt, betahml, betasmb), leverage (lev), momentum (mom), and others detailed in the Appendix[35]. These characteristics undergo week-by-week cross-sectional transformation using *CS-normalization*[36], which uniformly maps them into interval $[-0.5, 0.5]$, similar to the approach in Kelly et al. (2017).

To analyze the impact of all those variables jointly, one would need to reduce the sample further, as

[33] Additional details are available in the Appendix.

[34] These are the firm characteristics I could replicate for the study. For example, Kelly et al. (2017) use 36 firm characteristics. In comparison, I exclude capital intensity, the ratio of change in property, plants, and equipment to the change in total assets, fixed costs-to-sales, the ratio of sales, and general administrative costs to sales, but include SMB and HML exposures.

[35] Refer to Table **??** in the Appendix for a comprehensive list.

[36] Refer to the procedure outlined in the Appendix.

some details of the firm's balance sheet or previous stock performance, necessary for the estimation of the characteristics, are missing. This introduces an additional bias toward the largest and most liquid firms, amplifying the effect of the previous filters working in the same direction. The total sample for this analysis is 7,138,967 firm-date observations[37]. For further stability, I also classify a week per stock as an explosive up (down) week if it features at least one explosive up (down) detection date, forming 1,575,707 week-firm observations. This approach allows for reducing the level of *imbalance*[38] in the data, increasing the rate of explosion detection as a dependent variable, and avoiding extra noise when fitting the models that describe explosiveness via firm characteristics.

Figure 12 illustrates the *unconditional* relationship between explosiveness up detected at a 2.5% significance level[39] and each firm characteristic. It shows the percentage of firms, sorted by a particular characteristic at the start of an estimation week, that are anticipated to exhibit explosive behavior thereafter. The first observation is that, for most firm characteristics, the variation in explosiveness is very modest and stays around 6.3% across the majority of characteristic groups. Explosiveness up never goes below 3%, reaching a 5% threshold for only a few groups. Exceptions typically occur for groups with extreme firm characteristics, revealing a non-linear nature of the explosiveness association with the firm characteristics. The most pronounced non-linear dependencies can be observed in characteristics related to the past performance reflected in standard momentum (*mom*), intermediate momentum (*intmom*), long-term reversal (*ltrev*), closeness to 52-week high (*w52h*).[40] The strongest monotonic relationships observed in the data are when firms occur sorted by market capitalization (*mktcap*), total assets (*assets*), and return on net operating assets (*noa*), indicating that smaller firms are more likely to be explosive. Similarly, firms with higher bid-ask and idiosyncratic volatility are more likely to be explosive, though the former has lower explosiveness for extremely high values. Notably, firms with the lowest (highest) turnover are less (more) explosive up, but for other turnover levels, explosiveness up stays flat around 6.3%.

In a similar vein, Figure 13 illustrates the *unconditional* connection between explosiveness down and various firm characteristics. In contrast to explosiveness up, explosiveness down shows a weaker association

[37]The reported number is for the sample based on estimates with respect to no adjustment specification. Using specifications with adjustments reduces the size of the sample slightly.

[38]The imbalance term is related to machine learning literature, where in the categorization problem, the imbalance dataset is referred to as the case when one category (no explosion) is significantly more likely than another (explosion).

[39]I use no adjustment and $k = 1$.

[40]Momentum (mom) is a cumulative return from twelve months before the explosiveness prediction to two months before. Intermediate momentum is a cumulative return from twelve months before the explosiveness prediction to seven months before. Closeness to the 52-week high is the ratio of the stock price at the end of the previous calendar month and the previous 52-week high price.

with variables reflecting the size of firms and exhibits a more uniform pattern across most firm characteristics. Specifically, only firms with the smallest market capitalization seem less explosive. Additionally, stocks with the largest bid-ask spread and the smallest turnover are notably less explosive down. The momentum/reversal characteristics display a similar pattern, resembling the "smile" observed for explosiveness up, suggesting that an extreme return over different horizons precedes explosion down.

Given the observed (non-linear) relationship between firm characteristics and explosiveness, alongside the strong correlation among firm characteristics themselves, determining the distinct contributions of each individual firm characteristic is a challenging task. To partially address this issue, I use a variety of econometric and machine learning methods to understand the influence of these characteristics while simultaneously controlling for other predictors. In this section, the primary focus is on the linear relationship between explosiveness and firm characteristics, while the Appendix presents additional results about the unveiling of non-linear relationships.

I start the analysis by employing logistic regression, wherein the weekly explosion detection event is regressed against the normalized firm characteristics, presenting the results in Tables 14 and 15. Logistic regression proves to be suitable for capturing the first order *linear* dependencies. The CS-normalization ensures a comparable distribution of the explanatory variables, facilitating a straightforward interpretation of the logistic regression coefficients. These coefficients represent both the quantitative change in explosiveness associated with a characteristic's variation and the relative "importance" of the firm characteristic in the specification.[41]

The first column of Table 14 presents the coefficients derived from the logistic regression model of the explosion indicator against the normalized firm characteristics. Enclosed within parentheses are 90% pivot bootstrapped confidence intervals.[42]. Notably, the most influential explanatory variable is market capitalization with a coefficient of -0.29. This suggests a 1.70% decrease in detection probability solely due to the size when transitioning from the smallest to the largest firm. Following in economic significance is turnover, with a coefficient of 0.17, implying a 1% increase in detection probability. However, as illustrated in Figure 12, this effect appears to be primarily driven by firms with the smallest turnovers.

Additional notable contributors identified through logistic regression encompass previous performance and volume-based factors, including momentum (-0.18 with a -1.05% probability change), standard unexplained volume (-0.15 with a -0.88% probability change), short-term reversal (-0.08 with a -0.47% probabil-

[41] Here, *importance* denotes the variable's contribution to explaining the optimized log-likelihood in the context of logistic regression.

[42] The bootstrap procedure entails 1,000 draws. The utilization of bootstrapped confidence intervals is adopted for consistency with the confidence intervals obtained for penalized logistic regression.

ity change), intermediate momentum (0.14 with a 0.82% probability change), and closeness to the 52-week high (0.11 with a 0.64% probability change). Notably, these factors operate in different directions. They are likely to offset each other, as indicated, for instance, by the correlation between intermediate momentum and momentum (0.71) and between turnover and standard unexplained volume (0.25) in the sample. Furthermore, as depicted in Figure 12, the effects for all variables exhibit non-linear patterns.

The idiosyncratic volatility exhibits predictive power for explosiveness (0.10 with a 0.59% probability change), suggesting that stocks with higher volatility are more likely to be detected as explosive up. This observation is quite remarkable as, mechanistically, the detection of explosions with SADF-based procedures should be more challenging for stocks with higher volatility. This difficulty arises because the denominator of the rolled ADF procedures, reflecting the standard error for the estimated explosive coefficient in (4), increases with heightened volatility. Simultaneously, it is essential to acknowledge that idiosyncratic volatility (*idiovol*) is derived from lower-frequency daily data, as detailed in Section 5.3, and may not accurately reflect high-frequency volatility. The mechanical association highlighted above likely underlies the observed opposing effect of market beta (-0.12 with a -0.70% probability change), indicating a decreased likelihood of explosive events for stocks with more prominent exposure to the market.

Firm characteristics associated with balance sheet data also emerge as significant predictors of explosiveness. This encompasses Tobin's q (*q*; 0.10 with a 0.59% probability change), return on equity (*roe*; 0.11 with a 0.64% probability change), return on assets (*roa*; -0.09 with a -0.53% probability change), and book-to-market ratio (*bm*; -0.09 with a -0.53% probability change). Note that some of the reported coefficients may not be entirely robust to regularized versions of logistic regression. To illustrate this, I present alternative specifications in additional columns of Table 14, employing penalized logistic regression, commonly referred to as *Elastic Net*.[43] I employ three specifications incorporating Ridge (*L2*) penalty, Lasso (*L1*) penalty, and the combined average of these penalties, referred to as Elastic Net with $\alpha = 0.5$. The hyperparameter λ, controlling the strength of the penalty, is tuned to minimize likelihood loss through a cross-validation procedure.[44]

[43] Compared to the standard logistic likelihood optimization, $\min_{\beta_0,\beta} \frac{1}{N}\sum_{i=1}^{N} l\left(y_i, \beta_0 + \beta^T x_i\right)$, where $l\left(y_i, \beta_0 + \beta^T x_i\right)$ is the standard logistic cross-entropy based on indicator y_i (indicator of detection), x_i is the set of firm characteristics, and β-s are fitted parameters, the optimization in elastic net is augmented by penalty for overfilling estimated coefficients beta $\lambda\left[(1-\alpha)\|\beta\|_2^2/2 + \alpha\|\beta\|_1\right]$, resulting in problem $\min_{\beta_0,\beta} \frac{1}{N}\sum_{i=1}^{N} l\left(y_i, \beta_0 + \beta^T x_i\right) + \lambda\left[(1-\alpha)\|\beta\|_2^2/2 + \alpha\|\beta\|_1\right]$. See details in the Appendix.

[44] A 10-fold cross-validation approach is employed, wherein the sample is randomly divided into ten subgroups. For each group G_j, $j = 1\ldots,10$, estimates are derived on the residuals of the remaining nine groups G_{-j} to calculate the sum of ten cross-entropy results. The sum is minimized over λ to determine the hyperparameter λ used (refer to Hastie et al. (2009) for details and Stone (1974) as the original reference).

The Ridge regression notably diminishes the coefficients on market capitalization (-0.14 with a -0.82% probability change), momentum (-0.09 with a -0.53% probability change), book-to-market (-0.06 with a -0.35% probability change), and turnover (-0.11 with a -0.64% probability change). Relatively minor changes are observed in standard unexplained volume (-0.11 with a -0.64% probability change), market beta (-0.10 with a -0.59% probability change), and idiosyncratic volatility (0.08 with a 0.47% probability change). The coefficient on return on net assets (rna) even experiences a slight increase to 0.06. Despite these variations, most coefficients become smaller and are not economically or statistically significantly different from zero.

The Lasso and Elastic Net specifications yield coefficients that are comparable to the original non-penalized logistic regression, with a noticeable distinction. In these penalized models, return on assets (roa), return on equity (roe), assets, Tobin's q, sales to price ratio—variables likely compensating each other—are effectively set to zero. Similarly, closeness to the 52-week high, momentum, and intermediate momentum witness significant reductions in magnitude. The most pronounced effects are observed for classical characteristics such as market capitalization, book-to-market ratio, earnings-to-price, and market beta. Two other substantial contributors, turnover and standard unexplained volume, continue to exert varying impact, potentially offsetting each other within the sample.

The coefficients for the non-penalized and penalized logistic regressions for explosiveness down are comparably presented in Table 15. Without delving into extensive detail on each coefficient, a notable distinction arises in the reversal of the sign and subsequent reduction to zero for the market capitalization coefficient with the introduction of the penalty. Additionally, bid-ask spread (-0.20 with a -1.36% probability change), capital turnover (0.22 with a 1.49% probability change), and operating leverage (-0.20 with a -1.36% probability change) assume more pronounced roles.

The principal contributors to explosiveness down seem to be momentum (0.70 with a 4.80% probability change), closeness to the 52-week high (-0.53 with a -3.62% probability change), and intermediate momentum (-0.20 with a -1.42% probability change). Similar to the previous analysis, the substantial economic magnitudes of these contributors are likely to offset each other. Nevertheless, the coefficients remain both economically and statistically significant after the application of regularization, underscoring the closer association of explosiveness down with the past performance of the stock.

The preceding analysis underscores the pivotal role of firm characteristics in relation to the explosive behavior of stocks. In conjunction with firm fundamentals, past performance serves as a predictor for explosiveness, highlighting the need for variable control in subsequent analyses. The introduction of regularization highlighted potential fragility in the dependencies, signaling susceptibility to overfitting. However, the ultimate conclusion reaffirms the robustness and significance of size variables, momentum/reversal variables, and variables reflective of past volume.

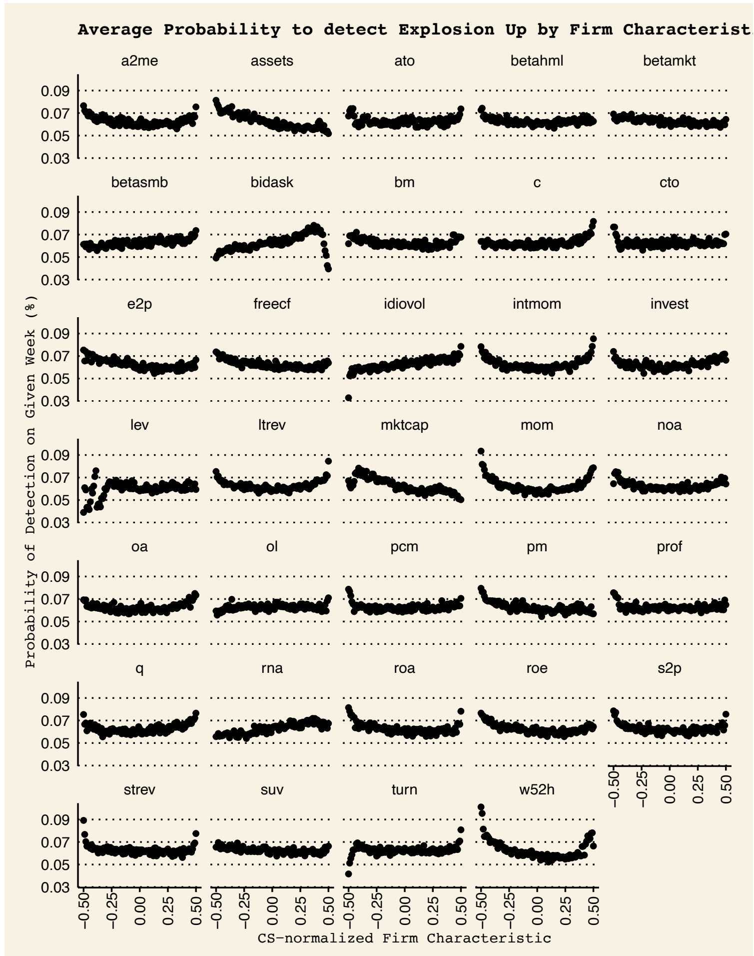

Figure 12: Explosion Up Detection by Firm Characteristic

This figure illustrates the average detection rate for explosions up aggregated by calendar weeks for stocks sorted by specific firm characteristics. The considered firm characteristics are from a set described in Freyberger et al. (2020). The firm characteristics are normalized to a range from -0.5 to 0.5 at the beginning of each estimated week to ensure a uniform distribution within the interval (CS-normalization). Based on these values, each week, firms are further sorted into 100 groups, and the explosiveness rate is measured and averaged across all firms classified into a group and all weeks. The (100) averages are reported on the Y-axis.

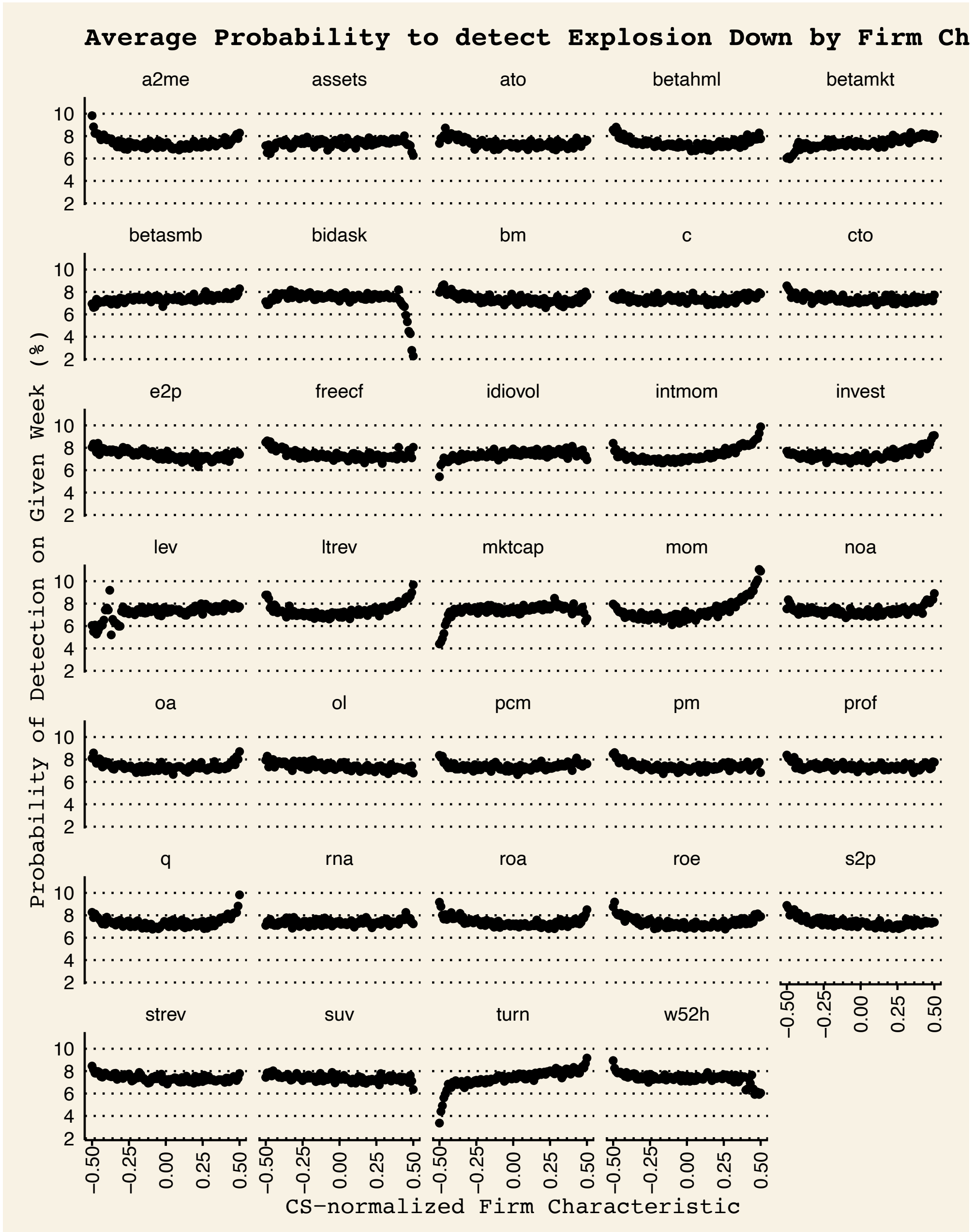

Figure 13: Explosion Down Detection by Firm Characteristic

This figure illustrates the average detection rate for explosions down aggregated for stocks sorted by specific firm characteristics. The explosiveness down is an indicator of detecting explosion down on at least one day in a calendar week. It is averaged across the stocks sorted into one of hundred groups per firm characteristic. The considered firm characteristics are from a set described in Freyberger et al. (2020). The firm characteristics are normalized to a range from -0.5 to 0.5 at the beginning of each estimated week to ensure a uniform distribution within the interval (CS-normalization). Based on these values, each week, firms are further sorted into 100 groups, and the explosiveness rate is measured within the group and across all weeks. The hundred averages per firm characteristic are reported on the Y-axis.

6 Short Squeezes

The previous sections showed that explosive behavior is a widespread phenomenon in individual stocks. The time-series property of the individual stocks did not attract that much attention until recently, when the market experienced a bunch of outstanding exuberant episodes in such stocks as GME, AMC, and in other broadly referred as "Meme Stocks". Not only did the short squeezes attract attention of social media attention but also pushed market aggregate volatility to histroical highs.

In October, 2022, the U.S. Securities and Exchange Commission (SEC) issued Staff Report[45] aiming to describe the market environment for the explosive episodes in January 2021 mentioning that *"the underlying causes of the meme stock phenomenon that are unrelated to market structure are a subject of speculation that is beyond the scope of this report."*

The paper's attention to short squeezes is highly relevant since they can be seen as amplified versions of "regular" explosive episodes and are often driven by similar market forces. The models presented in Sections 7 and 8 leverage some of the underlying mechanisms that are likely to contribute to short-squeeze events, drawing on a similar narrative of apparent forced short interest liquidation, which are particular cases of *inelastic demand* in the models.

6.1 Squeezes in January 2021

On the last week of January 2021, the market observed a sharp increase in the price of the Game Stop, Nokia, and AMC stocks. Starting January 21, the Game Stock price grew up from $39 to $347 though no positive news of a similar scale about the company's fundamental came over the period. The sharpe increase was broadly recognized as a bubble and firstly attributed to retail investors who blindly continued purchasing the stock despite the extreme price driving the price to its highs. On January 28, a popular trading platform Robinhood made the stocks unavailable for purchases later describing it by financial requirements and capital obligations[46]. Even this measure that effectively limited stocks' demand had a very modest effect on the price.

There are some aspects that might make the "bubble" different from many other bubbles studied in the literature. Fist of all, to big extend the life expectancy of the "bubble" was predetermined and quite short. The last week of January was not a random week for the sharpe price increase. January 29 is the settlement date for the options on a Game Stock. There were some indicators that a huge short position must be closed

[45] Staff Report on Equity and Options Market Structure Conditions in Early 2021

[46] https://blog.robinhood.com/news/2021/1/28/an-update-on-market-volatility

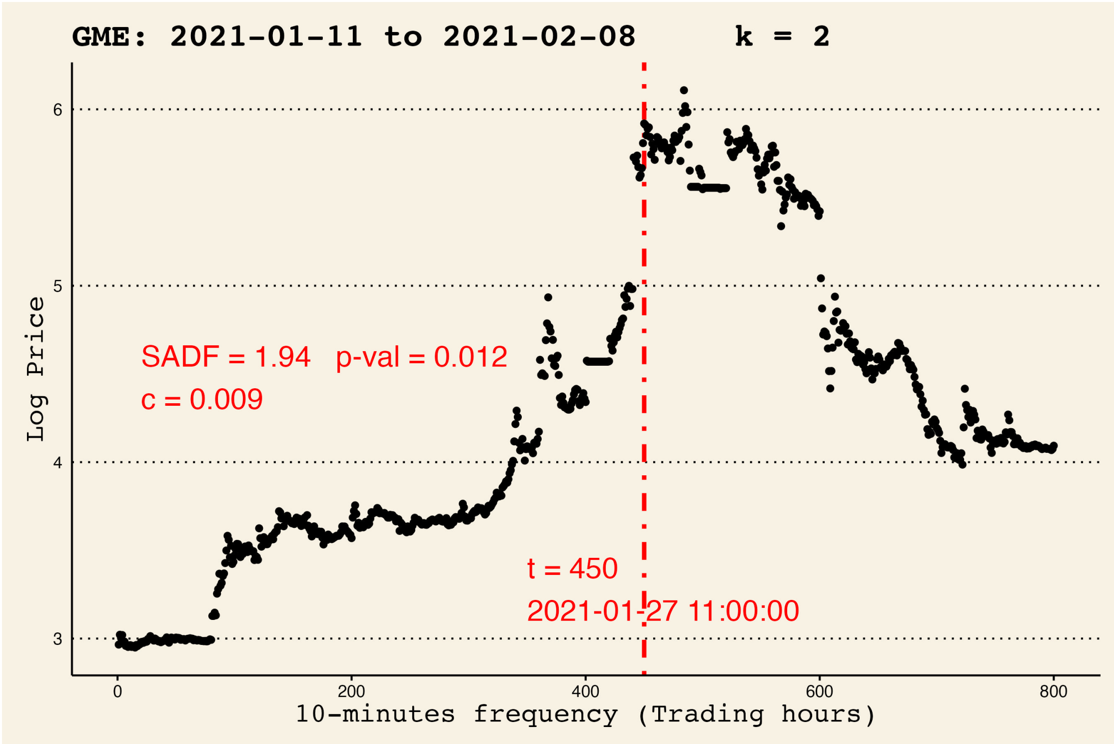

Figure 14: GME "Squeeze" in January 2021

before the date, and the expectations were that demand on the stock is almost inelastic in the stock price. That partially rationalizes the bubble as a chance to sell stock during a short-squeeze event, but also sets the limits for the life-expectancy of the bubble: after the short-squeeze the stock demand must return to price-sensitive level.

6.2 Empirical observations with GME squueze

Figure 14 illustrates the log-price dynamic of GME over twenty business days period since January 11, 2021. The supremum-ADF test identifies a bubble at a 1.2% significance level around the stock's peak on January 27. Notably, the stock exhibits highly volatile behavior, characterized by convex price growth since January 13, starting with a significant price jump. The date of January 13 is particularly noteworthy as it marks the dissemination date for short interest data, revealing a reported ratio of short interest to shares outstanding at 1.02. Although this short interest ratio was not the highest among recent disseminations, the fact that it did not substantially decrease indicated to traders the potential difficulties in liquidating and covering the significant short interest position. This short interest was also reflected in the high outstanding open interest of options, which required delta hedging to manage risk. The delta hedging strategy was dynamic

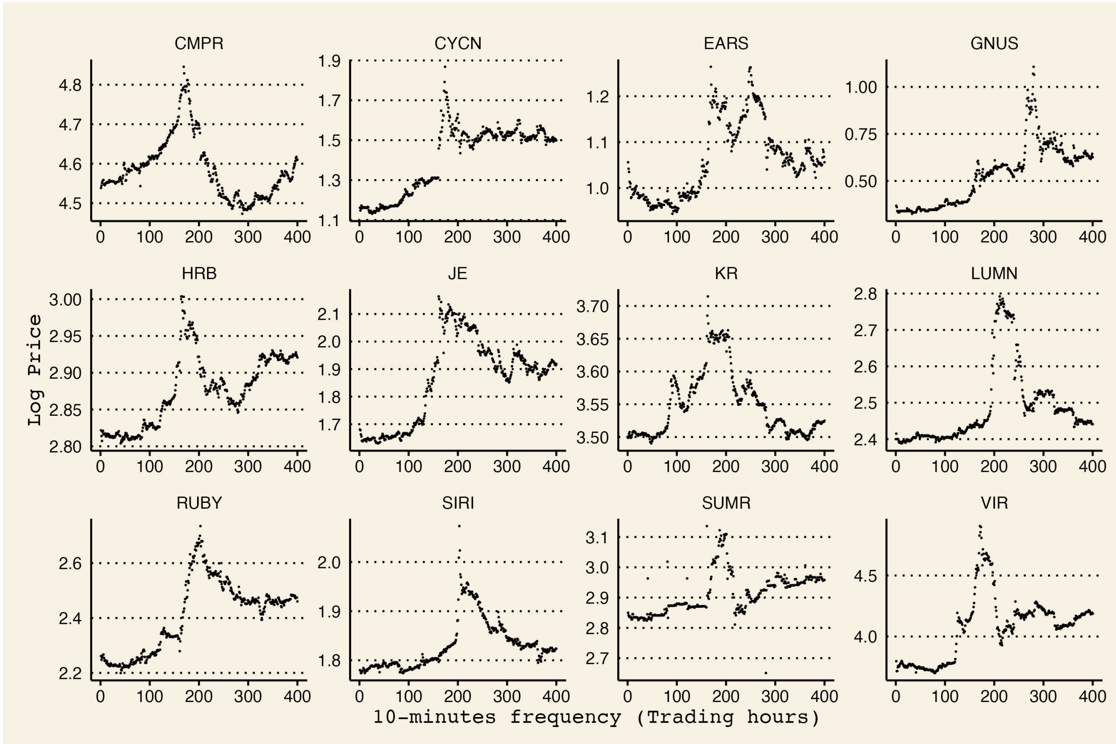

Figure 15: Some other "Squeezes" in January 2021

and adjusted as the stock price evolved effectively increasing outstanding short interest.

6.3 Model Explanation

At the same time, the equilibrium dynamic that can maintain the high prices with many potential sellers is unclear. In a classical model with many *rational* sellers of the same good, the price converges to a fundamental value of the good in any Nash-type equilibrium.[47] Hence, one must assume a very strong coordination power among the investors that would allow keeping the prices high. Even if the coordination is possible, say via platforms such as Reddit, the explosive price dynamics requires some explanation: why does not price jump immediately to some large “cooperative” price? If the price is driven by *irrational* investors only, then one might justify why big institutional investors do not exploit the abnormal prices by selling off their holdings.

The explanation proposed in this paper revolves around the existence of a strategic equilibrium that can lead to market explosiveness even with rational traders. Central to these events is the presence of *expected inelastic demand* for the asset. The amplifying mechanism responsible for generating explosiveness lies

[47] This will be discussed in more detials in the model section (Section 7) when studying the symmetric equilibrium.

in the uncertainty regarding the size of demand. Sellers must carefully consider when to sell due to their limited endowment of the asset and potential challenges in immediately borrowing stocks for further sale. Given the impracticality of trading all at once, the dynamic trading results in explosive price growth.

In the context of short squeezes, the inelastic demand arises from the need to cover an option position or deliver stocks to close short positions. However, the precise magnitude of this demand remains uncertain. It is challenging to determine how much short sellers will require to successfully cover their positions or at what point these sellers might default, thereby abruptly ending the inelastic demand.

The argument build around the role of expected inelastic demand is formalized in the next section. In Section 8, the mechanism is augmented by apparent presence of insiders that formalize the explosiveness as a price-discovery object as well.

7 Model with No Insiders

Setup

Consider an environment with one asset and two risk-neutral agent types: an inelastic impatient buyer and an infinite number of sellers. The sellers, denoted and indexed by s, are identical, endowed with a unit of an asset with a *reservation value* V. Their total mass is $\overline{\mathcal{S}}$, $\overline{\mathcal{S}} \in \mathbb{R}^+$, $0 \leqslant s \leqslant \overline{\mathcal{S}}$. There is no uncertainty about V. A buyer faces an *inelastic* demand for the asset, with a size of $\tilde{B} \geqslant 0$.This demand follows a continuous distribution with a cumulative distribution function (CDF) Φ_B over the interval $[0, \bar{B}]$, $\bar{B} \in \mathbb{R}^+ \cup \infty$ and enters the market to trade. Assume that the CDF function for the demand is twice differentiable, $\Phi_B \in \mathcal{C}^2$. The demand is unknown by sellers and is a private information of the buyer. The buyer incurs a marginal cost of not fulfilling their demand by a required unit of asset, $Y \in \mathbb{R}^+ \cup \infty$. Y is effectively the maximum price that sellers can "squieeze" out of the buyer. Assume that the cost is higher than the reservation value incentivizing the buyer to trade, $Y > V$.

Trading takes place in two steps. First, sellers choose the price, p_s, they want to sell with. Second, the buyer choses the set of sellers to buy from $\mathcal{B}$, with total mass B, $B = |\mathcal{B}|$. The terminal utilities of the sellers are

$$u_s = (p_s - V) \times \mathbb{I}\{s \in \mathcal{B}\}$$

The terminal utility of the buyer is

$$u_B = -\int_{s \in \mathcal{B}} p_s ds - [\tilde{B} - B]_+ \times Y.$$

The sellers act in the environment of uncertainty about the size of demand. Let us study Bayesian Nash Equilibria with pure strategies of buyers in the two stage model.

General Form

Without loss of generality, focus further only on equilibria where the sellers are sorted by the prices they submit in the first step: $p_s \geqslant p_{s'}$ if and only if $s \geqslant s'$, and the buyer randomly picks the seller to buy from when given similar prices and does not need to buy all supplied assets at given price. In other words, the sellers form a "queue" based on their indices to decide who will sell at higher price. The buyer's problem, by design, is trivial. They purchase the asset from sellers that require a lower price than their competitors and a price lower than Y. Assume that all sellers meet the latter condition in equilibrium, $B = \min\{\tilde{B}, \overline{\mathcal{S}}\}$, with the maximum price paid for an asset $p(B)$. Thus the buyer goes over the "queue" and buy until they satisfy the inelastic demand. The sellers' actions can feature multiple equilibria described in the following proposition.

Proposition 1. *Two types of Nash Equilibria exist in the model. The unique symmetric equilibrium with $p^* = V$ exists if $\bar{B} < \overline{\mathcal{S}}$ and only if $\bar{B} \leqslant \overline{\mathcal{S}}$. The asymmetric equilibria with continuous and non-decreasing p_s exist if $\overline{\mathcal{S}} < \bar{B}$ and only if $\overline{\mathcal{S}} \leqslant \bar{B}$. For those equilibria, $p(\overline{\mathcal{S}}) = Y$ and $p(0) = Y \times (1 - \Phi_B(\overline{\mathcal{S}})) + V \times \Phi_B(\overline{\mathcal{S}})$, if $Y < \infty$. The equilibrium is unique if $Y < \infty$.*

Let us discuss the main steps getting the result and develop reasoning behind possible explosiveness in the asymmetric equilibria deriving the sufficient condition. Focus on the case, $Y < \infty$. See the Appendix for discussion of $Y = \infty$, where an infinite set of asymmetric equilibria arises.

The most straightforward equilibrium is the symmetric one where all sellers sell at the same price, denoted as p^*. Given that the size of inelastic demand is unknown and can be as small as zero, sellers have an incentive to compete to sell first. Deviating to a price slightly lower, say $p' = p - \varepsilon$, is always beneficial, as long as $p^* \geqslant V$, since it guarantees selling. This *Bertrand-type* competition drives the price to converge to $p^* = V$ forming the unique symmetric equilibrium in the setup.

However, when $\bar{B} > \overline{\mathcal{S}}$ (indicating there could be less sellers than the true demand), the symmetric equilibrium ceases to exist. In this case, since $p^* = V$ guarantees zero profit for all sellers, any seller can find a profitable deviation, selling at a price $Y > p' > V$ with non-zero probability, resulting in non-zero utility. In this case, the symmetric equilibrium is replaced by the asymmetric equilibrium.

Therefore, when there are enough sellers to fulfill the demand, the inelastic nature of the demand may not lead to any price impact from the liquidity trader. To develop an equilibrium where price impact is observed, one must allow all traders to have a chance to sell.

Consider the case where $\bar{B} \geqslant \overline{\mathcal{S}}$ and look for asymmetric equilibria that, ex-ante, guarantee the same profit for sellers making them indifferent between chosen prices maintaining Nash Equilibrium. Using a similar argument to the symmetric case, one can show that it is impossible to have an equilibrium where a non-zero mass[48] of sellers aim for some price $p > V$ since the demand distribution is continuous. If the mass of sellers concentrates around the price $p = V$, then any of these sellers would find it advantageous to offer a higher price undercutting the others with $p_s > V$ (who must exist in the asymmetric equilibrium). This strategy allows them to make a non-zero profit on average. Consequently, this situation leads to a situation where every seller in the equilibrium can expect to earn some profit, resulting in a price level of $p_0 > V$.

Consider the asymmetric equilibrium where no mass is concentrated at a given price $V < p \leqslant Y$. Denote mass *density* of sellers aiming p by $m(p)$, $\int_V^Y m(p)dp = \overline{\mathcal{S}}$. Note that it is impossible to have $m(p) = 0$, for $p \in (p', p'')$, if $m(p') > 0$ and $m(p'') > 0$. Otherwise s with $p_s = p'$ can find a profitable deviation[49]. That means that $m(p)$ has support that is a connected set.

Denote the lower and upper limits of the price interval by $\underline{p}$ and $\bar{p}$ [50] respectively. In equilibrium, all sellers must make the same *endogenous* profit π^*. This indifference condition pins down the mass of sellers $m(p)$:

$$\pi_s = \pi^* = \mathbb{P}\left(\tilde{B} \geqslant \int_{\underline{p}}^{p_s} m(p)dp\right) \times (p_s - V) = \left[1 - \Phi_B\left(\int_{\underline{p}}^{p_s} m(p)dp\right)\right] \times (p_s - V) \tag{11}$$

Therefore, each seller, considering the actions of others, selects a price that strikes a balance between increasing their profit $(p_s - V)$ and the probability of achieving this profit, which is determined by the density of sellers offering a lower price. Equation (11) establishes the connection between the mass of sellers $m(p)$, the distribution of expected demand Φ_B, and the equilibrium profit level π^*.

$$\int_{\underline{p}}^{p_s} m(p)dp = \Phi_B^{-1}\left(1 - \frac{\pi^*}{p_s - V}\right)$$

First, note that the expected profit defines the lowest suggested price: $\underline{p} = p_0 = \pi^* + V$. Since demand is completely inelastic, sellers can submit an unlimited price as long as it does not exceed Y and does not violate the break-even condition for other sellers in equilibrium. For an outside observer, the equilibrium looks like an immediate jump in price when the inelastic buyer just enters the market. Second, note that the maximum price must be the marginal cost of the buyer, $\bar{p} = Y$. Otherwise, the "last" sellers in the "queue" have a profitable deviation selling at Y. Those two conditions effectively set the starting and terminal prices if $Y < \infty$:

$$\pi^* = (Y - V) \times (1 - \Phi_B(\overline{\mathcal{S}})).$$

[48] That means no equilibrium with infinite mass density of sellers.

[49] They would be strictly better off setting price $p'' - \varepsilon$ for sufficiently small ε.

[50] $\bar{p}$ might be infinite.

and

$$\underline{p} = Y \times (1 - \Phi_B(\overline{\mathcal{S}})) + V \times \Phi_B(\overline{\mathcal{S}}).$$

It is important to note two key observations. First, the equilibrium profit and the lowest available price both strictly increase with the marginal cost of the sellers and the probability that the inelastic demand overshoots the total supply of the asset in the economy. Second, the "initial," i.e., the lowest price, is determined by the weighted average of the maximum price at which buyers are willing to buy and the reservation value positively depending on both with a higher weight of the marginal cost as long as supply of asset is lower. This is an implicit consequence of having less competition in the model, even though the model assumes a competitive environment with an infinite number of sellers.

The gap between the initial price and the reservation value indicates that an external observer would observe a *jump* in the price when trading begins. This means that the price deviates from the reservation value solely based on the *expectation* of inelastic demand from buyers. A similar price jump can be observed on the date when the short interest data for GameStop stock was disseminated on January 13, 2022 (see Figure 1) and for some other cases of explosive episodes.

The mass of other prices is defined by the distribution density of Φ_B:

$$m(p_s) = \frac{\pi^* \times (p_s - V)^{-2}}{\phi_B\left(\Phi_B^{-1}\left(1 - \frac{\pi^*}{p_s - V}\right)\right)}, \qquad p_s > \underline{p}$$

Though the model is a two-period model[51] , one can envision how the price dynamics would appear from an external perspective if it were recorded while progressing through the "queue" along with the buyer. Consider $\tilde{B}$ enters the market at a constant speed C, this can be thought of as a throughput of the market, starting at $t = 0$. Then the observed price at time t (after just discussed initial jump), denoted as $p(t)$, is determined by the following equation:

$$\int_{\underline{p}}^{p(t)} m(p)dp = C \times t.$$

Taking the derivative twice with respect to t, we get

$$p''(t) = -C^2 \times \frac{m'(p(t))}{m(p(t))^3} \tag{12}$$

This expression reveals a crucial equilibrium condition for the price movement to exhibit a convex function. It states that the *endogenous* mass density, denoted as $m(p)$, must be a decreasing function. In other words, as the price increases, fewer sellers are willing to be patient and wait to sell later. The intuition

[51] A richer setup will be considered in Section 8.

behind this condition is clear: when the mass of sellers is uniformly distributed over prices, the price impact would be linear in response to the buying pressure from the impatient buyer. However, as the mass of sellers willing to provide at a lower price decreases, the probability of selling later also decreases, requiring higher compensation, which ultimately pushes the price impact higher, resulting in a convex price function.

The closed-form expression for the mass density is more complex (for details, please refer to the Appendix), but the general property is that the mass density decreases more rapidly as the tail of ϕ_B gets lighter, i.e., the average demand gets more concentrated around small values with little chance of getting extreme. The necessary condition for the path to be explosive can be reduced to the following restriction on distribution of demand:[52]

$$\frac{\phi_B'(z)}{\phi_B(z)} \geqslant -2 \times \frac{\phi_B(z)}{1-\Phi_B(z)}, \qquad 0 \leqslant z \leqslant \overline{S}.$$

Special cases of solutions: power and exponential distribution

To illustrate the idea, I derive the closed-form solutions for multiple distribution classes commonly appearing in the literature. All those distributions may result in explosive change of the price. The first is the class of power distributions for the size of demand, $\Phi_B(B) = 1-(1+B)^{1-\beta}$, $\beta > 2$.[53] Then the solution for the price gets form of,

$$p(t) = V + (Y - V)\left(\frac{1+Ct}{1+\overline{S}}\right)^{\beta-1},$$

The power parameter β plays a significant role in controlling the curvature of the price function, $p(t)$. A higher value of β leads to a steeper price increase in response to additional units purchased by the buyer, resulting in a faster price movement. Simultaneously, the initial price, denoted as $p(0)$, decreases. This initial price decrease occurs as the supply of the asset, represented by $\overline{S}$, increases. Additionally, when the marginal cost of the insider rises, the entire price curve shifts upward. This upward shift is more pronounced for the sellers who enter the market later. Thus, the power parameter β and other factors such as supply and marginal costs collectively influence the shape and dynamics of the price function.

In the example where Φ_B follows an exponential distribution, denoted as $\Phi_B(B) = 1-e^{-\beta B}$, the demand rate is memory-less. This means that each additional unit of demand dB is equally likely to occur with a probability of $\beta\, dB$. The price dynamics under this distribution result in exponential growth:

$$\dot{p}(t) = \beta C\,(p(t) - V)\,dt$$

[52] This condition guarantees negative $m'(p_s)$ derived in the Appendix (section A5.1).

[53] $2 \geqslant \beta > 1$ returns same result but the solution is not convex.

The rate of price growth is directly proportional to the parameter β. When the tail of the exponential distribution becomes thinner, meaning β increases, the rate of price growth also becomes larger. In other words, a higher β leads to faster exponential growth in the price. The other properties repeat established properties for the power distribution functions.

Finally, consider the case where the demand follows a half-normal distribution. This is in line with the commonly assumed normal distribution of signals and demands in market microstructure models. The density is considered only in the positive region as follows:

$$\phi_B(B;\beta) = \frac{2\sqrt{\beta}}{\sqrt{\pi}} \exp\left(-\beta B^2\right) \quad B \geqslant 0.$$

To maintain consistency with the previous notation, β is the inverse of two variance parameters of the corresponding Gaussian distribution. The closed-form solution for the price is given as:

$$p(t) = \frac{\pi^*}{2\Phi_{\mathcal{N}}[-Ct\sqrt{0.5\beta}]},$$

where $\Phi_{\mathcal{N}}$ is the CDF of the standard normal distribution, and $\pi^* = (Y-V)(2-2\Phi_{\mathcal{N}}(\overline{\mathcal{S}}\sqrt{0.5\beta}))$. Under this distribution, the price reacts even more aggressively to additional demand compared to the exponential case. As the variance of the Gaussian distribution shrinks and its tails become thinner, the reaction becomes more aggressive. The function $\Phi_{\mathcal{N}}^{-1}(-x)$ is increasing at a rate faster than exponential.

8 Model with Insider

Incorporating partial reversal into the model involves introducing an informed trader, the insider. In this section, I will build upon the model discussed in the previous section. To make the model more tractable, I will incorporate dynamic decision-making by the sellers as demand flows into the market. Demand can be generated by both an impatient buyer and an insider who mimics the other. Faced with uncertainty about who they are trading against, sellers continue to provide explosive prices, as discussed in the previous case, but prices now carry informational value.

8.1 Basic Setup

Consider a model featuring two types of agents: a buyer and an infinite number of sellers. Buyers can be either insiders or liquidity traders discussed in previous section. The information of buyers type is privately known but not revealed to sellers. The agents trade one asset with a reservation value V that reflects ex-ante sellers' expectation of the asset.

From the perspective of sellers, assume that there is a probability, denoted as $p_{I,0}$, that an agent with private knowledge of the true asset value $\tilde{V}$ enters the market. Importantly, the ex-ante expected value of the asset remains V. In essence, $p_{I,0}$ represents the likelihood of an insider's presence in the market and the probability that this insider knows $\tilde{V} \geqslant V$. I assume that $\tilde{V}$ is distributed over the interval $[V, \bar{V})$ with a continuous probability density function denoted as ϕ_V and a cumulative distribution function denoted as Φ_V. These functions are infinitely differentiable, and Φ_V belongs to the class $\mathcal{C}^{\infty}$.

Assume that inelastic *impatient* buyers discussed in the previous section enter the market with a probability of $1 - p_{I,0}$. They incur a limited marginal cost due to not fulfilling their demand Y. Additionally, similar to the previous setup, they face a demand of size $\tilde{B}$ with a continuous distribution. Its CDF and PDF are denoted as Φ_B and ϕ_B respectively, $\Phi_B \in \mathcal{C}^{\infty}$, and this demand has a support range of $[0, \bar{B}]$. Y and $\bar{B}$ can potentially be infinite.

Finally, each seller is endowed with a unit of the asset. The total mass of sellers and asset supply is denoted by $\overline{\mathcal{S}}$. $\overline{\mathcal{S}} \leqslant \bar{B}$ that guarantees that in case of really big demand $\tilde{B}$ the last seller has an opportunity to extract the value out of the buyer.

Trading takes place over an endogenously defined period $[0, \bar{t}]$, where $\bar{t}$ will be defined as the moment when the buyer stops buying. Over this interval, at every time t, a seller that is still endowed by a unit of the asset decides to provide the liquidity at a specific price $p_s(t)$ or wait. The buyer decides to buy Cdt additional units or stop buying, identifying $\bar{t} = t$, where C is the throughput of the market. If the buyer purchases, the transaction goes at the best available price $p(t)$.[54] The total mass of purchased asset is $B = \int_0^{\bar{t}} Cdt$. Denote the set of sellers to sell at time t by $\mathcal{B}_t$ and sellers to have sold at time $\bar{t}$ by $\mathcal{B} = \bigcup_{t \leqslant \bar{t}} \mathcal{B}_t$.

The terminal utilities of the *impatient inelastic* buyer is

$$u_B = -\int_t^{\bar{t}} p(t)db(t) - [\tilde{B} - B]_+ \times Y.$$

The terminal utilities of the insider is

$$u_I = -\int_t^{\bar{t}} p(t)db(t) + B \times \tilde{V}.$$

The terminal utilities of the sellers are

$$u_s = (p(s) - \tilde{V}) \times \mathbb{I}\{s \in \mathcal{B}\}$$

Sellers are risk neutral. Hence at time $t < \bar{t}$, the seller s's utility of submitting the best price $p(t)$ is

$$u_{s,t} = (p(t) - \mathbb{E}_t \tilde{V}) \times \mathbb{I}\{s \in \mathcal{B}_t\}.$$

[54] without loss of generality, assume that Cdt of sellers provide this price

8.2 General Form

Using a similar set of arguments that we have seen in the previous section model, the terminal sellers if trade will suggest the marginal cost of inelastic buyer price:

$$p(\bar{t}) = Y.$$

Denote the mass of sellers that sell at price $p(t)$ at time t by $m(p(t))$.The total trading volume before time t is

$$b(t) = \int_{\tau=0}^{t} m(p(\tau))d\tau, \qquad b(\bar{t}) = \overline{\mathcal{S}}.$$

The seller's challenge entails balancing several trade-offs. First, delaying the selling decision reduces the likelihood of a successful sale. Furthermore, postponing the decision can impact the probabilities of selling to an insider versus an uninformed buyer. Finally, both the future equilibrium price, which affects the expected profit gained from the buyer, and the conditional loss incurred by the informed seller who continues to buy in the market, are subject to change.

Let us formalize the tradeoffs analyzing equilibria where the price is a non-decreasing function of trading volume.[55] To get the expected profit of sellers at every stage introduce $p_I(t)$ and $p_B(t)$ the *endogenous* unconditional probability of informed insider buying the asset at time t.

$$\begin{aligned} p_I(t) &= Pr\Big[\tilde{V} > p(t)\Big] \times p_{I,0}, & p_I(\bar{t}) &= p_{I,0} \times (1 - \Phi_V(Y)) \\ p_B(t) &= Pr\Big[\tilde{B} > b(t)\Big] \times (1 - p_{I,0}), & p_B(\bar{t}) &= (1 - p_{I,0}) \times (1 - \Phi_B(\overline{\mathcal{S}})) \end{aligned} \tag{13}$$

The sum of $p_I(t)$ and $p_B(t)$ captures probability that at least $b(t)$ units of asset are purchased. The conditional probabilities to face insider and buyer conditional on the continuing trading demand at time t:

$$\begin{aligned} p_I^c(t) &= p_I(t)/(p_I(t) + p_B(t)) \\ p_B^c(t) &= p_B(t)/(p_I(t) + p_B(t)) \end{aligned} \tag{14}$$

Seller's expected profit, contingent upon the presence of demand at time t, consists of the gain to sell an uninformed buyer and the loss to the insider.:

$$\pi(t) = \underbrace{p_B^c(t) \times \big(p(t) - V\big)}_{\text{Gain from Buyer}} + \underbrace{p_I^c(t) \times E[p(t) - \tilde{V} | \tilde{V} \geqslant p(t)]}_{\text{Loss to Insider}}. \tag{15}$$

[55]That might require extra work to set the necessary conditions for ϕ_V and ϕ_B that $\dot{p}$ is always non-negative. Since the conditional probabilities of insider at given point of time may look nontrivial, it is not straightforward to show monotonicity for any set of the distributions as it would be in the model with no insider.

The indifference between selling now and a moment later requires

$$\dot{\pi}(t) = \pi(t) \times \Big(p_I^c(t) \times \frac{\phi_V(p(t))}{1-\Phi_V(p(t))} \times \dot{p}(t) + p_B^c(t) \times \frac{\phi_B(b(t))}{1-\Phi_B(b(t))} \times \dot{b}(t) \Big) \tag{16}$$

Alternatively, taking the derivative of (15) with respect to t:

$$\dot{\pi}(t) = \dot{p}(t) - \dot{p}_B^c(t) \times V - E[\tilde{V}|\tilde{V} \geqslant p(t)] \times \left(\dot{p}_I^c(t) - p_I^c(t) \times \frac{\phi_V(p(t))}{1-\Phi_V(p(t))} \times \dot{p}(t) \right). \tag{17}$$

Note that in cases where the model parameters guarantee the presence of a monotonically increasing equilibrium, with $\dot{p}$ remaining consistently non-negative, equation (16) also ensures that the profits of sellers do not change sign as time progresses. Since the equilibrium is sustainable only if sellers receive non-negative profits, this guarantees an increasing profit and also simplifies the *Trade Condition* to a requirement of non-negative profit for the terminal sellers:

$$\bar{\pi} = p_B^c(\bar{t}) \times \big(p(\bar{t}) - V\big) + p_I^c(\bar{t}) \times E[p(\bar{t}) - \tilde{V}|\tilde{V} \geqslant p(\bar{t})] \geqslant 0. \tag{18}$$

The violation of this condition, which might occur, for example, if the probability of facing an insider is too high, resembles the conditions for the No Trade theorems (Milgrom and Stokey (1982), Tirole (1982)). Once $p_I^c(\bar{t})$ goes to one, the loss to an informed trader dominates in (18), and discourages sellers to take an opposite side in the subsequent trades. If the conditions hold and a solution exists then one can inversely solve the system of ordinary differential equations backward characterizing the solution for any set of distributions and parameters.[56] I will focus further on a specific distributional case that provides a tractable closed-form solution,

8.3 Exponential distribution case

Closed Form Solution

Let us analyze the case when $\Phi_B(b) = 1 - \exp(-\beta b)$ and, conditional on buying, $\Phi_V(v) = 1 - \exp(-\alpha v)$, $V = 0$, that provides a tractable closed-form solution for the model. The solution is invariant up to shift in reservation value V and respective distribution $\Phi_V(v)$. Also, we have set that the speed of trading is linear,

$$b(t) = C \times t,$$

[56] See details in the Appendix.

and will analyze this case first by adding more degrees of freedom to it later. The linear trading speed means that the same amount of the asset is traded per unit of time. Without loss of generality, one can rescale time in terms of C so that $\bar{t} = 1$, and $C = \overline{S}$.

The *Trade Condition* (18) simplifies to

$$\bar{\pi} = \pi(\bar{t}) = \frac{e^{-\overline{S}\beta}(1-p_{I,0})Y}{e^{-\overline{S}\beta}(1-p_{I,0}) + e^{-Y\alpha}p_{I,0}} - \frac{\alpha^{-1}e^{-Y\alpha}p_{I,0}}{(e^{-\overline{S}\beta}(1-p_{I,0}) + e^{-Y\alpha}p_{I,0})} \geqslant 0$$

or

$$Y \geqslant \alpha^{-1}W\left(\frac{p_{I,0}}{1-p_{I,0}} \times e^{\overline{S}\beta}\right),$$

where W stands for a Lambert-W function[57]. Note that the condition is tractable. Sellers will sell as long as the maximum price they can "squeeze" out of the buyer is high enough to cover the costs related to the average knowledge possessed by insiders (α^{-1}) and the probability of facing them ($p_{I,0}$). Additionally, $e^{\overline{S}\beta}$ also requires that the probability of selling for the last sellers is not too small. In the Appendix, I provide the comparative statics that shows a similar dependence for $\bar{\pi}$, which strictly increases with Y and α and strictly decreases with $\overline{S}$, β, and $p_{I,0}$.

The next theorem characterizes the equilibrium under the exponential distribution assumptions.

Theorem 1. *If Trade Condition is satisfied, there exists unique (strictly) increasing $p(t)$ satisfying (13)-(16),*

$$p(t) = (c_2(t) + c_1(t)) + \alpha^{-1}W\left(\frac{p_{I,0}}{1-p_{I,0}} \times e^{-(c_2(t)+c_1(t))\alpha + \overline{S}\beta t}\right), \tag{19}$$

where

$$c_1(t) = e^{-\overline{S}\beta(1-t)} \times \bar{\pi}, \qquad c_2(t) = \frac{e^{-Y\alpha + \overline{S}\beta t}p_{I,0}}{1-p_{I,0}} \times \bar{\pi}.$$

$c_1(t)$ represents the price in the market in the absence of insiders. $p(t)$ is convex function. The sellers that sell at time t earn expected profit

$$\pi(t) = \frac{\bar{\pi}}{e^{-\overline{S}\beta t}(1-p_{I,0}) + e^{-\alpha p(t)}p_{I,0}} \times \left(e^{-\alpha Y}p_{I,0} + e^{-\overline{S}\beta}(1-p_{I,0})\right). \tag{20}$$

In the Appendix, I provide a formal derivation of the equilibrium dynamics. The core of explosiveness lies in the term $c_1(t)$, which captures the dynamics observed in the model without insiders. In this scenario, later sellers require larger compensation for the risk of waiting. The second component, $c_2(t)$, comes into play when insiders appear in the market. It is also explosive but balanced out by a new component with the Lambert function, which is not necessarily convex. The dominance of one over the other will depend on

[57] Lambert-W function is the inverse of $y \cdot e^y$, this is a monotonic, increasing and concave function over the non-negative region

the model's parameters, as discussed later. However, the combination of all three components is convex for all valid parameters that do not violate the Trade Condition. Moreover, the first two derivatives of the price with respect to time have a tractable closed form as well:

$$p'(t) = \frac{e^{\alpha p(t)}(1-p_{I,0})}{e^{\alpha p(t)}(1-p_{I,0}) + e^{\overline{S}t\beta} p_{I,0}} \times p(t) \times \overline{S}\beta \tag{21}$$

$$p''(t) = \frac{\overline{S}\beta e^{\alpha p(t)} \times \left(2p'(t) + p(t)\overline{S}\beta\right) \times (1-p_{I,0}) + \alpha e^{\overline{S}t\beta} \times p'(t)^2 \times p_{I,0}}{e^{\alpha p(t)}(1-p_{I,0}) + e^{\overline{S}t\beta} p_{I,0}} \tag{22}$$

Equation (21) shows that the growth rate of the price is primarily affected by $\overline{S}\beta$, which represents the speed at which the probability of selling to an uninformed buyer decreases. The probability of an insider, $p_{I,0}$, undermines explosiveness and, as the paper will discuss later, leads to a more considerable immediate price jump rather than an increase in price convexity. It is important to note that since the multiplier next to $p(t)$ in (21) is bounded from below, the price effectively exhibits an exponential growth rate, ultimately causing convexity.

The initial price just after the first purchase of the asset is non-zero:

$$p(0) = e^{-\overline{S}\beta} Y - \alpha^{-1} e^{-Y\alpha} \frac{p_{I,0}}{1-p_{I,0}} + \alpha^{-1} W \left(\frac{p_{I,0}}{1-p_{I,0}} \times e^{\frac{e^{-Y\alpha} p_{I,0}}{1-p_{I,0}} - e^{-\overline{S}\beta} Y\alpha} \right) \tag{23}$$

As long as the trade condition is satisfied, the difference between the first two terms remains positive and the price is strictly positive. Thus, the dynamics start with a jump in price. This is a common property of the equilibrium for different distributions Φ_V and Φ_B that reflects the existence of a non-zero spread.

The model features a deviation of the price from its fundamental value. Conditional on the arrival of an insider, the price discovery process occurs, and the price moves toward the new equilibrium price. Assuming that the insider is interested in realizing the value of the trade as soon as possible, one can make the assumption that once the purchase stops, the insider reveals the information to the public, justifying the lack of reversal in this case. While remaining agnostic about the way information is released, one may focus on the average reversal of the price in the long run:

$$\text{Reversal}(t) = \frac{p(t) - E[V|b(t)]}{p(t)}.$$

Characterization of the initial jump

In the first model without an insider, we observed that it features an initial jump to provide a non-zero profit for the first seller and keep them indifferent from deviating to higher prices. The mechanics remain the same in this model, but the adverse selection amplifies the effect. Therefore, the initial jump can be viewed as a component of the spread, reflecting both the market power of the sellers over the inelastic buyers and serving as a compensation for apparent losses to informed traders. The comparative statics for the initial jump are formulated in the proposition. The formal derivations are provided in the Appendix.

Proposition 2. *The initial jump, $p(0)$, is non-negative and bounded from below by $Y \times e^{-\overline{S}\beta}$, which corresponds to the jump in the no-insider case. It strictly increases as the probability of an insider, $p_{I,0}$, goes up, or as the average knowledge of the insider, α^{-1}, increases. It strictly decreases as the supply of the asset $\overline{S}$ goes up or the maximum price at which the buyer purchases, Y, decreases. It increases as the size of the expected inelastic demand, β^{-1}, goes up. As the number of insiders approaches the no-trade region the initial jump converges to Y.*

It is worth mentioning that as the economy becomes more populated by insiders, the sensitivity of the initial price change increases, and there is less room for explosiveness to occur. This is because the distance between the initial price, $p(0)$, and Y shrinks, making explosive episodes less likely. One can think of events such as earnings announcements, where there is a growing probability that market participants can either learn or process the information marginally faster than others. This can lead to the prediction that explosive episodes are less common than ordinary jumps in these cases.

Characterization of the price path

In the provided solution, all price paths are explosive, but there should be a way to compare the explosiveness of these events. The natural candidate for this comparison is some level of convexity in the price path. I will use the average curvature of the price path, denoted by κ^{av}, which captures the average second price derivative over the possible price path:

$$\kappa^{av} = \int_0^1 p''(t)dt.$$

The following proposition summarizes the findings regarding it in the equilibrium.

Proposition 3. *The price $p(t)$ is a strictly increasing and convex function. It increases point-wise as the probability of an insider, $p_{I,0}$, increases. In the extreme case of no insiders, we have:*

$$p(t) = e^{-\overline{S}\beta(1-t)} \times Y.$$

In the limit case where the market is primarily populated by insiders but the Trade Condition is not violated, $p(t)$ approaches a linear curve:

$$p'(t) = \alpha^{-1}\overline{S}\beta + O(1 - p_{I,0}).$$

The average curvature of the price path strictly increases as long as the probability of insiders decrease.

The result suggests that the level of explosiveness after the initial big jump, captured by the average curvature, decreases as the share of insiders grows. Moreover, the price impact also becomes almost linear after the jump. This result is interesting because the price impact in this context has a different modeling structure compared to a significant portion of market microstructure models where the price impact is defined by the strategy of hiding behind the noise traders (e.g., Kyle (1985)). In those models, linear price impact is often observed, but the situation here is quite the opposite from the perspective of sellers (market makers). In the limiting case of my model, they primarily trade against informed traders and, on rare occasions, face inelastic buyers. This model could be a good framework for studying post-jump drifting.

Numerical Example

Let us begin with a numerical example that aims to produce a realistic output consistent with the average observation from the data. We set the maximum marginal cost of the buyer to be 15%, limiting the explosion's size in terms of numbers. The probability of an insider is 2%, with an average possessed knowledge of $\alpha^{-1} = 10\%$. We set the supply to $\overline{S} = 15$ and $\beta = 0.35$. The simulated paths for price, $p(t)$, and profit, $\pi(t)$, conditional on ongoing demand, are illustrated in Figure 16. These paths exhibit the desired explosiveness. The initial price jump for this specification is approximately 20 basis points, with $p(0) = 0.23\%$. At time $t = 0.5$, the price only reaches 2.62%, but subsequent demand leads to a significant increase in the price. Figure 16 displays the probability of ongoing demand and the changing conditional probability of an insider. In this parametrization, the probability of an insider significantly increases with rising demand. This is necessary to match the appropriate relative reversal, as reported in Figure 16, which varies between 14% to 22%.

Y	α	$p_{I,0}$	β	$\overline{S}$	V
15%	10	2%	0.35	15	0

Next, let us relax the parameter $\alpha = 80$, allowing for larger probabilities of an insider, consistent with the Trade Condition, to analyze established facts regarding the jump size and curvature of the price in the limit cases. It is worth noting that by doing so, we significantly reduce the knowledge of the insider, and unless the probability of an insider is exceptionally high, it means that as time approaches $t = 1$, the conditional probability of an insider approaches zero. This dynamics is illustrated in Figure 18. The figure also highlights the non-trivial dependence of the probability to sell on $p_{I,0}$, as well as the potential non-monotonic nature of conditional probabilities when selling to an insider or an uninformed buyer.

The price dynamics confirm the theoretical findings from the previous subsection. Starting from the no-insider case, as indicated by the red line, the price exhibits pointwise increases as the probability of an insider rises. The most dramatic change occurs for the immediate jump, $p(0)$, which is a reaction to the intention to buy. In the extreme scenario with 99.99% of informed buyers, indicated by the black line, the price would jump by 9%, which is significantly more than the ex-ante expected knowledge possessed by the insider. This reflects the revealing nature of the intention to buy, as seen in the No Trade theorems. As the probability of an insider approaches one, the price path also loses curvature, resulting in an almost linear price path, indicating a constant price impact.

As the probability of an insider increases, the sellers' profit naturally decreases, eventually converging to almost zero at all times. Conversely, the relative reversal of the price change shifts from complete reversal in the no-insider case to no reversal in the case where almost all participants are insiders. Furthermore, the relative reversal generally increases as the magnitude of the price rises.

Analysis of varying proportions of insiders in the population might lead to the *incorrect conclusion* that there exists a tradeoff for the model: one might observe either pronounced explosiveness (as in the case when $p_{i,0}$ is small) or a persistent change in the price (as in the case when $p_{i,0} \to 1$). However, this is not the case; as for this specific specification, we intentionally relaxed the knowledge possessed by insiders to satisfy the Trade Condition. In this scenario, the significant demand reveals the nature of the buyer in a way that the updated belief $p_I(t)$ decreases as considerable buying pressure is realized. This behavior is depicted in the right plot of Figure 19, where $p_I(t)$ exhibits an inverse smile shape.

To illustrate that in the model, the possessed knowledge parameter takes care of the false tradeoff, let us examine a specification where the parameters replicate the base case but instead of altering the probability of an insider, I modify parameter α^{-1}. Figure 25 demonstrates that the sellers' profit and relative reversal change significantly while there is minimal alternation in the price path when α^{-1} changes from 0.01 to 0.12. The interpretation is that if the probability of an insider is relatively tiny, the sellers provide liquidity in a manner close to how they would act under the assumption of no insiders. The insider's presence for the initial sellers in the "queue" is perceived as a minor risk, and thus, there is no need to impose a significant bid-ask spread, and little initial jump is observed. As demand flows into the market, the probability of an insider increases, thereby balancing out the reversal and leading to persistent change in the price.

8.4 Discussion and possible extensions

The model simplifies many aspects of trading for tractability and to emphasize the core mechanism generating explosion. That is the expectation of inelastic demand. The assumption can be alternatively viewed as

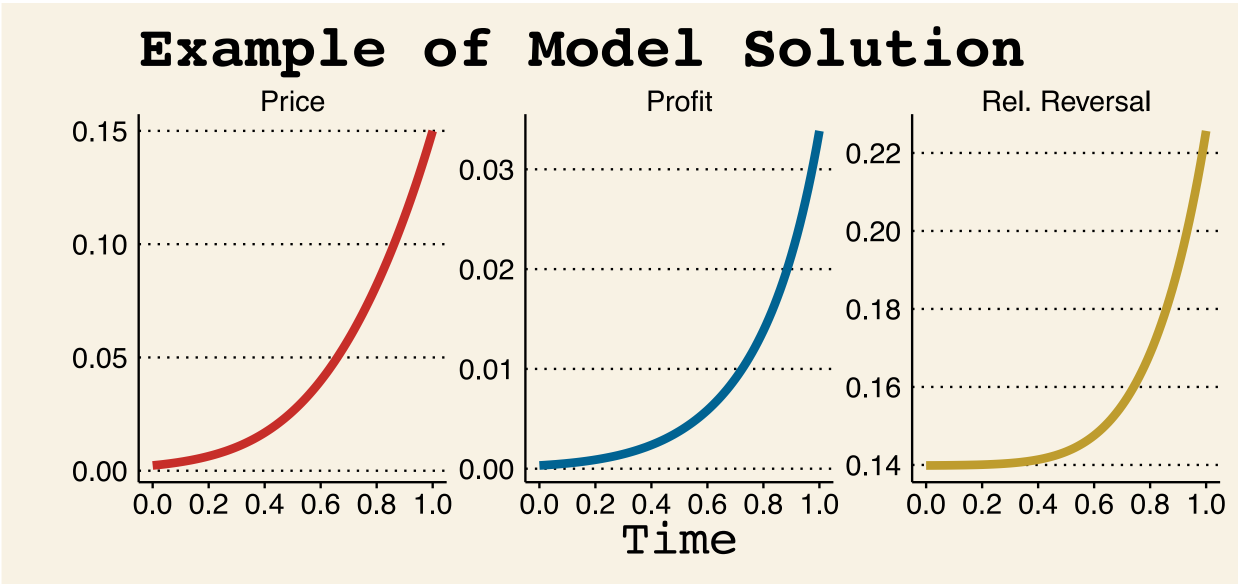

Figure 16: Price, $p(t)$, profit, $\pi(t)$, and reversal(t)

This figure illustrates the price, profit, and (expected) relative reversal path conditional on parameters set in Section 8.3.

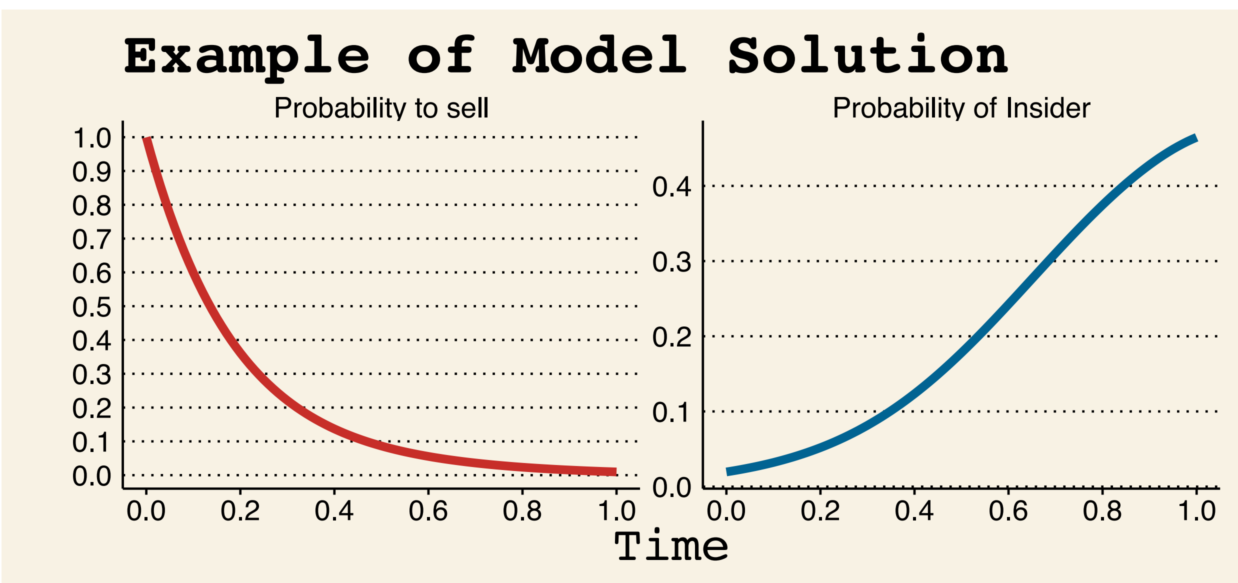

Figure 17: Probability to sell, $p_I(t) + p_B(t)$, and probability of insider, $p_I^c(t)$

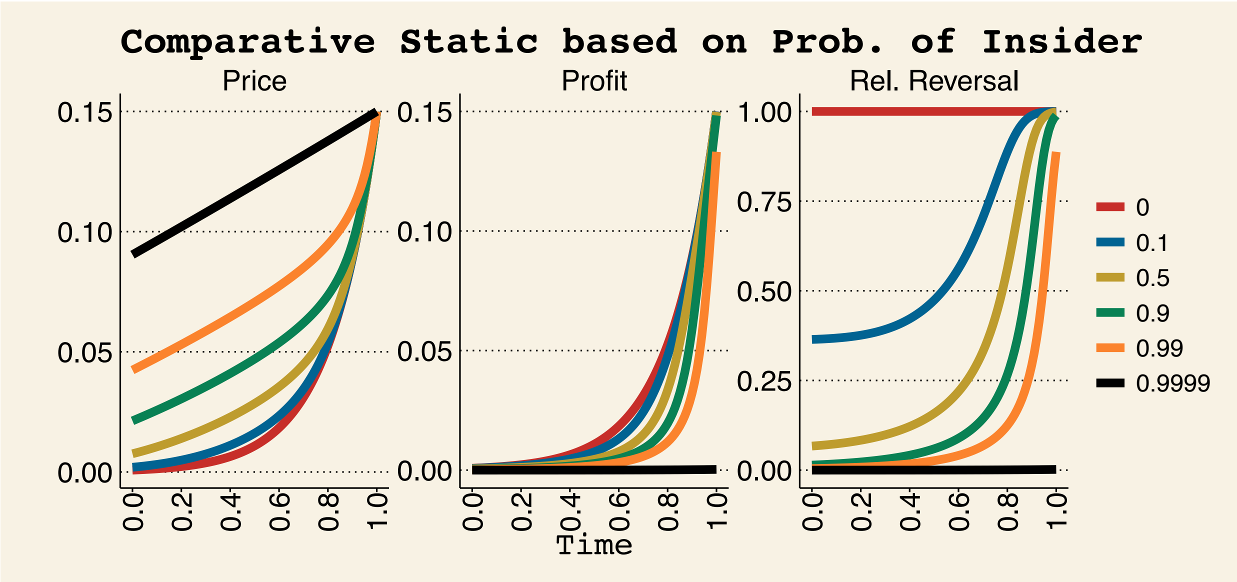

Figure 18: Price, $p(t)$, profit, $\pi(t)$, and reversal(t)

This figure illustrates the sensitivity of price, profit, and relative reversal to the probability of an insider, which is captured by $p_{i,0}$. The knowledge possessed by insider $\alpha^{-1} = 1.25\%$ is relaxed to avoid violations of Trade Condition. The other parameters replicate the suggested parameters from the numerical example in Section 8.3.

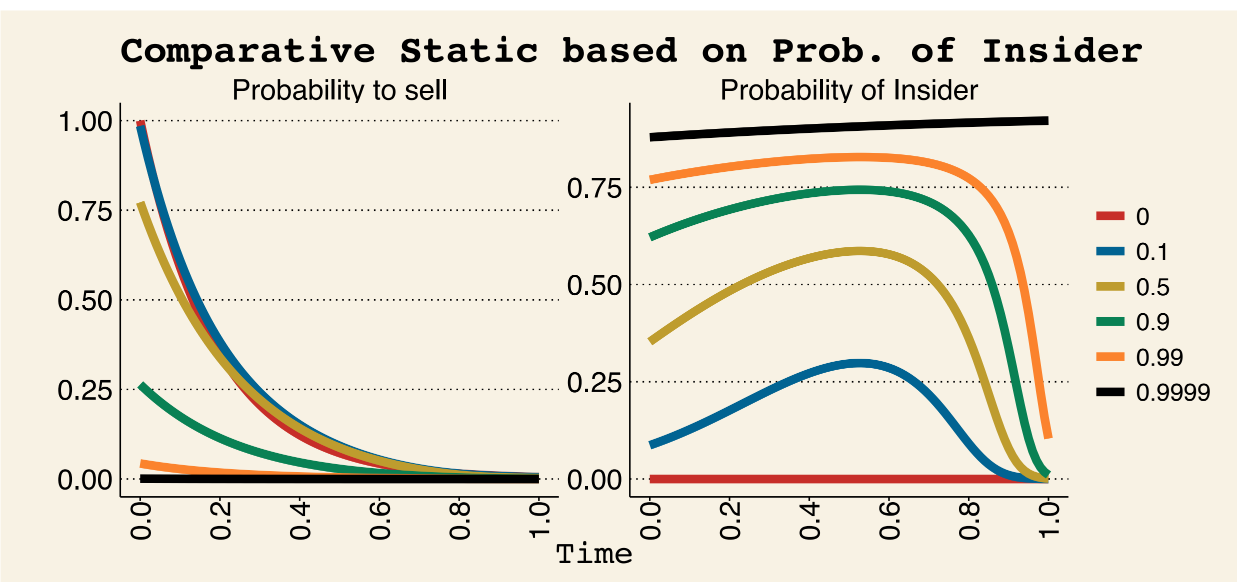

Figure 19: Probability to sell, $p_I(t) + p_B(t)$, and probability of insider, $p_I^c(t)$

the existence of correlated noise trading in the demand. In the models, this assumption is taken to extreme, though, there are ways to expand and refine it for more tractability.

Partially Uninformed Trader: Instead of a fully inelastic buyer, one could introduce a partially uninformed trader who has some information about the probability of the asset. This would introduce a non-binary type of reversal, potentially adjusting the filtering problem in the model. With a limited supply of the asset, a similar explosive mechanism as in the current model might occur, but with a more considerable initial jump.

Noise Traders: To control the trading speed, one could add noise traders who randomly enter the market and are willing to trade at the current price, providing (1) some competition to sellers and (2) some chances for hiding behind them for the buyers. This would incentivize the buyer to buy slowly and avoid ad-hoc assumptions on the throughput of the market system. To balance this out with an incentive to buy at a non-zero speed, the model could include a cost for the buyer associated with carrying non-trading.

Introduction of informative signals To enhance the model's price dynamics resemblance to actual data and incorporate some noise around the explosive path while also revealing the fundamental value $\tilde{V}$, one could introduce public signals about V that arrive at a given frequency. This change would incentivize the insider to trade faster and provide the reversal after time $\bar{t}$, in accordance with the newly incorporated information. Including public signals would add an element of realism to the model, as it aligns with the way real-world markets react to information and news releases.

9 Explosions around short interest fluctuations

9.1 Discussion of mechanism

The model provides clear predictions regarding the underlying mechanisms that may contribute to the occurrence of explosive events. First, it highlights the importance of having an *ex-ante* expectation of inelastic demand. Second, it emphasizes the significance of limited asset supply in the market. Third, it underscores the role of actual buying pressure as a catalyst for these explosive occurrences. To empirically test this mechanism, one may seek events that impact these critical factors. Simultaneously identifying events that affect one of these ingredients while maintaining sufficient statistical power during testing is a challenging endeavor. For instance, there is evidence indicating that index-inclusion events appear to predict explosive events.[58] Nevertheless, due to the infrequency of such events and the limited power to detect explosive occurrences, this test does not offer sufficient statistical strength for a conclusive judgment. In this section,

[58]The analysis will be added later to the Appendix of the paper.

I propose an event that is likely to influence all the ingredients in a manner that increases the likelihood of explosive events. These are short interest dissemination. Moreover, this event provides a substantial number of observations, facilitating empirical testing of its relevance.

I concentrate on the events surrounding the public disclosure of short interest data, which currently follows a pre-scheduled bi-monthly pattern when the Financial Industry Regulatory Authority (FINRA) releases data for the most publicly traded assets, including the individual stocks examined in this paper. Historically, this disclosure has been a significant source of public information regarding the extent of short selling activity in these assets. Despite the emergence of other sources of short interest data from private vendors in recent years,[59] FINRA's data still possesses robust predictive power for stock returns.

Within the context of stock explosiveness, the release of this data is expected to impact various critical factors simultaneously. It influences the expectations of inelastic demand, the available supply of the asset in the market, and the likelihood of insider trading activity. Furthermore, this collective impact suggests that the release of data indicating an unusually high level of short interest should, at the very least, raise the probability of an explosion up.

To illustrate this, consider that a substantial outstanding short interest ultimately prompts those holding these positions to close or cover them, potentially increasing the probability of an inelastic surge in buying pressure. This heightened demand, under such circumstances, is often referred to as a "short squeeze event." Additionally, in the case of shorting, the supply of the asset is likely to decrease, as the lenders who have loaned their stocks may be less inclined to lend more. Lastly, it is widely believed that short interest is often held by insiders, making them more likely to have established negative positions prior to the data's release, potentially incorporating the information into the stock price and reducing the immediate entrance of insiders into the market. Altogether, this suggests empirical testing to examine an increase in explosions up in stocks with high short interest following the public release of data.

Building on previous findings that heavily shorted stocks tend to yield negative returns after announcements, a mechanical factor comes into play that can potentially reduce the likelihood of detecting an increase in stock explosiveness up while enhancing the detection of explosiveness down. This effect arises because a larger proportion of observations result in negative returns. Increase in explosions down is also consistent with an increase of crash-probability studied in Callen and Fang (2015).

In this section, we will demonstrate that despite this influence, the incidence of upward explosions still increases, underscoring the robustness of the mechanism linked to changes in beliefs about market primitives. Furthermore, by implementing an event-study-type analysis, examining explosiveness just before and

[59] See Kim et al. (2022) for discussion of private-sector short itnerest data.

just after dissemination, we will find that short interest appears to be the only economically and statistically significant factor in predicting subsequent explosiveness. Moreover, other static return characteristics react more moderately to the information inflow, emphasizing the importance of changes in time series dynamics rather than solely distributional properties of the returns.

9.2 Data

The firm-level short interest data originates from FINRA releases. This dataset begins in 1973 and comprises monthly reports before January 2007. Starting in 2007, short interest was collected and disseminated twice per month. The data was made available to the public eight business days after collection, resulting in an average of 10-11 business days between dissemination dates. I accessed this data through the Compustat North America database, which was made available via WRDS (Wharton Research Data Services).

In line with previous research on short interest (e.g., Asquith et al. (2005); Boehmer et al. (2008), Boehmer et al. (2008), Boehmer et al. (2010)), the primary measure of short selling is the short interest ratio that is the percentage of total shares outstanding:

$$SIR_{s,d} = \frac{\text{Shares Shorted}_{s,d}}{\text{Shares Outstanding}_{s,d}},$$

where d refers to the day following the dissemination of the short interest date [60]. Previous literature on short interest has demonstrated that the SIR is significantly influenced by firm characteristics, industry factors, and prior stock performance. To mitigate the impact of these predictors, I employ an alternative short-selling measure called *partialled out short interest ratio*, $SIR^{po}_{s,d}$. This measure represents the residual obtained from the following cross-sectional regression:

$$\begin{aligned} SIR_{s,d} = &\sum_{X=Size,BtM,IV,Mom,STRev} \sum_{q=1}^{5} \beta_{X,q} I_{X,s,q,d} + \\ &+ \sum_{f\in Mkt,HML,SMB} \sum_{q=1}^{5} \gamma_{X,q} Beta_{s,q,d,f} + \sum_{i\in Industries} \omega_i I_{i,s,d} + u_{s,d}, \qquad (24) \\ SIR^{po}_{s,d} \overset{\text{def}}{\equiv} &\; u_{s,d}. \end{aligned}$$

Here, $I_{X,s,q,d}$ is an indicator of a firm s being in quintile q by firm characteristic X on day d, where X can represent stock market equity, book-to-market ratio, idiosyncratic volatility, momentum, or short-term reversal. $I_{i,s,d}$ represents an industry control, and $Beta_{s,q,d,f}$ is an indicator of the firm s to be in quintile q by factor f exposure on day s. All variables, except for momentum and last-month return, have a 31-day lag.

[60] the short interest is released after the regular trading hours. Hence, the information is incorporated into price on the next business date.

For more details and a summary of the variables, please refer to Table 16. To account for the informational content of the short interest disclosure, I will also utilize the changes in both measures since the last reporting date.

$$\begin{aligned}\Delta SIR_{s,d} &= SIR_{s,d} - SIR_{s,d-1/2\ month}\\ \Delta SIR^{po}_{s,d} &= SIR^{po}_{s,d} - SIR^{po}_{s,d-1/2\ month}\end{aligned} \tag{25}$$

The partialled out difference, $\Delta SIR^{po}_{s,d}$, more accurately captures the surprise in the change of SIR for the stock and fits better for estimation of informational content on the dissemination date. (see e.g. Hanauer et al. (2023))

Table 16 represents the summary on all four measures over the sample of stocks for which explosiveness is estimated over the subsequent 10 days interval. It is worth noting that the sample exhibits a slight bias toward more liquid and larger stocks compared to the sample used in short interest literature, as we require the stocks to be consistently traded in TAQ over the estimation period. In this sample, the median and average Short Interest Ratio (SIR) are 3.17% and 5.16%, corresponding to 4.74 and 6.66 days-to-cover[61]. The 90th percentile of SIR and SIR^{po} are 12.29% and 5.22%. Additionally, the 90th percentile of the change in SIR, SIR^{po}, and days-to-cover are 0.64%, 1.45%[62], and 0.84 days, respectively. Since SIR is heavyFor most of the a

9.3 Empirical Findings

Returns after the dissemination

Let us firstly establish the relationship between the short interest ratio (SIR) and explosiveness. I begin by examining the probability of a market explosion over 10-day periods starting from one of the next days after the release of short interest data. In line with the short-interest literature, I categorize a stock as having high SIR if its SIR falls within the top decile for a given date. This categorization serves as the treatment variable in the regression analysis.

Before delving into the explosiveness analysis, we must assess the difference in average returns between high-SIR stocks and other stocks over the same 10-day interval under investigation. Table 17 presents several results regarding the predictive power of this treatment effect on the abnormal average returns of these stocks. The first column reports the estimates from a simple regression analysis, indicating that high-SIR stocks exhibit, on average, log-returns approximately 66 basis points lower than those of other stocks

[61] Days-to-cover is defined as SIR normalized by the average 31-days lagged average monthly turnover

[62] The fact that the partialled out short interest is more volatile suggests that individual short interest levels are quite persistent, and the aggregate movements used for adjustment introduce additional variation in the measure

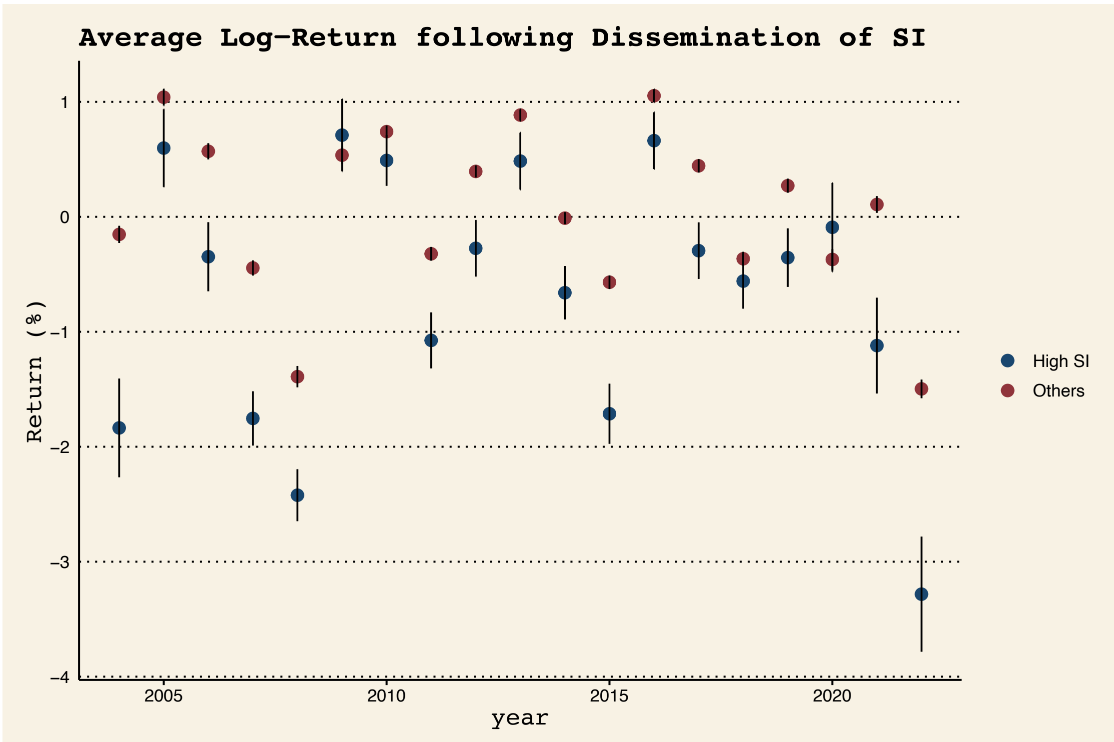

Figure 20: Log-Return for top-decile SIR vs Other Stocks

This plot shows the average 10-day log-return following the dissemination of short-interest data. The blue dots correspond to the equally-weighted average return of stocks with reported SIR in the top decile, while the red dots correspond to the equally-weighted return of other stocks. The 95%-confidence interval around the averages is based on standard errors clustered for firm-year interaction. To be included, the stock must have a closing price of at least $5 thirty days before the disclosure and a good price in the TAQ dataset over the 10-day interval, allowing for the estimation of explosiveness measures.

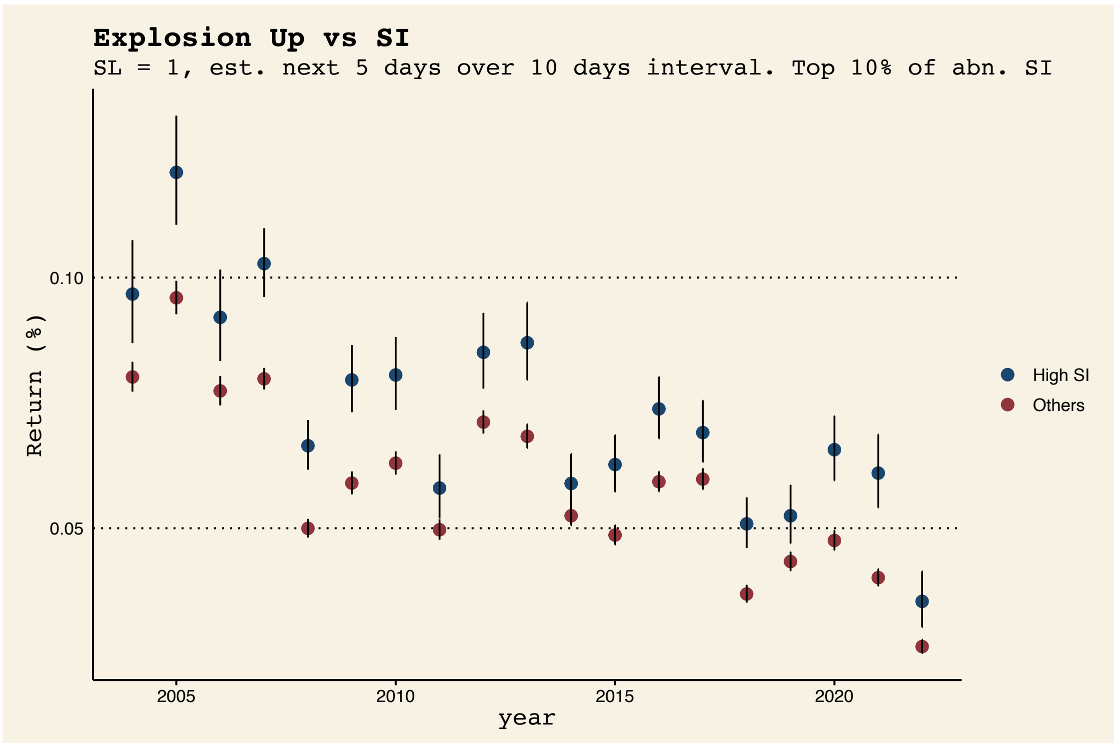

Figure 21: Explosion up for top-decile SIR^{po} vs Other Stocks

This plot shows the estimated probability of explosion up (in percentage terms) following the dissemination of short-interest data. The blue dots correspond to logit-based estimate to be explosive up for stocks with reported SIR^{po} in the top decile, while the red dots correspond to other stocks. A stock is considered explosive up if we detect an explosion up according SADF procedure with $k = 1$ and $r = 0.2$ over at least one of the five subsequent (overlapping) 10-day intervals following the disclosure date. The 95%-confidence interval around the averages is based on standard errors clustered for firm-year interaction. To be included, the stock must have a closing price of at least $5 thirty days before the disclosure, and have a good price in TAQ dataset over the 10-days interval allowing estimation of explosiveness measure.

during the same period. The following two columns confirm the robustness of this result, accounting for risk adjustment according to CAPM and FF3 adjustments. This robustness extends to other factor adjustments, although these specific adjustments align with those used in estimating explosiveness.

Columns 4 to 6 present the results of the same least square regressions, augmented by a set of controls and fixed effects that will be used in the explosiveness analysis. These fixed effects include factors such as the stock and date interacted with industry. The controls use both a set of standard continuous variables that capture firm characteristics, such as log-size, book-to-market ratio, liquidity, volatility, and institutional ownership. Additionally, a set of dummy variables is used to capture non-linearity in the dependence. This includes quintile categorization based on size, volatility, liquidity, and exposures to market return, SMB-portfolio return, and HML-portfolio return. Additionally, controls include previous returns, which capture momentum, as well as the returns from the previous month and the most recent changes in assets and the firm's operational profitability.

Remarkably, even with this extensive set of controls, the average return and abnormal returns remain statistically significant and economically negative, ranging from -9.14% to -7.05% in annualized terms.[63]. The table reports only continuous variables and includes the significant quintile-based dummy variable for clarity. In the sample, following the dissemination of short interest and after controlling for fixed effects, larger, with lower book value, and more liquid stocks had significantly lower returns[64]. This highlights the necessity to control for the factors when performing the explosiveness analysis.

This finding remains consistent across the years of observation. Figure [20] illustrates the difference in average returns between heavily shorted and other stocks for each year, with 95% confidence intervals based on standard errors clustered by firm-year interactions. These results corroborate the predictive power of short interest, as documented previously.

Explosiveness Up for heavily shorted stocks

Now, shifting our focus from analyzing the reaction of returns to short interest, let us delve into the association between explosiveness and short interest. More specifically, we will examine the following specification:

$$E^{\alpha,up}_{s,d} = \beta \times I\{s \text{ is in top 10\% by } SIR_{s,d}\}_d + Controls_{s,d} + \varepsilon_{s,d}, \tag{26}$$

where $E^{\alpha,up}_{s,d}$ represents an indicator that signals the detection of an explosive up episode after date d. To ensure a more precise estimate, we will aggregate the explosiveness data over multiple daily estimates, effec-

[63]use $(1+r)^{252/10}-1$ to annulize the values from 10-days interval

[64]A higher quintile of the Amihud Liquidity measure indicates a higher price impact for the stock

tively utilizing G overlapping 10-day intervals. For this section, we primarily use $G = 5$ as our specification. This means that $E_{s,d}^{\alpha,up} = 1$ if the SADF procedure detects an explosion over one of the intervals, such as $[d, d+9]$, $[d+1, d+10]$, and so on, up to $[d+4, d+13]$. This approach must be more accurate and must mitigate the potential noise that may arise immediately following the data release. Nevertheless, the results remain stable when using a similar estimate based on only one, two, or three 10-day intervals, even with greater magnitude in absolute terms. Details are provided in the appendix.

Figure 21 captures the summary on explosiveness up by year. For every year[65], we estimate logit-model with treatment of being in top SIR^{po} decile[66] and provide 95%-confidence interval based on year-firm clustered standard errors around the estimate. Consistently, in nearly all years, the probability of upward explosiveness is significantly higher for stocks in the top decile of short interest. The difference ranges from 2.50% in 2005 to as low as 0.64% in 2014. The finding is robust to the definitions of explosiveness, selected significance level[67], across various risk-adjustment schemes and selected lag-order for SADF estimates.

Table 19 presents a more comprehensive analysis of the relationship between upward explosions and being a high-SI stock, using the linear specification with date-industry and firm fixed effects, along with a set of controls previously discussed for returns. According to Panel A of the table, in the overall sample, using a linear model, heavily shorted stocks are .53% more likely to be identified as explosive at a 1% significance level after the announcement. This likelihood increases to .60% when we use partialled out SIR to identify the most explosive stocks. The effect generally slightly decreases when we use detected explosions based on risk-adjusted price data. Alternatively, using a logit fixed effect model where the probability of detection is used as the dependent variable, we find that the probability of detection changes by 9-11% for heavily shorted stocks compared to other stocks. The coefficients are all statistically significant at the 1% level. Importantly, the high short-interest indicator emerges as one of the strongest predictors, both in terms of statistical significance and economic impact, for explosiveness up following the dissemination date.

The other most significant controls, as reported in Table 19, include size, liquidity, institutional ownership, and idiosyncratic variance, along with previous returns, namely momentum and short reversal. Notably, these controls highlight important factors affecting explosiveness detection. For instance, the relationship between previous returns and explosiveness suggests that stocks are less likely to experience explosive movements if they have already grown significantly over the previous year, particularly if this change occurred earlier. This observation aligns with the idea that stocks tend to undergo reversals, resulting in a lower probability of detecting upward explosions.

[65] I miss 2003, since the year is incpomplete.

[66] at this point we do not use controls, they will be used in later specifications

[67] up to significance level 5%

Size appears to be negatively correlated with explosiveness detection, with larger stocks showing an even stronger decrease in the largest quintiles. Controlling for firm effects, this could be partly attributed to the same reversal narrative, but it likely reflects increased market attention, making it less likely for these stocks to separately exhibit explosive dynamics. Economically, the size effect is the largest causing almost 2% drop in explosiveness up following detection if taken one standard deviation of it keeping everything else equal.

Moreover, the most liquid stocks with greater institutional ownership tend to be less prone to explosiveness. This could be attributed to the ease of shorting in the presence of institutional investors, which can act as stabilizing forces for the explosions. On average, stocks with a history of high idiosyncratic volatility are more likely to exhibit explosiveness, possibly reflecting uncertainty about their underlying fundamentals. Despite the continuous variable's statistical and economic significance—causing an approximate 0.30% increase in explosiveness for every one-standard-deviation rise—this relationship is intricate and nonlinear. The non-linearity is in part offset by the negative dummies associated with being in the IV quintile portfolios.

To delve into this in more detail, let us consider Table 20, which reports coefficients on idiosyncratic volatility, along with interactions with the high short interest indicator, while keeping all other controls and fixed effects in place. The top quintile of stocks by idiosyncratic volatility is expected to be around 0.30% higher due to the continuous variable. However, the expected impact is approximately the same when considering the controlled $Q5 - IV$ coefficient, which indicates being in the top 20% of volatile stocks. Furthermore, the interaction with high short interest rates is also non-linear, with the largest effect varying from .80% to 1.3% across specifications, occurring for the most and least volatile stocks. This non-linear dynamic is related to two mechanisms that influence the power of explosiveness identification. On one hand, higher volatility makes it more challenging to detect explosive movements, as it requires greater statistical power to reject the null hypothesis. On the other hand, more volatile stocks are characterized by greater uncertainty about their fundamentals, increasing the potential for explosiveness to emerge.

Explosiveness Up following dissemination of changed SI

The previous specification demonstrates that short interest data predicts explosiveness up, in line with the model's mechanism prediction, despite the fact that, on average, stocks tend to go downward. The next specification aims to capture the informational content delivered on the dissemination date itself. In this specification, we use the change in short interest ratio as the explanatory variable and compare the detection of explosiveness measured just before and just after the dissemination date:

$$E^{\alpha,up}_{s,d+1} - E^{\alpha,up}_{s,d-L} = \beta \times I\{s \text{ is in top } 10\% \text{ by } \Delta SIR_{s,d}\}_d + Controls_{s,d} + \varepsilon_{s,d}. \tag{27}$$

Here, L represents the lag before the dissemination date. The idea behind this specification is that if the stock properties do not change significantly over L days, we should observe changes in the model parameters following the public dissemination of information. The choice of L involves a trade-off. Selecting a large L may introduce a reverse causality problem because previous explosive episodes can lead to increased shorting of the stock. Selecting a smaller L results in increased overlapping between the intervals for detecting explosions with $E^{\alpha,up}_{s,d+1}$ and $E^{\alpha,up}_{s,d-L}$, potentially reducing the captured effect. However, small overlapping is less likely to be problematic, as it is empirically more challenging to detect explosions toward the end of the estimation interval than at the beginning.

Finally, given the nature of short interest data, which is the short interest settled on the collection date, one might expect some change in explosiveness immediately after the collection if market participants have alternative methods to estimate short interest data rather than waiting for the dissemination date. This could potentially weaken the derived results.

Taking into account these concerns, we present multiple specifications with varying values of L (e.g., 7, 10, 12 days), with the primary specification using $L = 10$. Similar to our previous analyses, we examine various aggregations of explosiveness detection based on subsequent dates, denoted as G. G varies from one (using a single interval) to five, with the primary specification using $G = 3$ intervals. The choice of $L = 7$ is motivated by estimating explosiveness immediately after the collection date, 7 days before the dissemination of information to the public, to address the reverse causality concern. On the other hand, $L = 12$ allows for zero overlapping between underlying estimation periods for $E^{\alpha,up}_{s,d+1}$ and $E^{\alpha,up}_{s,d-L}$ with $G = 3$.

Table 21 reports the first set of the results on the explosiveness on 27 for the four type of aggregations of explosiveness (G) and the three primary lag levels L. As the explosiveness measure, the detection at 1% significance level with $k = 1$ and CAPM-adjustment for price is used. Consistently for all specification explosiveness up is greater after the dissemination of data on the top change in the short interest. The increase in probability is measured by .30% when using just one 10-day interval and slightly increases when using more intervals to find an explosion.

Importantly, in the difference based specification High SI appears to be the only consistently significant and economically meaningful coefficient for change in explosiveness up. Table 22 provides results on both explosiveness down and explosiveness up change around the dissemination. As before, I provide results for three specifications based on different adjustment. The table contains information on regression coefficients of all continuous control variables and the few quintile based control variables that appeared significant at least in one of the specification. Focusing on the last three columns, we can see explosiveness up does not have other factors that would appear significant in more than one coefficient, size, shortterm reversal, momentum , idiosyncratic variance do not predict any difference and have economically neglectable coefficients. That allows to interpret the coefficient as the reaction of explosiveness up specifically to the informational content in the short interest dissemination. Explosiveness grows up consistent with the narative developed in the model: a public signal of increasing probability of inelastic demand and lower probability of insiders buying the asset.

Finally, in this section, we do not closely study explosiveness down. The reason for this is that it is challenging to separate the mechanism related to liquidity from the informational content in the data studied before. Previous literature has shown that negative information from short interest tends to not be incorporated immediately but rather up to 30 days later (e.g., Hanauer, 2023), indicating a gradual process. My results are consistent with this finding, suggesting that the transition might also happen explosively.

The first three columns of Table 22 also show that the informational content at the dissemination of short intest increases the probability of explosiveness down by .52% to .73%, with larger magnitudes than explosiveness up. The results are robust to alternative specifications (see the Appendix for other specifications involving explosiveness down), suggesting that at least partially, the price discovery following the data disclosure happens explosively. Nevertheless, other factors such as short-term reversal and idiosyncratic volatility contribute to the regression, indicating that the mechanism driving explosiveness down around the dissemination is more complex than the one attributed to explosiveness up.

Static Return Moments following dissemination of changed SI

The open question pertains to the extent to which the detected explosiveness is influenced by factors other than the properties of returns and informational content. An alternative hypothesis could be that increasing short interest precedes significant firm news or stock turmoil. To address these concerns, we have already compared explosiveness just after the data collection date to the period following dissemination. If short interest precedes news, it is more likely that the news would be closer to the short interest settlement date rather than 8 business days later. To investigate these questions further, let us examine how the statistical moments of high-frequency returns change in the sample following the dissemination.

Table 23 provides the averages of the realized empirical moments within the 10-day interval used for SADF estimation. This table includes heavily shorted stocks indicated by the top-decile change in partialled out Short Interest Ratio (SIR^{po}) and all other stocks. Despite the substantial change of -0.28% in the average return for heavily shorted stocks following the dissemination, the reaction of all other moments is relatively modest. For instance, volatility and the absolute values of quintiles for intraday returns show an average change of less than half of basis point for both heavily shorted and other stocks.

The extreme returns, which often represent overnight returns in the sample, change slightly, primarily on the minimum side. The average minimum captured return decreases by 8 basis points for heavily shorted stocks versus 1 basis point for other stocks. The maximum return goes up by 4 basis points for heavily shorted stocks compared to 2 basis points for other stocks. This results in slightly smaller skewness and larger kurtosis in the sample. However, all these changes are economically small.

When we control for firm characteristics and include the same fixed effects used in the analysis of returns and explosiveness, the results remain consistent. Table 24 reports the coefficients for the indicator of being in the top decile by the change in partialled out Short Interest Ratio in the fixed-effect regression. The economic magnitudes are similar to those in the unconditional analysis. There is a statistically significant increase in intra-period volatility, with stocks being 0.20 basis points more volatile following the disclosure of high short interest. This increase may be mainly attributed to the decline in minimum returns by 8 basis points and the rise in maximum returns by 3 basis points. While these results are consistent with the appearance of explosiveness, which itself can increase volatility, the changes are relatively small.

The 10% and 90% quintiles of returns actually mildly decrease in absolute terms, skewness decreases by 5.7%, while kurtosis increases by 2.82. The fact that most of the estimated coefficients are statistically significant suggests that the moments of high-frequency returns are quite stable after controlling for aggregate fluctuations and other firm-specific factors. However, it is unlikely that there is a fundamental change in the moments that would substantially impact the detection of explosiveness.

On the other hand, the average increase in the probability of explosiveness up and down is more substantial. The increase in the probability of explosiveness is noteworthy when considering unconditional changes. Over the 10 days before and after the dissemination of short interest, the probability of explosiveness increases by approximately 0.31% for stocks in the top decile of SIR change and 0.42% for other stocks. The average unconditional probability of explosiveness before dissemination is 2.58% and 3.10%, which translates to around a 12% increase in explosiveness. The fixed-effect regression results were discussed in the previous section and have approximately same magnitude (see Table 21).

This underscores the importance of the informational content of short interest dissemination for explosiveness. It is unlikely that the information about high short interest precedes fundamental changes in the static return properties, except for predicting subsequent negative returns. This is consistent with the idea that short interest data does not just influence beliefs about where the stock will go, but also who will likely trade in the future, creating the conditions for explosive transitional dynamics.

10 Monte-Carlo Simulations

10.1 Empirical challenges

The accuracy of the SADF test relies on the speed of convergence to the asymptotic distribution mentioned in equation (6). The underlying assumptions necessary for achieving this asymptotic distribution are relatively mild, such as strict stationarity and ergodicity in the error terms of equation (5). However, as documented in the econometric literature, finite sample unit-root tests are susceptible to both type 1 and type 2 errors (e.g., Elliott et al. (1992), Ng and Perron (2001)). This susceptibility may be exacerbated by the fact that innovations in individual stock returns exhibit high skewness and fat tails.

Furthermore, the presence of stochastic volatility, a common assumption in stock return models, can pose challenges for the size and power of unit root tests (see e.g., Zhang et al. (2013)). Specifically, it introduces a positive bias toward rejecting the unit root hypothesis in finite sample sizes when using the ADF procedure. This bias is particularly pronounced when stochastic volatility approaches non-stationary or non-stationary behavior.

If stock prices followed a continuous path, as often modeled in jump-diffusive processes, increasing the frequency of time stamps could address many obstacles by expanding the sample size. However, in practice, a trade-off arises when selecting time intervals for analysis. More frequent time-stamps introduce additional microstructure noise into price fluctuations, while excessively long intervals may exacerbate finite-sample issues. In this paper, I strike a balance by choosing a 5-10 minute intervals. Longer intervals may be suitable for detecting long-term bubbles, but this would require a separate investigation beyond the paper's scope.

Another challenge in the analysis involves dealing with overnight returns, which have markedly different statistical properties than 5-minute intraday returns, featuring roughly five times larger standard deviations. These returns could be seen as scheduled jumps, leading to a naturally higher probability of explosive detection. Therefore, it is essential to investigate the influence of jump events on explosive detection. Finally, the presence of standard drift and autocorrelation can impact the frequency of explosive event detection.

10.2 False Positive

To evaluate test size and power, I perform simulations using finite samples of a size comparable to the periods discussed in Section 3. These simulations will utilize data generating processes as described in equations (1) and (5). To explore the impact of the mentioned issues, I conduct multiple simulations: one with Gaussian error terms that mirror the autoregressive structure found in actual data and match the moments of realized returns, a second with bootstrapped error terms to align with the actual data distribution, and a third in which data is generated incorporating stochastic volatility.

To be more specific, I undertake three types of simulations employing normal, independent, and identically distributed error terms as outlined in equation (5). To create realistic processes, I begin by estimating autoregression parameters in accordance with equation (5), extracting both the autoregression coefficients and residuals. The moments of simulated Gaussian terms are taken from the empirical moments of the residuals. Subsequently, I conducted Monte Carlo simulations. In these simulations, I randomly select[68] the coefficients based on their empirical distribution within a given year or set them as constants. To assess the impact of negative autocorrelation, I set these constants to either zero, which transforms the process into a standard random walk, or the median values of the estimated autocorrelations, which are negative.[69]

The results of 19,000,000 simulations for each specification (100,000 simulations per year) can be found in the first three columns of Table 25. To check the result if the model is misspecified I also estimate it with one additional or one less lag. The results are reported for 1% significance level. The default lag-order is 2.

In the *Random Walk (RW)* column, it becomes evident that in finite samples, the test size is only marginally smaller than the theoretically expected detection rate of 0.5%. This implies that to be more precise, the quantiles used for the test could be adjusted slightly lower, i.e., relaxed. This observation holds true even when we introduce constant (negative) autocorrelations into returns. However, as soon as we introduce randomness into the autocorrelation terms together with a random drift, which mirrors what is observed in the data, we observe a rise in false positive detections. This phenomenon may be attributed to the presence of a strong drift or significant autocorrelation terms in the data, which, in finite samples, occasionally leads the test to interpret it as explosiveness mistakenly. Not surprisingly, the more substantial effect on false detections occurs for downward explosiveness since extreme negative autocorrelations are more common. Finally, in the *Overnight* specification, where I introduce a scheduled overnight jump and draw the random autocorrelation terms, the false detection rate is close to targeted level of .5%. For this specification, the simulated intraday and overnight returns are drawn from the Gaussian distributions fitted into the actual intraday and overnight returns.[70]

The effects become more pronounced when we deviate from normality and generate return innovations with fat tails based on their empirical distribution. In the first column, assuming a random walk with

[68] I randomly draw the set of parameters $\{\alpha, \phi_j, j = 1, \dots, k\}$ with replacement.

[69] For example, in Table 25, where $k = 2$, the values are -0.07 and -0.03.

[70] It is enough to estimate the average and standard deviation for the samples and draw Gaussian returns with respective parameters.

bootstrapped errors leads to an increase in false negative values to slightly over 1%. Not setting the autocorrelation terms to median values (which produces nearly identical results) let us concentrate on two autoregression specifications. In one scenario, we simulate autocorrelations terms while setting the drift to zero, while in the other, we simulate both the drift and the coefficients. This enables us to determine which factor, drift or autocorrelation, has a more significant impact on the likelihood of detecting explosions. Surprisingly, autocorrelation alone does not appear to introduce detection issues, whereas the introduction of drift may result in an additional 35 basis points in false detections. Finally, to make our simulations closely resemble actual data, we introduce scheduled overnight jumps into the returns based on the distribution of actual overnight returns. This results in a slight increase in false positive detections by a few basis points.

The final specification addresses the concern that return volatility may exhibit clustering, which could potentially elevate the probability of localized explosions. Intuitively, if a sequence of low-volatility returns is followed by a clustered episode of high volatility, there is a greater likelihood of multiple successive realizations in one direction. This can be easier to interpret as an explosive term when fitted in the rolling procedure. To explore this, I estimate the Heston model using the GMM method in accordance with Ellickson et al. (2018), and then conduct simulations of paths based on the model while preserving the AR coefficients estimated in the market. As observed, this indeed increases the probability of Type I error to approximately 1.2%.

In general, all the discussed specifications exhibit robustness when extra lags are included in model estimation. However, using fewer lags tends to result in a lower likelihood of detecting explosions. The primary reason for this is that the absence of negative correlations in the returns redistributes them to the explosive term.

Finally, to emphasize the role of time-series dynamics in the detection of explosiveness, as a separate specification, I use a procedure that preserves the static return distributions from the actual data sample over the estimation period and autocorrelation of high-frequency returns. Instead of drawing the random sample from the aggregate distribution, it resamples residuals from (5) from within the estimation window[71] to simulate the bootstrapped time series. Figure 22 reports the resulting empirical distribution of p-values of the respective SADF estimates. Like in the case of previous Monte Carlo studies, one can see that the static return creates false detections, with the most substantial false detection at a 1% significance level of around

[71] I draw without replacement 770 residuals for a given stock and a given 10-day window. In the case of zero autocorrelation, this would guarantee preserving the cumulative return as well as all static return properties within the window. In the case of stationarity and no volatility clustering, that must also guarantee a similar cumulative return. The aggregate sample consists of 320 random stock-10-day window pairs per day of observation, totaling 1,559,040 samples.

2.2%. Still, this number does not closely match the detection rate of 6.4% for the actual data with the same parameters.[72]

To sum up, we have observed that the distributional characteristics of returns, as well as the presence of significant drift, can lead to false detections. Conversely, there is also the possibility of reverse causality where the estimated substantial drift in the data may actually be caused by explosions, but is incorrectly attributed to drift in the absence of the explosive term. However, it is important to note that none of the specifications were able to closely match the frequency of explosion detection in the actual data. This highlights the insufficiency of the conventional AR-based model in providing a comprehensive description of the data, implying the presence of additional time-series dynamics at work.

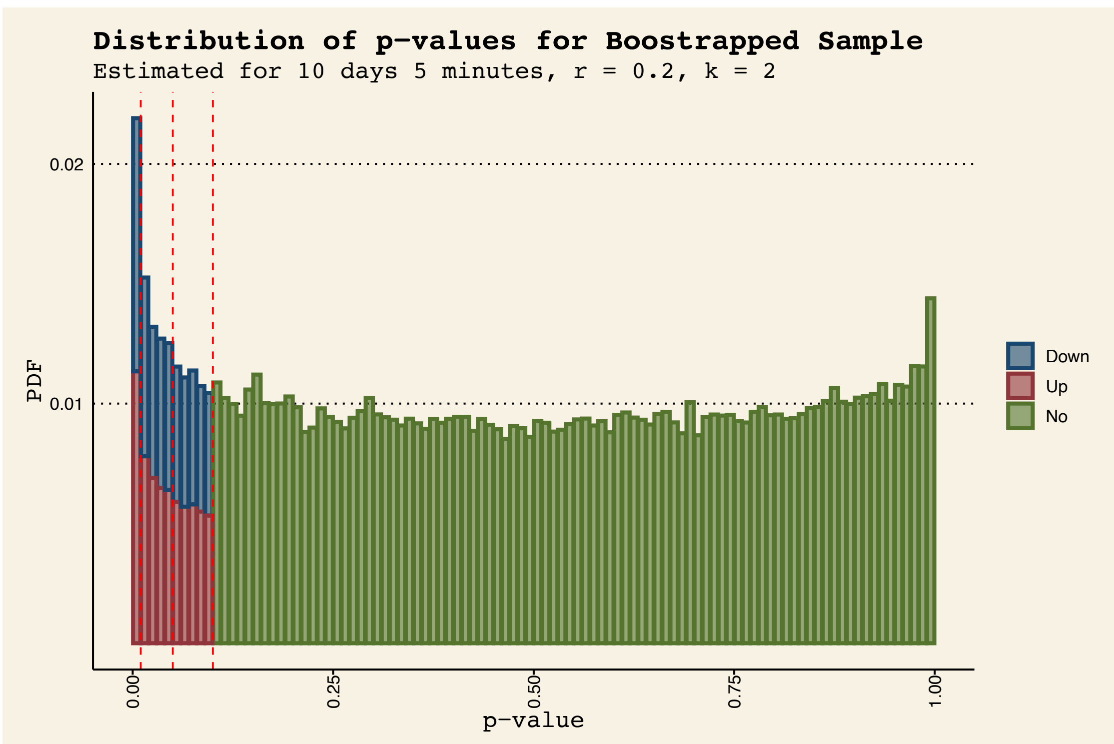

Figure 22: EPDF of p-values for bootstrapped time-series

11 Conclusion

In this paper, the novel phenomenon of explosiveness in individual stocks is examined, revealing that it is a widespread and impactful occurrence that cannot be attributed to random variations under traditional asset

[72] The used parameters are $k = 2$, $r = 0.2$, and $SL = 1\%$. Reducing/increasing the lag-order k leads to a marginal decrease/increase in the detection rate for both measures.

pricing models. These explosive episodes are not simply outliers or fleeting "mini-bubbles"; instead, they play a crucial role in the price discovery process. Explosive episodes are shown to exert a lasting influence on asset prices, albeit with partial reversals. They can be categorized into two distinct types: explosions up and explosions down, each with its unique characteristics. Explosiveness up tends to be more idiosyncratic, while explosions down exhibit clustering and cyclical behavior across markets.

The paper begins by introducing the theoretical concept of explosiveness as a local unit root violation. It underscores that detecting the theoretical explosion does not necessarily aim at identifying a martingale violation but rather focuses on the *ex-post* detection of a stochastic form of price convexity. The paper further explores the application of the SADF procedure for detecting explosive episodes in high-frequency data and presents estimates based on 20 years of data encompassing the cross-section of the U.S. common stock universe. The estimates reveal a detection frequency more than six times higher than anticipated, occurring on at least 2.4% of dates for individual stocks. The prevalence of this phenomenon prompts questions about the appropriate modeling approach and the underlying mechanisms driving it.

Aggregate portfolios, encompassing both market and standard double-sorted portfolios, exhibit a higher susceptibility to explosive downward movements but seldom display upward explosiveness. The explosiveness within these aggregate portfolios is concentrated and correlated with the explosiveness of individual stocks, yet it falls short of describing the majority of cases. This observation implies the utility of risk adjustment before estimating the occurrence of explosions.

The average magnitude of individual stock explosions is 10%. I demonstrate that both directions of explosiveness feature a partial reversal of 10% to 15% of the change before detection. This provides predictable and tradable returns, separating explosions from large overnight and high-frequency jumps. The portfolios formed based on explosive stocks generate persistent alpha over the 20 years of observations that standard factors cannot explain.

The study establishes links between these explosive episodes and liquidity measures, such as trading volume and buying pressure, highlighting their intricate connection to the underlying market structure. Unlike jumps, explosive episodes feature a pronounced buying pressure of around a 9% increase in abnormal order imbalance. In contrast, the change in abnormal trading volume by 46% and 36% for up and down explosion dates features smaller magnitudes than the jumps, pointing out the different nature of explosions.

A series of models is developed to elucidate the mechanisms that lead to explosiveness, and these models illustrate that explosiveness naturally emerges as an equilibrium outcome. These models delve into the interplay between informed traders, inelastic demand, and market dynamics, revealing that explosiveness primarily hinges on factors such as the average knowledge possessed by insiders, the probability of their presence in the market, and the distributional characteristics of inelastic demand.

Notably, explosiveness can manifest whenever the distribution of the demand size exhibits thin tails, diminishing the likelihood of large demands but not precluding their occurrence entirely. The presence of insiders is crucial for informative explosiveness, leading to a situation where the price experiences only partial reversals.

To empirically test the models' mechanisms based on expected inelastic demand and a lower presence of informed traders, the study examines explosiveness around short-interest dissemination dates. Despite the well-documented negative impact of high short interest on future returns, the analysis reveals that stocks with the same feature tend to exhibit a higher probability of explosiveness up. This association is scrutinized further through an event-type analysis, which assesses explosiveness before and after the dissemination date. The findings suggest that stocks experiencing a negative surprise due to heavy shorting, as measured by changes in the short interest ratio, tend to be more explosive immediately after the information becomes available. This variable proves to be the only consistent predictor among a broad range of firm characteristics, underscoring the pivotal role of short interest data in driving changes in explosiveness around the dissemination date. In contrast, other static risk measures do not display the same economically meaningful shifts, emphasizing the role of explosions as a subsequent price appreciation mechanism.

References

Abreu, D. and M. K. Brunnermeier (2003). Bubbles and crashes. Econometrica 71(1), 173–204.

Aït-Sahalia, Y., I. Kalnina, and D. Xiu (2020). High-frequency factor models and regressions. Journal of Econometrics 216(1), 86–105.

Allen, F. and D. Gale (2000). Financial contagion. Journal of political economy 108(1), 1–33.

An, B.-J., A. Ang, T. G. Bali, and N. Cakici (2014). The joint cross section of stocks and options. The Journal of Finance 69(5), 2279–2337.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. The review of economics and statistics 89(4), 701–720.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of exchange rate realized volatility. Journal of the American Statistical Association 96, 42–55.

Andersen, T. G., Y. Li, V. Todorov, and B. Zhou (2021). Volatility measurement with pockets of extreme return persistence. Journal of Econometrics.

Andersen, T. G., R. Riva, M. Thyrsgaard, and V. Todorov (2023). Intraday cross-sectional distributions of systematic risk. Journal of Econometrics 235(2), 1394–1418.

Asquith, P., P. A. Pathak, and J. R. Ritter (2005). Short interest, institutional ownership, and stock returns. Journal of Financial Economics 78(2), 243–276.

Back, K., C. H. Cao, and G. A. Willard (2000). Imperfect competition among informed traders. The journal of finance 55(5), 2117–2155.

Barberis, N., A. Shleifer, and R. Vishny (1998). A model of investor sentiment. Journal of financial economics 49(3), 307–343.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2009). Realized kernels in practice: Trades and quotes.

Barndorff-Nielsen, O. E., S. Kinnebrock, and N. Shephard (2008). Measuring downside risk-realised semivariance. CREATES Research Paper (2008-42).

Bernanke, B. S. and K. N. Kuttner (2005). What explains the stock market's reaction to federal reserve policy? The Journal of finance 60(3), 1221–1257.

Bernard, V. L. and J. K. Thomas (1989). Post-earnings-announcement drift: delayed price response or risk premium? Journal of Accounting research 27, 1–36.

Bertsimas, D. and A. Lo (1998). Optimal control of execution costs. Journal of Financial Markets 1, 1–50.

Blanchard, O. J. (1979). Speculative bubbles, crashes and rational expectations. Economics letters 3(4), 387–389.

Boehmer, E., Z. R. Huszar, and B. D. Jordan (2010). The good news in short interest. Journal of Financial Economics 96(1), 80–97.

Boehmer, E., C. M. Jones, and X. Zhang (2008). Which shorts are informed? The Journal of Finance 63(2), 491–527.

Boehmer, E. and J. Wu (2013). Short selling and the price discovery process. The Review of Financial Studies 26(2), 287–322.

Bollerslev, T. (2022). Realized semi (co) variation: Signs that all volatilities are not created equal. Journal of Financial Econometrics 20(2), 219–252.

Bollerslev, T., S. Z. Li, and V. Todorov (2016). Roughing up beta: Continuous versus discontinuous betas and the cross section of expected stock returns. Journal of Financial Economics 120(3), 464–490.

Bollerslev, T. and V. Todorov (2011a). Estimation of jump tails. Econometrica 79(6), 1727–1783.

Bollerslev, T. and V. Todorov (2011b). Tails, fears, and risk premia. The Journal of Finance 66(6), 2165–2211.

Bollerslev, T., V. Todorov, and S. Z. Li (2013). Jump tails, extreme dependencies, and the distribution of stock returns. Journal of Econometrics 172(2), 307–324.

Bouchaud, J.-P., J. D. Farmer, and F. Lillo (2009). How markets slowly digest changes in supply and demand. In Handbook of financial markets: dynamics and evolution, pp. 57–160. Elsevier.

Brunnermeier, M. K. and M. Oehmke (2014). Predatory short selling. Review of Finance 18(6), 2153–2195.

Brunnermeier, M. K. and L. H. Pedersen (2009). Market liquidity and funding liquidity. The review of financial studies 22(6), 2201–2238.

Callen, J. L. and X. Fang (2015). Short interest and stock price crash risk. Journal of Banking & Finance 60, 181–194.

Campbell, J. Y. and P. Perron (1991). Pitfalls and opportunities: what macroeconomists should know about unit roots. NBER macroeconomics annual 6, 141–201.

Campbell, J. Y. and R. J. Shiller (1987). Cointegration and tests of present value models. Journal of political economy 95(5), 1062–1088.

Caporin, M., A. Kolokolov, and R. Renò (2017). Systemic co-jumps. Journal of Financial Economics 126(3), 563–591.

Chakrabarty, B., R. Pascual, and A. Shkilko (2015). Evaluating trade classification algorithms: Bulk volume classification versus the tick rule and the lee-ready algorithm. Journal of Financial Markets 25, 52–79.

Chen, T. and C. Guestrin (2016). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pp. 785–794.

Chen, Y., M. Kim, J. M. McInnis, and W. Zhao (2023). Interest in the short interest: The rise of private sector data. Available at SSRN 4468667.

Cheung, A., E. Roca, and J.-J. Su (2015). Crypto-currency bubbles: an application of the phillips–shi–yu (2013) methodology on mt. gox bitcoin prices. Applied Economics 47(23), 2348–2358.

Christensen, K., R. Oomen, and R. Renò (2022). The drift burst hypothesis. Journal of Econometrics 227(2), 461–497.

Christensen, K., R. C. Oomen, and M. Podolskij (2014). Fact or friction: Jumps at ultra high frequency. Journal of Financial Economics 114(3), 576–599.

Cont, R., A. Kukanov, and S. Stoikov (2014). The price impact of order book events. Journal of financial econometrics 12(1), 47–88.

Corbet, S., B. Lucey, and L. Yarovaya (2018). Datestamping the bitcoin and ethereum bubbles. Finance Research Letters 26, 81–88.

Coskun, Y., U. Seven, H. M. Ertugrul, and A. Alp (2020). Housing price dynamics and bubble risk: the case of turkey. Housing Studies 35(1), 50–86.

Da, R. and D. Xiu (2021). When moving-average models meet high-frequency data: Uniform inference on volatility. Econometrica 89(6), 2787–2825.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam (1998). Investor psychology and security market under- and overreactions. the Journal of Finance 53(6), 1839–1885.

De Bondt, W. F. and R. Thaler (1985). Does the stock market overreact? The Journal of finance 40(3), 793–805.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990b). Noise trader risk in financial markets. Journal of political Economy 98(4), 703–738.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990a). Positive feedback investment strategies and destabilizing rational speculation. the Journal of Finance 45(2), 379–395.

Diba, B. T. and H. I. Grossman (1988). Explosive rational bubbles in stock prices? The American Economic Review 78(3), 520–530.

Diether, K. B., K.-H. Lee, and I. M. Werner (2009). Short-sale strategies and return predictability. The Review of Financial Studies 22(2), 575–607.

Drechsler, I. and Q. F. Drechsler (2014). The shorting premium and asset pricing anomalies. Technical report, National Bureau of Economic Research.

Easley, D., M. L. De Prado, and M. O'Hara (2011). The microstructure of the flash crash. Journal of Portfolio Management 37(2), 118–128.

Ellickson, B., M. Sun, D. Whang, and S. Yan (2018). Estimating a local heston model. Available at SSRN 3108822.

Elliott, G., T. J. Rothenberg, and J. H. Stock (1992). Efficient tests for an autoregressive unit root.

Engelberg, J. E., A. V. Reed, and M. C. Ringgenberg (2018). Short-selling risk. The Journal of Finance 73(2), 755–786.

Evans, G. W. (1991). Pitfalls in testing for explosive bubbles in asset prices. The American Economic Review 81(4), 922–930.

Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. the Journal of Finance 47(2), 427–465.

Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. Journal of financial economics 116(1), 1–22.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987). Expected stock returns and volatility. Journal of financial Economics 19(1), 3–29.

Freyberger, J., A. Neuhierl, and M. Weber (2020, May). Dissecting Characteristics Nonparametrically. The Review of Financial Studies 33(5), 2326–2377.

Gabaix, X. and R. S. Koijen (2021). In search of the origins of financial fluctuations: The inelastic markets hypothesis. Technical report, National Bureau of Economic Research.

Grossman, S. J. and M. H. Miller (1988). Liquidity and market structure. the Journal of Finance 43(3), 617–633.

Hanauer, M. X., P. Lesnevski, and E. Smajlbegovic (2023). Surprise in short interest. Journal of Financial Markets, 100841.

Harrison, J. M. and D. M. Kreps (1978). Speculative investor behavior in a stock market with heterogeneous expectations. The Quarterly Journal of Economics 92(2), 323–336.

Harvey, C. R. and A. Siddique (2000). Conditional skewness in asset pricing tests. The Journal of finance 55(3), 1263–1295.

Harvey, D. I., S. J. Leybourne, and R. Sollis (2015). Recursive right-tailed unit root tests for an explosive asset price bubble. Journal of Financial Econometrics 13(1), 166–187.

Hastie, T., R. Tibshirani, J. H. Friedman, and J. H. Friedman (2009). The elements of statistical learning: data mining, inference, and prediction, Volume 2. Springer.

Hau, L., H. Zhu, R. Huang, and X. Ma (2020). Heterogeneous dependence between crude oil price volatility and china's agriculture commodity futures: Evidence from quantile-on-quantile regression. Energy 213, 118781.

Hirschey, N. (2021). Do high-frequency traders anticipate buying and selling pressure? Management Science 67(6), 3321–3345.

Holden, C. W. and A. Subrahmanyam (1992). Long-lived private information and imperfect competition. The Journal of Finance 47(1), 247–270.

Homm, U. and J. Breitung (2012). Testing for speculative bubbles in stock markets: a comparison of alternative methods. Journal of Financial Econometrics 10(1), 198–231.

Hong, H., W. Li, S. X. Ni, J. A. Scheinkman, and P. Yan (2015). Days to cover and stock returns. Technical report, National Bureau of Economic Research.

Hong, H. and J. C. Stein (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. The Journal of finance 54(6), 2143–2184.

Hong, H. and J. C. Stein (2007). Disagreement and the stock market. Journal of Economic perspectives 21(2), 109–128.

Hopman, C. (2007). Do supply and demand drive stock prices? Quantitative Finance 7(1), 37–53.

Hu, J. (2014). Does option trading convey stock price information? Journal of Financial Economics 111(3), 625–645.

Huberman, G. and W. Stanzl (2005). Optimal liquidity trading. Review of finance 9(2), 165–200.

Jacod, J., Y. Li, and X. Zheng (2017). Statistical properties of microstructure noise. Econometrica 85(4), 1133–1174.

Jegadeesh, N. and S. Titman (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. The Journal of finance 48(1), 65–91.

Jegadeesh, N. and S. Titman (1995). Overreaction, delayed reaction, and contrarian profits. The Review of Financial Studies 8(4), 973–993.

Jiang, G. J. and K. X. Zhu (2017). Information shocks and short-term market underreaction. Journal of Financial Economics 124(1), 43–64.

Kapadia, N. and M. Zekhnini (2019). Do idiosyncratic jumps matter? Journal of Financial Economics 131(3), 666–692.

Kelly, B. T., S. Pruitt, and Y. Su (2017). Instrumented Principal Component Analysis. SSRN Electronic Journal.

Khanna, N. and R. D. Mathews (2012). Doing battle with short sellers: The conflicted role of blockholders in bear raids. Journal of Financial Economics 106(2), 229–246.

Kilic, M. and I. Shaliastovich (2019). Good and bad variance premia and expected returns. Management Science 65(6), 2522–2544.

Kim, M., J. McInnis, and W. Zhao (2022). Interest in the short interest: The rise of private sector data.

Kirilenko, A., A. S. Kyle, M. Samadi, and T. Tuzun (2017). The flash crash: High-frequency trading in an electronic market. The Journal of Finance 72(3), 967–998.

Koijen, R. S. and M. Yogo (2019). A demand system approach to asset pricing. Journal of Political Economy 127(4), 1475–1515.

Koijen, R. S. and M. Yogo (2020). Exchange rates and asset prices in a global demand system. Technical report, National Bureau of Economic Research.

Korajczyk, R. A. and D. Murphy (2019). High-frequency market making to large institutional trades. The Review of Financial Studies 32(3), 1034–1067.

Kozak, S., S. Nagel, and S. Santosh (2020, February). Shrinking the cross-section. Journal of Financial Economics 135(2), 271–292.

Kyle, A. P. S., A. A. Obizhaeva, and Y. Wang (2017). Beliefs aggregation and return predictability. SSRN Electronic Journal.

Kyle, A. S. (1985). Continuous auctions and insider trading. Econometrica: Journal of the Econometric Society, 1315–1335.

Kyle, A. S. (1989). Informed speculation with imperfect competition. The Review of Economic Studies 56(3), 317–355.

Lakonishok, J., I. Lee, N. D. Pearson, and A. M. Poteshman (2007). Option market activity. The Review of Financial Studies 20(3), 813–857.

Lakonishok, J., A. Shleifer, and R. W. Vishny (1994). Contrarian investment, extrapolation, and risk. The journal of finance 49(5), 1541–1578.

Lamont, O. A. and J. C. Stein (2004). Aggregate short interest and market valuations. American Economic Review 94(2), 29–32.

Lee, S. S. and P. A. Mykland (2012). Jumps in equilibrium prices and market microstructure noise. Journal of Econometrics 168(2), 396–406.

Lehmann, B. N. (1990). Fads, martingales, and market efficiency. The Quarterly Journal of Economics 105(1), 1–28.

Li, J., V. Todorov, and G. Tauchen (2017). Adaptive estimation of continuous-time regression models using high-frequency data. Journal of Econometrics 200(1), 36–47.

Li, Z. M., R. J. Laeven, and M. H. Vellekoop (2020). Dependent microstructure noise and integrated volatility estimation from high-frequency data. Journal of Econometrics 215(2), 536–558.

Li, Z. M. and O. Linton (2022). A remedi for microstructure noise. Econometrica 90(1), 367–389.

Liu, X. (2015). Short-selling attacks and creditor runs. Management Science 61(4), 814–830.

Lou, D. (2012). A flow-based explanation for return predictability. The Review of Financial Studies 25(12), 3457–3489.

Lynch, A., B. Nikolic, X. S. Yan, and H. Yu (2014). Aggregate short selling, commonality, and stock market returns. Journal of Financial Markets 17, 199–229.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. Journal of financial economics 3(1-2), 125–144.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. Journal of financial economics 8(4), 323–361.

Mikusheva, A. (2012). One-dimensional inference in autoregressive models with the potential presence of a unit root. Econometrica 80(1), 173–212.

Milgrom, P. and N. Stokey (1982). Information, trade and common knowledge. Journal of economic theory 26(1), 17–27.

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. The Journal of finance 32(4), 1151–1168.

Muravyev, D., N. D. Pearson, and J. P. Broussard (2013). Is there price discovery in equity options? Journal of Financial Economics 107(2), 259–283.

Muravyev, D., N. D. Pearson, and J. M. Pollet (2018). Why does options market information predict stock returns? Available at SSRN 2851560.

Ng, S. and P. Perron (2001). Lag length selection and the construction of unit root tests with good size and power. Econometrica 69(6), 1519–1554.

Ni, S. X., N. D. Pearson, and A. M. Poteshman (2005). Stock price clustering on option expiration dates. Journal of Financial Economics 78(1), 49–87.

Obizhaeva, A. A. and J. Wang (2013). Optimal trading strategy and supply/demand dynamics. Journal of Financial markets 16(1), 1–32.

Pan, J. and A. M. Poteshman (2006). The information in option volume for future stock prices. The Review of Financial Studies 19(3), 871–908.

Pan, W.-F. (2020). Does investor sentiment drive stock market bubbles? beware of excessive optimism! Journal of Behavioral Finance 21(1), 27–41.

Pelger, M. (2020). Understanding systematic risk: A high-frequency approach. The Journal of Finance 75(4), 2179–2220.

Phillips, P. C., S. Shi, and J. Yu (2015b). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the s&p 500. International economic review 56(4), 1043–1078.

Phillips, P. C., S. Shi, and J. Yu (2015a). Testing for multiple bubbles: Limit theory of real-time detectors. International Economic Review 56(4), 1079–1134.

Phillips, P. C. and S.-P. Shi (2018). Financial bubble implosion and reverse regression. Econometric Theory 34(4), 705–753.

Phillips, P. C., Y. Wu, and J. Yu (2011). Explosive behavior in the 1990s nasdaq: When did exuberance escalate asset values? International economic review 52(1), 201–226.

Poterba, J. M. and L. H. Summers (1988). Mean reversion in stock prices: Evidence and implications. Journal of financial economics 22(1), 27–59.

Rapach, D. E., M. C. Ringgenberg, and G. Zhou (2016). Short interest and aggregate stock returns. Journal of Financial Economics 121(1), 46–65.

Savor, P. G. (2012). Stock returns after major price shocks: The impact of information. Journal of financial Economics 106(3), 635–659.

Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. Journal of political Economy 111(6), 1183–1220.

Shiller, R. J. et al. (1981). Do stock prices move too much to be justified by subsequent changes in dividends?

Shleifer, A. (1986). Do demand curves for stocks slope down? The Journal of Finance 41(3), 579–590.

Shleifer, A. and R. W. Vishny (1997). The limits of arbitrage. The Journal of Finance 52(1), 35–55.

Staiger, D. O. and J. H. Stock (1997). Instrumental variables regression with weak instruments. Econometrica 65(3), 557–586.

Stock, J. H. and J. H. Wright (2000). Gmm with weak identification. Econometrica 68(5), 1055–1096.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. Journal of the royal statistical society: Series B (Methodological) 36(2), 111–133.

Tirole, J. (1982). On the possibility of speculation under rational expectations. Econometrica: Journal of the Econometric Society, 1163–1181.

Todorov, V. and T. Bollerslev (2010). Jumps and betas: A new framework for disentangling and estimating systematic risks. Journal of Econometrics 157(2), 220–235.

Van Kervel, V. and A. J. Menkveld (2019). High-frequency trading around large institutional orders. The Journal of Finance 74(3), 1091–1137.

Vayanos, D. and P. Woolley (2013). An institutional theory of momentum and reversal. The Review of Financial Studies 26(5), 1087–1145.

Wang, X., X. S. Yan, and L. Zheng (2020). Shorting flows, public disclosure, and market efficiency. Journal of Financial Economics 135(1), 191–212.

Zhang, J.-Y., Y. Li, and Z.-M. Chen (2013). Unit root hypothesis in the presence of stochastic volatility, a bayesian analysis. Computational Economics 41, 89–100.

Please find full **version of the paper with Appendix** on **my website** . It discusses cleaning procedure, proofs, and other details not mentioned in the paper. For clarity, only referenced figures and tables appear here:

Tables

			Moments of HF Returns				
	N obs.	$\bar{r}_{t,t+l}$	Max	Min	s.d.	Skew	Kurt
Explosions Up							
No adj.							
Before	208,200	0.69	2.58	-2.38	0.31	0.29	38.51
Explosion	208,200	7.54	3.11	-2.11	0.31	1.49	47.51
After	208,200	0.21	2.62	-2.46	0.31	0.23	40.19
FF5							
Before	199,284	0.00	2.29	-2.15	0.28	0.20	38.10
Explosion	199,284	5.46	2.86	-1.95	0.29	1.40	48.65
After	199,284	0.01	2.38	-2.24	0.29	0.18	40.52
Explosions Down							
No adj.							
Before	241,669	0.15	2.58	-2.46	0.31	0.18	37.76
Explosion	241,669	-6.65	2.46	-3.14	0.34	-0.94	43.84
After	241,669	0.25	3.00	-2.80	0.35	0.34	40.32
FF5							
Before	232,066	-0.11	2.28	-2.19	0.28	0.12	37.82
Explosion	232,066	-4.68	2.20	-2.76	0.30	-0.87	43.86
After	232,066	0.05	2.66	-2.46	0.32	0.26	40.08

Table 3: The return properties 10 days before, after, and within explosive episode.

The second column reports the average 10-day return over the given period (10 days prior to the estimated explosive period, within the 10 days when the explosion is detected, and 10 days after the period). The remaining columns summarize the properties of the high-frequency returns within each of the three periods. The explosive intervals are selected using a non-overlapping procedure. Additionally, the ten-day intervals before and after the explosive interval are chosen for the same stocks. For a stock to be considered, it must have reliable price data for a thirty-day interval, an initial price of at least \$5, pass the bounce-back test around detection, and satisfy liquidity condition. See appendix for details.

				Moments of HF Returns				
Expl	N obs.	$\bar{r}_{t,t+l}$	r^{Expl}	Max	Min	s.d.	Skew	Kurt
No	104,921	0.66		1.00	-0.97	0.12	0.30	31.95
Down	8,654	-3.61	-5.06	1.12	-1.31	0.14	-0.57	31.42
Up	2,897	3.33	3.44	0.94	-0.70	0.10	1.17	34.28

Table 4: Portfolio return characteristics in explosion window

The first two columns display the average percentage return over a 10-day period and the return at the point of explosion. The remaining columns provide information on extreme short-term (5-minute) or overnight returns and the statistical moments (standard deviation, skewness, kurtosis) of their distribution. The initial row represents non-explosive episodes, determined at a 5% significance level for $k = 1$ and $r = 0.2$.

	No-Adjustment	Market	CAPM	FF3
Explosion				
1%	6.20	5.68	5.74	5.61
2.5%	9.01	8.27	8.31	8.13
5%	12.27	11.30	11.32	11.09
10%	17.18	15.88	15.86	15.55
Explosion Up				
1%	2.81	2.78	2.77	2.73
2.5%	4.10	4.07	4.03	3.98
5%	5.63	5.57	5.50	5.44
10%	7.98	7.84	7.70	7.64
Explosion Down				
1%	3.39	2.90	2.97	2.88
2.5%	4.90	4.20	4.28	4.15
5%	6.65	5.73	5.82	5.65
10%	9.21	8.05	8.15	7.91

Table 5: Frequency of explosion detection

The table provides a summary of the frequency of detecting explosions, both upward and downward, at 1%, 2.5%, 5%, and 10% significance levels. The analysis utilizes a lag order of $k = 1$, a 10-day observation period, and a timestamp frequency of 5 minutes. Stocks with an initial price below $5 are excluded from the analysis. Each column corresponds to the detection of explosiveness with different price risk adjustments. The "Just Market" specification involves scaling down the price by the gross market return since the beginning of the estimation period for explosion. "CAPM" and "FF3" correspond to adjustment by the gross return on the Market and Market, HML, SMB portfolios, respectively, with weights based on pre-estimated risk exposures (betas) of the stocks. These betas are estimated using high-frequency data over a 50-day period prior to the start of the estimation period for explosion.

	p-value $\leqslant 0.01$				p-value $\leqslant 0.05$			
	No Adj.	JM	CAPM	FF3	No Adj.	JM	CAPM	FF3
No Adj.	1.00	0.63	0.64	0.58	1.00	0.79	0.80	0.74
JM	0.69	1.00	0.83	0.72	0.85	1.00	0.94	0.87
CAPM	0.69	0.81	1.00	0.80	0.83	0.93	1.00	0.93
FF3	0.64	0.72	0.82	1.00	0.79	0.88	0.94	1.00

Table 6: Overlapping in detection of explosion

The table presents the frequency with which a 10-day episode marked as explosive at the 1% significance level, based on a specific adjustment specification (defined per row of the table), is also identified as explosive at the 1% or 5% significance level according to an alternative specification (defined per column of the table).

	Average	S.D.	Skewness	Kurtosis	10%	25%	50%	75%	90%
At explosion up	1.18	2.54	11.15	212.18	0.16	0.30	0.59	1.18	2.30
At explosion down	-1.23	2.51	-10.61	194.94	-2.45	-1.26	-0.63	-0.33	-0.18

Table 7: Distribution of the HF-return at explosion detection

	No-Adjustment		Market		CAPM		FF3	
	No	Yes	No	Yes	No	Yes	No	Yes
Panel 1: Share of Stocks Explosive Up conditional on Mkt-explosive 10 days								
1%	2.82	2.89	2.68	3.90	2.70	3.46	2.68	3.25
5%	5.64	5.50	5.38	7.52	5.35	6.89	5.32	6.53
Panel 2: Share of Stocks Explosive Up conditional on Mkt-Up-explosive 10 days								
1%	2.82	7.50	2.68	3.58	2.70	3.21	2.68	2.97
5%	5.64	14.16	5.38	6.86	5.35	6.12	5.32	5.81
Panel 3: Share of Stocks Explosive Down conditional on Mkt-explosive 10 days								
1%	2.84	8.84	2.76	3.89	2.86	3.80	2.77	3.65
5%	5.65	16.20	5.49	7.50	5.63	7.20	5.47	6.91
Panel 4: Share of Stocks Explosive Down conditional on Mkt-Down-explosive 10 days								
1%	2.84	10.41	2.76	3.94	2.86	3.90	2.77	3.73
5%	5.65	19.07	5.49	7.61	5.63	7.33	5.47	6.97
Panel 5: Share of Stock Explosions Up on the Mkt-Explosive date								
1%		5.32		7.52		6.34		5.66
5%		5.12		7.39		6.57		5.93
Panel 6: Share of Stock Explosions Down on the Mkt-Explosive date								
1%		19.39		7.31		6.84		6.56
5%		18.75		7.27		6.50		6.23

Table 8: Association of Explosive Episodes in the Market and Individual Stocks

The table summarizes the frequency of detecting explosive episodes, both up and down, in relation to the explosiveness of the market portfolio. The first four panels of the table provide information on the frequency of detecting an explosion in individual stocks, either up (Panels 1-2) or down (Panels 3-4), within a 10-day interval conditionally on whether there was no detection or a detection of explosiveness (either up or down) in the market portfolio. The sub-columns "No" and "Yes" indicate whether the market was detected to be explosive at 5% significance level. The fifth and sixth panels of the table examine whether the *detection date* of the detected stock explosion coincides with a market-explosive *detection date*. The reported numbers is the share of detection dates at various significance levels labeled by explosions for the market.

	Dependent variable: $r^{Ex.Port}$ - r_f					
	Based on Explosion Up			Based on Explosion Down		
	CAPM	FF-C	FF5	CAPM	FF-C	FF5
	(1)	(2)	(3)	(4)	(5)	(6)
Alpha	−0.010***	−0.010***	−0.010***	0.012***	0.013***	0.013***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
‘Mkt-RF‘	0.646***	0.590***	0.598***	0.855***	0.784***	0.771***
	(0.018)	(0.019)	(0.019)	(0.020)	(0.021)	(0.022)
SMB		0.268***	0.262***		0.384***	0.384***
		(0.032)	(0.033)		(0.040)	(0.042)
HML		0.094***	0.053		0.019	0.077*
		(0.028)	(0.033)		(0.035)	(0.042)
RMW			0.005			−0.018
			(0.046)			(0.060)
CMA			−0.129**			0.186**
			(0.058)			(0.074)
Mom		−0.021	−0.031		−0.078***	−0.062**
		(0.019)	(0.020)		(0.025)	(0.026)
Observations	4,001	4,001	4,001	3,850	3,850	3,850
R^2	0.252	0.268	0.269	0.329	0.346	0.347
Adjusted R^2	0.252	0.267	0.267	0.329	0.345	0.346
Residual Std. Error	0.011 (df = 3999)	0.010 (df = 3996)	0.010 (df = 3994)	0.014 (df = 3848)	0.014 (df = 3845)	0.014 (df = 3843)
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 9: Alpha of explosive portfolio

The equally weighted portfolios are constructed daily, guided by the explosion detection criteria. Each portfolio consists of a minimum of 10 stocks.

	Dependent variable: $r^{Ex.Port}$ - r_f					
	Based on Explosion Up			Based on Explosion Down		
	CAPM	FF-C	FF5	CAPM	FF-C	FF5
	(1)	(2)	(3)	(4)	(5)	(6)
Alpha	0.006***	0.007***	0.007***	0.009***	0.009***	0.009***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
‘Mkt-RF‘	−0.645***	−0.590***	−0.596***	0.837***	0.768***	0.752***
	(0.017)	(0.019)	(0.019)	(0.019)	(0.021)	(0.021)
SMB		−0.266***	−0.259***		0.364***	0.359***
		(0.031)	(0.033)		(0.040)	(0.042)
HML		−0.094***	−0.058*		0.028	0.094**
		(0.027)	(0.033)		(0.035)	(0.041)
RMW			0.008			−0.049
			(0.045)			(0.059)
CMA			0.115**			0.210***
			(0.057)			(0.072)
Mom		0.018	0.028		−0.085***	−0.066***
		(0.019)	(0.019)		(0.024)	(0.025)
Observations	4,001	4,001	4,001	3,850	3,850	3,850
R^2	0.256	0.271	0.272	0.326	0.343	0.344
Adjusted R^2	0.255	0.270	0.271	0.326	0.342	0.343
Residual Std. Error	0.010 (df = 3999)	0.010 (df = 3996)	0.010 (df = 3994)	0.014 (df = 3848)	0.013 (df = 3845)	0.013 (df = 3843)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 10: Alpha of explosive portfolio using Bid and Offer

The equally weighted portfolios are constructed daily, guided by the explosion detection criteria. Each portfolio consists of a minimum of 10 stocks. The actual bid and offer prices are employed for calculating the returns. The actual bid and offer prices are employed for calculating the returns. A buy-sell (sell-buy) strategy is applied following upward (downward) explosions, meaning that to compute returns, the offer (bid) price is used as the numerator and the bid (offer) price is used as the denominator.

Jump Up	Jump Down	N obs.	Abn. B.Pr. (%)	Abn. TV. (%)	ES (b.p.)	Abn. ES (b.p)	Abn. ES (Norm) (%)
No Explosions							
No	No	13,385,037	0.0	-0.9	20.9	-1.2	-3.9
No	Yes	97,607	1.1	68.8	38.8	5.9	31.3
Yes	No	114,159	2.4	65.6	39.6	3.8	21.4
Yes	Yes	5,548	2.3	88.6	75.4	13.6	51.6
Explosions Up							
No	No	183,719	9.3	44.7	20.6	-0.5	1.7
No	Yes	268	13.7	150.6	71.3	7.7	34.0
Yes	No	17,034	9.6	192.4	31.2	3.7	28.9
Yes	Yes	434	14.6	253.4	64.4	6.9	37.6
Explosions Down							
No	No	216,848	-6.7	35.6	21.5	1.9	12.8
No	Yes	20,518	-0.9	174.4	33.7	9.3	52.1
Yes	No	795	-7.5	91.4	63.2	18.6	68.3
Yes	Yes	1,049	-4.2	175.7	60.9	22.1	92.2

Table 11: Average Liquidity Measures around Explosions and Jumps

The table presents average daily abnormal buying pressure, abnormal trading volume, and effective spread measures on the day of detection and the preceding day for explosions, jumps, or both. Data for daily measures are sourced from Intraday Indicators by WRDS. Daily buying pressure is defined as the order imbalance, representing the difference between the dollar trading volume assigned to buyer-initiated and seller-initiated trades, classified using Lee and Ready's (1991) algorithm. The abnormal characteristics in the fourth, fifth, and eighth columns represent the normalized difference between the respective daily measure and the rolling average of the characteristic from the previous 30 days. Abnormal trading volume and buying pressure in the fourth and fifth columns are normalized by the average daily trading volume over the previous 30 days. The abnormal effective spread in the last column is normalized by the average effective spread over 30 days. The reported characteristics represent the average of these abnormal liquidity measures over two days: event detection and the preceding date. The sixth and seventh columns report (not normalized) the average dollar effective spread and abnormal effective spread in basis points. .

Dep.Var	Abn. B.Pr		Abn. TV		Abn. ES	
	(1)	(2)	(3)	(4)	(5)	(6)
Expl. Up	8.5***	8.2***	45.2***	44.1***	0.99***	1.0***
	(0.12)	(0.13)	(0.60)	(0.65)	(0.03)	(0.04)
Expl. Down	-5.6***	-5.5***	32.3***	32.2***	2.3***	2.2***
	(0.10)	(0.10)	(0.56)	(0.60)	(0.06)	(0.06)
Jump Up	2.1***	1.8***	62.3***	65.8***	4.6***	4.8***
	(0.09)	(0.11)	(1.5)	(1.7)	(0.11)	(0.13)
Jump Down	1.2***	1.7***	66.3***	70.0***	5.8***	6.0***
	(0.12)	(0.14)	(1.7)	(1.9)	(0.17)	(0.19)
Expl. Up × Jump Up	-1.5***	-2.0***	84.0***	84.0***	-0.71***	-0.57***
	(0.26)	(0.29)	(2.6)	(3.0)	(0.15)	(0.18)
Expl. Down × Jump Up	-3.9***	-3.5***	-43.0***	-48.6***	7.5***	7.5***
	(0.72)	(0.89)	(7.3)	(7.6)	(1.0)	(1.2)
Expl. Up × Jump Down	2.8*	2.9	8.6	-3.9	-1.7**	0.24
	(1.7)	(2.5)	(9.4)	(12.3)	(0.80)	(1.0)
Expl. Down × Jump Down	4.2***	4.7***	73.2***	75.5***	0.50**	0.44
	(0.25)	(0.29)	(5.5)	(5.8)	(0.25)	(0.28)
Observations	14,043,016	10,907,987	14,043,016	10,907,987	14,043,016	10,907,987
R^2	0.02984	0.03409	0.14605	0.15684	0.07837	0.08842
Firm fixed effects	✓	✓	✓	✓	✓	✓
Date fixed effects	✓	✓	✓	✓	✓	✓

Table 12: Liquidity Measures around Explosions and Jumps

The table presents the association between explosive events and the dependent variables: abnormal trading volume, abnormal buying pressure, and abnormal effective spread averaged over two days around the detection. Refer to Table 11 for detailed definitions. The 1st, 3rd, and 5th columns report two-way fixed-effect regression with firm and date controls. The 2nd, 4th, and 6th columns report two-way fixed-effect regression with firm and date controls augmented by standard firm-characteristic controls. Standard errors, reported in parentheses, are two-way clustered by firm and date. The control variables include idiosyncratic variance, market equity (ME), book-to-market ratio (BtM), factor exposures with respect to the FF3 model over a 120-day interval (β_{Mkt}, β_{SMB}, and β_{HML}), short-term reversal, momentum, turnover, change of assets, and operating profitability, Amihud Liquidity, and Institutional Ownership. $Turnover$ is the average daily trading volume over a 20-day period, lagged by 20 days. Momentum is the return over the period from $t = 250$ to $t = 20$ days before the date. Short-term reversal is the preceding monthly return laggged by ME, BtM, Change of Assets, and Operating profitability are defined consistently with Fama-French. Idiosyncratic volatility (IV) is realized idiosyncratic volatility in the prior days using an exponentially weighted moving average (EWMA) model consistent with RiskMetrics. The sample period spans from 2003 to 2022, and statistical significance is indicated at the 10%, 5%, and 1% levels by *, **, ***, respectively.

Year Group	Explosion	Small	Medium	Large	Small	Medium	Large
2003-2008	No	328,145	1,520,104	1,746,256	95.9	96.9	97.3
	Up	7,190	24,481	24,596	2.1	1.6	1.4
	Dn	6,765	23,765	23,977	2.0	1.5	1.3
2009-2013	No	196,027	1,257,021	1,480,401	96.2	97.2	97.5
	Up	4,159	19,307	19,373	2.0	1.5	1.3
	Dn	3,516	17,341	18,728	1.7	1.3	1.2
2014-2018	No	298,795	1,272,315	1,406,493	96.8	97.5	97.4
	Up	5,248	16,336	17,294	1.7	1.3	1.2
	Dn	4,542	16,059	20,012	1.5	1.2	1.4
2018-2022	No	225,631	1,002,090	1,239,464	97.4	97.7	97.5
	Up	3,628	11,360	14,296	1.6	1.1	1.1
	Dn	2,460	12,045	17,237	1.1	1.2	1.4
All	No	1,048,598	5,051,530	5,872,614	96.5	97.3	97.4
	Up	20,225	71,484	75,559	1.9	1.4	1.3
	Dn	17,283	69,210	79,954	1.6	1.3	1.3

Table 13: Number of observations for return decomposition analysis in Subsection 5.3.

The table reports summary on the number of stocks in the explosion groups. The stocks are pre-classified into three equal groups based on day-specific size, measured 30 days prior. As small stocks are more susceptible to exclusion during cleaning procedures, the primary sample is predominantly composed of medium and large stocks

	Logit		Ridge		ENet		Lasso	
Variable	Estimate	(Q5, Q95)	Estimate	(Q5, Q95)	Estimate	(Q5, Q95)	Estimate	(Q5, Q95)
Intercept	-2.71	(-2.71, -2.70)	-2.71	(-2.71, -2.70)	-2.71	(-2.71, -2.70)	-2.71	(-2.71, -2.70)
a2me	0.01	(0.00, 0.02)	-0.02	(-0.03, -0.01)	0.00	(0.00, 0.00)	0.00	(0.00, 0.00)
assets	-0.04	(-0.11, 0.03)	-0.09	(-0.10, -0.07)	-0.05	(-0.10, 0.01)	-0.05	(-0.10, 0.01)
ato	-0.03	(-0.07, 0.00)	-0.01	(-0.03, 0.01)	-0.01	(-0.02, 0.01)	-0.01	(-0.02, 0.01)
betahml	0.02	(-0.01, 0.04)	0.02	(0.01, 0.04)	0.00	(-0.02, 0.00)	0.00	(-0.02, 0.00)
betamkt	-0.12	(-0.14, -0.10)	-0.10	(-0.11, -0.08)	-0.11	(-0.13, -0.09)	-0.11	(-0.13, -0.09)
betasmb	-0.04	(-0.06, -0.02)	-0.02	(-0.04, 0.00)	-0.01	(-0.02, 0.02)	-0.01	(-0.02, 0.02)
bidask	0.03	(-0.01, 0.07)	0.05	(0.03, 0.07)	0.00	(-0.04, 0.00)	0.00	(-0.04, 0.00)
bm	-0.09	(-0.15, -0.03)	-0.06	(-0.08, -0.04)	-0.09	(-0.11, -0.05)	-0.09	(-0.11, -0.05)
c	-0.04	(-0.08, -0.01)	-0.02	(-0.04, -0.01)	0.00	(0.00, 0.02)	0.00	(0.00, 0.02)
cto	-0.01	(-0.09, 0.06)	0.00	(-0.02, 0.02)	0.00	(0.00, 0.00)	0.00	(0.00, 0.00)
e2p	-0.11	(-0.15, -0.07)	-0.07	(-0.09, -0.05)	-0.09	(-0.13, -0.07)	-0.09	(-0.13, -0.07)
freecf	-0.02	(-0.06, 0.01)	-0.01	(-0.03, 0.01)	0.00	(0.00, 0.02)	0.00	(0.00, 0.02)
idiovol	0.10	(0.07, 0.13)	0.08	(0.06, 0.10)	0.08	(0.06, 0.11)	0.08	(0.06, 0.11)
intmom	0.14	(0.12, 0.17)	0.07	(0.06, 0.09)	0.06	(0.03, 0.09)	0.06	(0.03, 0.09)
invest	0.07	(0.04, 0.10)	0.05	(0.03, 0.07)	0.05	(0.03, 0.07)	0.05	(0.03, 0.07)
lev	0.05	(0.02, 0.09)	0.04	(0.02, 0.06)	0.02	(0.00, 0.04)	0.02	(0.00, 0.04)
ltrev	0.03	(0.01, 0.05)	0.02	(0.00, 0.03)	0.02	(0.00, 0.04)	0.02	(0.00, 0.04)
mktcap	-0.29	(-0.36, -0.22)	-0.14	(-0.16, -0.13)	-0.24	(-0.30, -0.19)	-0.25	(-0.32, -0.20)
mom	-0.18	(-0.22, -0.14)	-0.09	(-0.11, -0.07)	-0.07	(-0.10, -0.04)	-0.07	(-0.10, -0.04)
noa	-0.04	(-0.08, -0.01)	-0.01	(-0.03, 0.01)	0.00	(0.00, 0.00)	0.00	(0.00, 0.00)
oa	0.06	(0.04, 0.08)	0.04	(0.02, 0.05)	0.04	(0.02, 0.06)	0.04	(0.02, 0.06)
ol	-0.00	(-0.06, 0.08)	-0.00	(-0.01, 0.02)	0.00	(0.00, 0.00)	0.00	(0.00, 0.00)
pcm	0.01	(-0.04, 0.05)	-0.02	(-0.04, 0.00)	-0.00	(-0.00, 0.03)	-0.00	(-0.00, 0.02)
pm	-0.04	(-0.08, 0.00)	-0.04	(-0.07, -0.03)	-0.02	(-0.04, 0.01)	-0.02	(-0.04, 0.01)
prof	-0.05	(-0.10, 0.00)	-0.00	(-0.01, 0.02)	0.00	(0.00, 0.02)	-0.00	(-0.00, 0.02)
q	0.10	(0.02, 0.18)	0.02	(0.01, 0.04)	0.00	(-0.01, 0.00)	0.00	(-0.01, 0.00)
rna	0.05	(0.01, 0.09)	0.06	(0.04, 0.08)	0.03	(-0.01, 0.06)	0.03	(-0.01, 0.06)
roa	-0.09	(-0.16, -0.01)	-0.03	(-0.04, -0.01)	-0.00	(-0.00, 0.04)	0.00	(0.00, 0.04)
roe	0.11	(0.04, 0.17)	0.02	(0.00, 0.03)	0.00	(0.00, 0.00)	0.00	(0.00, 0.00)
s2p	0.07	(-0.01, 0.14)	0.00	(-0.02, 0.01)	0.00	(0.00, 0.00)	0.00	(0.00, 0.00)
strev	-0.08	(-0.10, -0.05)	-0.05	(-0.07, -0.04)	-0.03	(-0.05, -0.01)	-0.03	(-0.05, -0.01)
suv	-0.15	(-0.18, -0.13)	-0.11	(-0.13, -0.10)	-0.11	(-0.13, -0.08)	-0.11	(-0.13, -0.09)
turn	0.17	(0.14, 0.19)	0.11	(0.09, 0.12)	0.12	(0.09, 0.14)	0.12	(0.09, 0.14)
w52h	0.11	(0.08, 0.15)	0.01	(-0.01, 0.03)	0.00	(-0.02, 0.00)	0.00	(-0.02, 0.00)

Table 14: Logistic Regression of Explosiveness Up against Firm Characteristics

This table presents estimates from logistic and penalized logistic regressions using the indicator of explosion up detection over a week as the dependent variable. The explanatory variables include a set of firm characteristics described in Freyberger et al. (2020). The results are reported for elastic net penalization with hyperparameter $\alpha = 0$ (Ridge regression), $\alpha = 0.5$, and $\alpha = 1$ (Lasso regression). The penalty hyperparameter λ, which controls the strength of the penalty, is determined by minimizing the error through 10-fold cross-validation (see Hastie et al. (2009) for details). Bootstrapped 90% confidence intervals are also provided for the results. The sample covers 1,575,707 week-firm observations since September 2003 to December 2022.

	Logit		Ridge		ENet		Lasso	
Variable	Estimate	(Q5, Q95)	Estimate	(Q5, Q95)	Estimate	(Q5, Q95)		
Intercept	-2.54	(-2.54, -2.53)	-2.54	(-2.54, -2.53)	-2.54	(-2.54, -2.53)	-2.54	(-2.54, -2.53)
a2me	0.03	(0.02, 0.03)	0.02	(0.00, 0.03)	0.00	(0.00, 0.00)	0.00	(0.00, 0.00)
assets	-0.17	(-0.24, -0.10)	-0.09	(-0.12, -0.04)	0.00	(0.00, 0.02)	0.00	(0.00, 0.01)
ato	-0.05	(-0.09, -0.02)	-0.03	(-0.06, 0.00)	-0.01	(-0.02, 0.02)	-0.01	(-0.02, 0.02)
betahml	0.07	(0.05, 0.09)	0.06	(0.04, 0.07)	0.05	(0.03, 0.07)	0.05	(0.03, 0.07)
betamkt	0.01	(-0.01, 0.03)	0.02	(0.00, 0.04)	0.01	(-0.01, 0.02)	0.01	(-0.01, 0.02)
betasmb	0.04	(0.02, 0.07)	0.04	(0.02, 0.06)	0.03	(0.01, 0.05)	0.03	(0.01, 0.05)
bidask	-0.20	(-0.23, -0.16)	-0.18	(-0.21, -0.15)	-0.18	(-0.21, -0.15)	-0.18	(-0.21, -0.15)
bm	-0.14	(-0.20, -0.09)	-0.10	(-0.13, -0.06)	-0.08	(-0.11, -0.05)	-0.08	(-0.11, -0.06)
c	-0.02	(-0.06, 0.01)	-0.01	(-0.03, 0.02)	0.00	(0.00, 0.03)	0.00	(0.00, 0.03)
cto	0.22	(0.15, 0.29)	0.10	(0.05, 0.13)	0.00	(-0.02, 0.00)	0.00	(-0.01, 0.00)
e2p	-0.04	(-0.07, -0.01)	-0.05	(-0.08, -0.03)	-0.04	(-0.08, -0.02)	-0.04	(-0.07, -0.01)
freecf	-0.07	(-0.10, -0.04)	-0.06	(-0.09, -0.04)	-0.04	(-0.07, -0.01)	-0.04	(-0.07, -0.01)
idiovol	-0.05	(-0.08, -0.03)	-0.02	(-0.04, 0.01)	-0.01	(-0.02, 0.01)	-0.01	(-0.02, 0.01)
intmom	-0.21	(-0.24, -0.19)	-0.14	(-0.16, -0.12)	-0.15	(-0.17, -0.12)	-0.15	(-0.17, -0.12)
invest	0.03	(0.00, 0.06)	0.06	(0.04, 0.09)	0.07	(0.05, 0.09)	0.07	(0.05, 0.09)
lev	0.05	(0.01, 0.08)	0.06	(0.04, 0.09)	0.05	(0.02, 0.07)	0.05	(0.03, 0.08)
ltrev	0.10	(0.08, 0.12)	0.10	(0.08, 0.12)	0.10	(0.08, 0.12)	0.10	(0.08, 0.12)
mktcap	0.12	(0.04, 0.19)	0.07	(0.02, 0.10)	0.00	(-0.03, 0.00)	0.00	(-0.03, 0.00)
mom	0.70	(0.67, 0.74)	0.58	(0.55, 0.60)	0.62	(0.58, 0.64)	0.63	(0.60, 0.66)
noa	0.03	(-0.01, 0.07)	0.04	(0.02, 0.07)	0.05	(0.03, 0.08)	0.05	(0.03, 0.08)
oa	0.02	(0.00, 0.04)	0.02	(0.00, 0.04)	0.01	(0.00, 0.02)	0.01	(0.00, 0.02)
ol	-0.20	(-0.27, -0.13)	-0.09	(-0.11, -0.04)	-0.02	(-0.04, 0.01)	-0.01	(-0.02, 0.03)
pcm	0.05	(0.02, 0.09)	0.02	(-0.01, 0.05)	0.00	(0.00, 0.01)	0.00	(0.00, 0.01)
pm	-0.03	(-0.07, 0.00)	-0.02	(-0.04, 0.01)	0.00	(0.00, 0.03)	0.00	(0.00, 0.03)
prof	-0.10	(-0.14, -0.06)	-0.06	(-0.08, -0.02)	-0.04	(-0.07, -0.02)	-0.04	(-0.07, -0.02)
q	-0.04	(-0.12, 0.03)	-0.03	(-0.07, -0.01)	0.00	(0.00, 0.00)	0.00	(0.00, 0.00)
rna	-0.04	(-0.08, 0.00)	-0.01	(-0.03, 0.03)	0.00	(0.00, 0.01)	0.00	(0.00, 0.00)
roa	-0.13	(-0.20, -0.06)	-0.08	(-0.12, -0.04)	-0.03	(-0.06, 0.01)	-0.03	(-0.06, 0.01)
roe	0.11	(0.04, 0.17)	0.06	(0.01, 0.09)	0.00	(0.00, 0.00)	0.00	(0.00, 0.00)
s2p	0.02	(-0.06, 0.09)	-0.00	(-0.04, 0.05)	0.00	(0.00, 0.00)	0.00	(0.00, 0.01)
strev	0.13	(0.10, 0.15)	0.09	(0.07, 0.11)	0.10	(0.08, 0.12)	0.10	(0.08, 0.12)
suv	-0.09	(-0.11, -0.07)	-0.09	(-0.10, -0.07)	-0.09	(-0.11, -0.07)	-0.09	(-0.11, -0.07)
turn	0.14	(0.12, 0.17)	0.15	(0.13, 0.17)	0.14	(0.11, 0.16)	0.14	(0.12, 0.16)
w52h	-0.53	(-0.56, -0.50)	-0.43	(-0.46, -0.41)	-0.45	(-0.47, -0.42)	-0.45	(-0.48, -0.41)

Table 15: Logistic Regression of Explosiveness Down against Firm Characteristics

This table presents estimates from logistic and penalized logistic regressions using the indicator of explosion down detection over a week as the dependent variable. The explanatory variables include a set of firm characteristics described in Freyberger et al. (2020). The results are reported for elastic net penalization with hyperparameter $\alpha = 0$ (Ridge regression), $\alpha = 0.5$, and $\alpha = 1$ (Lasso regression). The penalty hyperparameter λ, which controls the strength of the penalty, is determined by minimizing the error through 10-fold cross-validation (see Hastie et al. (2009) for details). Bootstrapped 90% confidence intervals are also provided for the results. The sample covers 1,575,707 week-firm observations since September 2003 to December 2022.

Variables	*Mean*	*S.D.*	*Q-10%*	*Q-25%*	*Q-50%*	*Q-75%*	*Q-90%*
SIR	5.16	5.77	0.48	1.45	3.17	6.83	12.29
ΔSIR	0.01	0.85	-0.60	-0.20	0.00	0.20	0.64
SIR^{po}	0.14	4.78	-4.12	-2.48	-0.78	1.37	5.22
ΔSIR^{po}	0.00	1.33	-1.45	-0.65	0.02	0.69	1.45
DtC	6.66	6.67	1.30	2.50	4.74	8.59	14.02
ΔDtC	0.01	1.04	-0.81	-0.33	0.00	0.32	0.84
$\log r_{t,t+9}$ (%)	-0.09	8.98	-9.14	-3.77	0.22	4.07	8.74
$r^{Adj.CAPM}_{t,t+9}$ (%)	-0.11	8.37	-8.08	-3.66	-0.22	3.16	7.72
$r^{Adj.FF3}_{t,t+9}$ (%)	0.02	8.22	-7.69	-3.44	-0.13	3.17	7.65
IV (%)	2.16	1.51	0.90	1.23	1.79	2.63	3.75
Am.Liq	0.64	6.33	0.00	0.00	0.00	0.02	0.17
Log-Size	12.95	1.83	10.68	11.64	12.84	14.12	15.42
ME	7.37	36.09	0.11	0.30	1.01	3.59	13.18
Op. Prof	0.21	1.06	-0.05	0.11	0.20	0.30	0.46
Δ Assets	0.16	0.57	-0.09	-0.01	0.06	0.17	0.40
BtM	0.58	0.61	0.12	0.26	0.49	0.79	1.14
β_{Mkt}	0.94	0.51	0.31	0.65	0.94	1.23	1.53
β_{SMB}	0.70	0.79	-0.20	0.16	0.63	1.15	1.68
β_{HML}	0.08	0.89	-0.90	-0.36	0.09	0.54	1.05
ST-Reversal	0.00	0.13	-0.13	-0.05	0.01	0.06	0.13
Momentum	0.08	0.42	-0.40	-0.13	0.09	0.29	0.52

Table 16: Summary Statistics for SI analysis

Summary statistics for the variables used in the analysis around the dissemination of short interest (SI). The first two columns show the average and standard deviation of the variables in the sample, while the other columns display the respective 10%, 25%, 50%, 75%, and 90% quantiles. The sample includes U.S. companies with share codes 10 and 11, primary exchange listings on NYSE, Amex, or NASDAQ, matched data in TAQ from September 2003 to December 2022, and estimated explosiveness following the dissemination date. SIR: Short Interest Ratio, disclosed by FINRA and normalized by shares outstanding from CRSP. ΔSIR: Change in reported SIR since the last reported SI for the firm (typically a month before February 2007, and half a month after). SIR^{po}: The residual of a cross-sectional regression of SIR against the indicator of being in a quintile-group on a given day based on idiosyncratic variance, market equity (ME), book-to-market ratio (BtM), factor exposures with respect to FF3 model over a 120-day interval (β_{Mkt}, β_{SMB}, and β_{HML}), short-term reversal, momentum, and industry identified by the first two digits of the SIC code. *Turnover* is average daily trading volume over a 20-day period, lagged by 20 days. Days-to-Cover (DTC) is SIR normalized by turnover. *Amihud Liquidity* measure is defined according to standard practice. ST-Reversal is the 20-day return prior to the dissemination date. *Momentum* is return over the period from $t = 250$ to $t = 20$ days before the date. ME, BtM, Change of Assets, and Operating profitability are defined consistently with Fama-French. *Idiosyncratic volatility* (IV) is realized idiosyncratic volatility in the prior days using an exponentially weighted moving average (EWMA) model consistent with RiskMetrics. $\log r_{t,t+9}$ is log-return over a 10-day interval (excluding the first overnight return, as used in explosiveness estimation). $r^{Adj.CAPM}_{t,t+9}$ and $r^{Adj.FF3}_{t,t+9}$ are adjusted 10-day CRSP returns expressed in percentage terms, based on pre-estimated factor exposures lagged by 31 days over a 120-day interval using the CAPM and Fama-French 3 models. All variables are winsorized at 0.1% from both sides.

Dep.Var	OLS			FE-LS		
	Ret (%)	Ret CAPM-Adj. (%)	Ret FF3-Adj. (%)	Ret (%)	Ret CAPM-Adj. (%)	Ret FF3-Adj. (%)
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-0.16***	-0.05***	0.07***			
	(0.009)	(0.010)	(0.009)			
High SIR^{po}	-0.66***	-0.35***	-0.28***	-0.38***	-0.29***	-0.29***
	(0.03)	(0.03)	(0.03)	(0.07)	(0.05)	(0.05)
Turnover				-4.0**	-0.39	0.47
				(1.5)	(2.3)	(2.0)
IV				-5.5	-3.6	-2.7
				(3.6)	(4.1)	(4.0)
$Am.Liq$				-0.0009	0.003**	0.003**
				(0.001)	(0.0008)	(0.0010)
Log-Size				-0.65***	-0.89***	-0.88***
				(0.08)	(0.09)	(0.08)
BtM				0.09*	0.18***	0.18***
				(0.04)	(0.04)	(0.04)
Δ Assets				-0.03	-0.04	-0.03
				(0.03)	(0.03)	(0.02)
Op. Prof				0.008	-0.01	-0.009
				(0.009)	(0.01)	(0.01)
IOR				-0.15	-0.16	-0.17
				(0.10)	(0.12)	(0.12)
ST-Reversal				-2.5***	-3.3***	-3.5***
				(0.45)	(0.51)	(0.50)
Momentum				-0.10	-0.39**	-0.40**
				(0.14)	(0.15)	(0.13)
$Q5-Am.Liq$				0.58**	0.64**	0.66**
				(0.17)	(0.19)	(0.17)
Standard-Errors		IID		3-way clustered: Date & Size-BtM Portfolio & Firm		
Observations	1,412,719	1,412,717	1,412,714	1,295,065	1,295,064	1,295,061
R^2	0.00036	9.75×10^{-5}	6.12×10^{-5}	0.23944	0.09764	0.07039
Date-Industry fixed effects				✓	✓	✓
Firm fixed effects				✓	✓	✓
Quintile fixed effects				✓	✓	✓

Table 17: Return following SI dissemination

The table presents the regression analysis of predictive power of Short Interest Ratio in 10-day subsequent returns following the dissemination of short interest (SI). The dependent variables are the log-return over a 10-day interval (excluding the first overnight return, as used in explosiveness estimation), and adjusted 10-day CRSP returns, expressed in percentage terms. The adjustments are based on pre-estimated factor exposures lagged by 31 days over a 120-day interval, according to CAPM and Fama-French 3 models. The primary regressor in the first three columns is the treatment of being in the top decile of partialled out short interest, without the inclusion of controls. In the next three columns, fixed-effect regressions with a variety of control variables are presented. The definitions of the control variables are consistent with those in the summary statistics table. The control variables also encompass indicators for being in a quintile-group on a given day based on IV, ME, BtM, β_{Mkt}, β_{SMB}, and β_{HML}. For clarity and simplicity, only the coefficient on the most liquid stocks, $Q5-Am.Liq$, is included in the table, as it is the only consistently significant coefficient. The regression coefficients are reported with standard errors in columns 1-3 and are 3-way clustered by Date, Firm, and Size-BtM Portfolios in columns 4-6. The sample period spans from 2003 to 2022, and statistical significance is indicated at the 10%, 5%, and 1% levels by *, **, ***, respectively.

Panel A: *Fixed-Effect Least Square Estimates*						
	Use *SIR* as High SI			Use SIR^{po} as High SI		
Dep.Var	Expl.Up (%)	Expl.Up CAPM-Adj. (%)	Expl.Up FF3-Adj. (%)	Expl.Up (%)	Expl.Up CAPM-Adj. (%)	Expl.Up FF3-Adj. (%)
	(1)	(2)	(3)	(4)	(5)	(6)
High SI	0.53***	0.49***	0.31**	0.60***	0.64***	0.45***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.12)	(0.12)
Observations	1,060,913	1,039,901	1,039,901	1,060,913	1,039,901	1,039,901
R^2	0.07486	0.06192	0.05646	0.07487	0.06194	0.05647
Date-Industry fixed effects	✓	✓	✓	✓	✓	✓
Firm fixed effects	✓	✓	✓	✓	✓	✓

Panel B: *Fixed-Effect Logit Estimates*						
	Use Q-10 by *SIR* as High SI			Use Q-10 by SIR^{po} as High SI		
Dep.Var	Expl.Up	Expl.Up CAPM-Adj.	Expl.Up FF3-Adj.	Expl.Up	Expl.Up CAPM-Adj.	Expl.Up FF3-Adj. (%)
	(1)	(2)	(3)	(4)	(5)	(6)
High SI	0.09***	0.09***	0.06**	0.11***	0.11***	0.08***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	881,370	883,236	889,078	881,370	883,236	889,078
R^2	0.06670	0.05482	0.04961	0.06673	0.05487	0.04963
Date-Industry fixed effects	✓	✓	✓	✓	✓	✓
Firm fixed effects	✓	✓	✓	✓	✓	✓

Table 18: Explosiveness Up following SI dissemination.

The table presents the predictive power of the short interest ratio in 10-day subsequent returns following the dissemination of short interest (SI). The dependent variables are indicators of detecting an upward explosion over at least one of the five subsequent (overlapping) 10-day intervals following the dissemination of short interest. Columns (1-3) and (4-6) report the results for detection using raw price data (adjusted only for dividends and stock splits), CAPM-adjusted price data, and FF3-adjusted price data, respectively. The first three columns present results where the Short Interest Ratio (*SIR*) is used to select the top decile of heavily shorted stocks as the treatment group. The other three columns present results where the partialled-out *SIR* is used. The control variables and fixed effects are consistent with the controls used for return regressions reported in Table 17. Significant coefficients for all controls are reported in Table 19. The sample period spans from 2003 to 2022, and statistical significance is indicated at the 10%, 5%, and 1% levels by *, **, ***, respectively.

Dep.Var	Use SIR as High SI			Use SIR^{po} as High SI		
	Expl.Up (%)	Expl.Up CAPM-Adj. (%)	Expl.Up FF3-Adj. (%)	Expl.Up (%)	Expl.Up CAPM-Adj. (%)	Expl.Up FF3-Adj. (%)
	(1)	(2)	(3)	(4)	(5)	(6)
High SI	0.53***	0.49***	0.31**	0.60***	0.64***	0.45***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.12)	(0.12)
Turnover	4.8	4.5	5.5*	4.6	4.1	5.1
	(3.2)	(3.5)	(3.1)	(3.2)	(3.5)	(3.1)
IV	20.2***	15.3***	15.6***	20.1***	15.3***	15.7***
	(4.8)	(5.2)	(5.1)	(4.8)	(5.3)	(5.1)
$Am.Liq$	-0.005	0.002	0.004	-0.005	0.002	0.004
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Log-Size	-1.1***	-1.1***	-1.2***	-1.1***	-1.1***	-1.2***
	(0.11)	(0.12)	(0.11)	(0.11)	(0.12)	(0.11)
BtM	0.10	0.11	0.14	0.10	0.11	0.13
	(0.10)	(0.11)	(0.11)	(0.10)	(0.11)	(0.11)
Δ Assets	0.02	0.05	0.03	0.02	0.05	0.03
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Op. Prof	-0.05	-0.04	-0.01	-0.05	-0.04	-0.01
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
IOR	-1.1***	-1.3***	-1.4***	-1.1***	-1.3***	-1.5***
	(0.24)	(0.25)	(0.23)	(0.24)	(0.24)	(0.23)
ST-Reversal	-3.2***	-3.3***	-2.9***	-3.2***	-3.3***	-2.9***
	(0.39)	(0.40)	(0.36)	(0.39)	(0.40)	(0.36)
Momentum	-0.60***	-0.61***	-0.55***	-0.61***	-0.63***	-0.56***
	(0.14)	(0.14)	(0.13)	(0.14)	(0.14)	(0.13)
$Q2-BtM$	-0.22*	-0.22*	-0.19*	-0.24**	-0.23**	-0.19*
	(0.12)	(0.11)	(0.11)	(0.12)	(0.11)	(0.11)
$Q2-\beta_{SMB}$	-0.17**	-0.06	-0.11	-0.17*	-0.06	-0.11
	(0.09)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)
$Q3-ME$	-0.58*	-0.45	-0.28	-0.53*	-0.40	-0.25
	(0.30)	(0.30)	(0.29)	(0.30)	(0.30)	(0.29)
$Q3-BtM$	-0.42***	-0.27**	-0.31**	-0.44***	-0.29**	-0.33**
	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)
$Q3-\beta_{SMB}$	-0.32***	-0.19*	-0.17*	-0.31***	-0.19*	-0.16
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
$Q4-Am.Liq$	0.80***	0.29	0.07	0.83***	0.34	0.11
	(0.30)	(0.27)	(0.28)	(0.30)	(0.27)	(0.28)
$Q4-ME$	-0.88***	-0.57*	-0.30	-0.81**	-0.50	-0.25
	(0.34)	(0.34)	(0.33)	(0.34)	(0.34)	(0.33)
$Q4-BtM$	-0.61***	-0.52***	-0.50***	-0.63***	-0.53***	-0.51***
	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)
$Q4-\beta_{Mkt}$	-0.23*	-0.29**	-0.34***	-0.22*	-0.29**	-0.34***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
$Q4-\beta_{SMB}$	-0.53***	-0.33***	-0.27**	-0.52***	-0.32***	-0.26**
	(0.10)	(0.11)	(0.11)	(0.10)	(0.11)	(0.11)
$Q5-Am.Liq$	-0.38	-0.75**	-0.72**	-0.33	-0.68*	-0.67*
	(0.38)	(0.36)	(0.36)	(0.38)	(0.36)	(0.37)
$Q5-ME$	-0.65*	-0.38	-0.30	-0.62	-0.34	-0.27
	(0.39)	(0.39)	(0.39)	(0.39)	(0.39)	(0.39)
$Q5-BtM$	-0.44**	-0.28	-0.33	-0.45**	-0.30	-0.33
	(0.21)	(0.20)	(0.20)	(0.21)	(0.20)	(0.20)
$Q5-\beta_{Mkt}$	-0.38**	-0.34**	-0.40***	-0.36**	-0.32**	-0.39***
	(0.15)	(0.15)	(0.14)	(0.15)	(0.15)	(0.14)
$Q5-\beta_{SMB}$	-0.67***	-0.54***	-0.48***	-0.66***	-0.52***	-0.47***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Observations	1,060,913	1,039,901	1,039,901	1,060,913	1,039,901	1,039,901
R^2	0.07486	0.06192	0.05646	0.07487	0.06194	0.05647
Date-Industry fixed effects	✓	✓	✓	✓	✓	✓
Firm fixed effects	✓	✓	✓	✓	✓	✓

Table 19: Explosiveness Up following SI dissemination. Least Square Estimates.

The table presents the predictive power of the short interest ratio in 10-day subsequent returns following the dissemination of short interest (SI). The dependent variables are indicators of detecting an upward explosion over at least one of the five subsequent (overlapping) 10-day intervals following the dissemination of short interest. See details in decription of Table 18

Dep.Var	Use Q-10 by *SIR* as High SI			Use Q-10 by SIR^{po} as High SI		
	Expl.Up (%)	Expl.Up CAPM-Adj. (%)	Expl.Up FF3-Adj. (%)	Expl.Up (%)	Expl.Up CAPM-Adj. (%)	Expl.Up FF3-Adj. (%)
	(1)	(2)	(3)	(4)	(5)	(6)
$Q1-IV$ × High SI	1.0***	0.93***	0.68**	0.80***	0.92***	0.70***
	(0.31)	(0.28)	(0.26)	(0.23)	(0.21)	(0.21)
$Q2-IV$ × High SI	0.23	0.45**	0.06	0.21	0.53***	0.17
	(0.20)	(0.20)	(0.20)	(0.18)	(0.18)	(0.18)
$Q3-IV$ × High SI	0.67***	0.34*	0.36**	0.67***	0.52***	0.47***
	(0.17)	(0.18)	(0.17)	(0.17)	(0.17)	(0.17)
$Q4-IV$ × High SI	0.22	0.43**	0.21	0.41*	0.54***	0.30
	(0.20)	(0.20)	(0.19)	(0.22)	(0.21)	(0.21)
$Q5-IV$ × High SI	0.87***	0.64**	0.53*	1.3***	0.83***	0.88***
	(0.28)	(0.29)	(0.30)	(0.31)	(0.31)	(0.32)
IV	20.3***	15.4***	15.7***	20.3***	15.4***	15.8***
	(4.8)	(5.2)	(5.1)	(4.8)	(5.3)	(5.1)
$Q2-IV$	0.13	0.06	-0.04	0.14*	0.07	-0.03
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
$Q3-IV$	0.07	0.01	-0.11	0.08	0.03	-0.09
	(0.10)	(0.11)	(0.10)	(0.11)	(0.11)	(0.10)
$Q4-IV$	-0.04	-0.09	-0.18	-0.05	-0.06	-0.16
	(0.14)	(0.14)	(0.14)	(0.14)	(0.15)	(0.14)
$Q5-IV$	-0.37*	-0.22	-0.30	-0.38*	-0.18	-0.31
	(0.21)	(0.23)	(0.24)	(0.21)	(0.23)	(0.23)
Observations	1,060,913	1,039,901	1,039,901	1,060,913	1,039,901	1,039,901
R^2	0.07487	0.06192	0.05647	0.07488	0.06194	0.05648
Date-Industry fixed effects	✓	✓	✓	✓	✓	✓
Firm fixed effects	✓	✓	✓	✓	✓	✓

Table 20: Explosiveness Up interacted with IV following SI dissemination.

The table presents the predictive power of the short interest ratio interacted with idiosyncratic volatility in 10-day subsequent returns following the dissemination of short interest (SI). The dependent variables are indicators of detecting an upward explosion over at least one of the five subsequent (overlapping) 10-day intervals following the dissemination of short interest. Columns (1-3) and (4-6) report the results for detection using raw price data (adjusted only for dividends and stock splits), CAPM-adjusted price data, and FF3-adjusted price data, respectively. The first three columns present results where the Short Interest Ratio (*SIR*) is used to select the top decile of heavily shorted stocks as the treatment group. The other three columns present results where the partialled-out *SIR* is used. The control variables and fixed effects are consistent with the controls used for return regressions reported in Table 17. The sample period spans from 2003 to 2022, and statistical significance is indicated at the 10%, 5%, and 1% levels by *, **, ***, respectively.

Dep.Var	$E^{\alpha,up}_{s,d+1} - E^{\alpha,up}_{s,d-L}$			
	G = 1 day	G = 2 days	G = 3 days	G = 5 days
Panel A: *Using* $L = 7$				
High SI	0.28***	0.32***	0.34***	0.33***
	(0.09)	(0.10)	(0.11)	(0.12)
Observations	1,037,241	1,037,045	1,036,804	1,033,987
R^2	0.04878	0.04941	0.04991	0.05040
Panel B: *Using* $L = 10$				
High SI	0.31***	0.39***	0.39***	0.44***
	(0.09)	(0.10)	(0.11)	(0.13)
Observations	1,037,223	1,037,027	1,036,786	1,033,969
R^2	0.04543	0.04733	0.04859	0.05003
Panel C: *Using* $L = 12$				
High SI	0.29***	0.36***	0.39***	0.51***
	(0.08)	(0.10)	(0.12)	(0.13)
Observations	1,036,846	1,036,650	1,036,409	1,033,592
R^2	0.04641	0.04896	0.05017	0.05091
Date-Industry fixed effects	✓	✓	✓	✓
Firm fixed effects	✓	✓	✓	✓

Table 21: Difference in Explosiveness Up around SI Dissemination

The table presents a fixed-effect regression analysis of the predictive power of the change in the short interest ratio on the probability of detecting an upward explosion following the dissemination of short interest (SI). The dependent variables are indicators of detecting an upward explosion over at least one of the G subsequent (overlapping) 10-day intervals following the SI dissemination. These indicators are adjusted based on the same measure lagged by L days before the dissemination. Explosiveness is detected at the 1%-significance level using $k = 1$ and *CAPM*-adjusted stock prices. The control variables and fixed effects are consistent with those used for the return regressions reported in Table 17, except for short-term reversal, which is lagged to just prior to date $d - L$. The standard errors are two-way clustered by Date and Firm. The sample period spans from 2003 to 2022, and statistical significance is indicated at the 10%, 5%, and 1% levels by *, **, ***, respectively.

Dep.Var	Explosiveness Down			Explosiveness Up		
	$E^{\alpha,dn}_{s,d+1} - E^{\alpha,dn}_{s,d-12}$			$E^{\alpha,up}_{s,d+1} - E^{\alpha,up}_{s,d-12}$		
	No adjustment	CAPM-Adj.	FF3-Adj. (%)	No adjustment	CAPM-Adj.	FF3-Adj. (%)
	(1)	(2)	(3)	(4)	(5)	(6)
High SI	0.52***	0.73***	0.66***	0.34***	0.39***	0.46***
	(0.13)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Turnover	6.2	3.1	2.1	4.0	3.1	3.2
	(4.6)	(3.6)	(3.4)	(3.5)	(3.8)	(3.3)
IV	-1.7	0.32	-5.1	-0.003	-2.8	3.4
	(6.2)	(5.8)	(5.9)	(5.0)	(5.7)	(5.2)
Am.Liq	-0.002	-0.002	0.002	-0.005	-0.003	-0.005
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Log-Size	-0.06	0.07	-0.02	-0.01	0.001	0.13
	(0.19)	(0.14)	(0.12)	(0.11)	(0.12)	(0.11)
BtM	0.04	-0.03	-0.09	-0.04	0.04	0.07
	(0.13)	(0.13)	(0.12)	(0.10)	(0.10)	(0.11)
Δ Assets	-0.03	0.002	0.03	0.06	0.08	0.05
	(0.08)	(0.07)	(0.07)	(0.05)	(0.05)	(0.05)
Op. Prof	-0.03	0.02	0.005	-0.06**	-0.03	-0.009
	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
IOR	0.32	0.32	0.18	-0.007	-0.26	-0.47***
	(0.26)	(0.25)	(0.25)	(0.21)	(0.19)	(0.18)
STReversal	0.53	1.5***	1.3***	0.20	-0.02	-0.05
	(0.58)	(0.48)	(0.44)	(0.40)	(0.40)	(0.37)
Momentum	0.12	-0.11	-0.07	0.22	0.22	0.08
	(0.20)	(0.17)	(0.16)	(0.14)	(0.15)	(0.13)
$Q2-\beta_{HML}$	0.10	0.18	0.22*	-0.10	-0.06	-0.06
	(0.13)	(0.13)	(0.12)	(0.09)	(0.10)	(0.09)
$Q3-\beta_{Mkt}$	0.23	0.25*	0.38**	0.04	0.04	-0.04
	(0.16)	(0.15)	(0.16)	(0.11)	(0.11)	(0.10)
$Q4-IV$	0.33	0.46***	0.47***	0.16	0.20	-0.03
	(0.22)	(0.18)	(0.16)	(0.14)	(0.15)	(0.14)
$Q5-IV$	0.32	0.40	0.50*	0.20	0.32	-0.02
	(0.33)	(0.28)	(0.27)	(0.23)	(0.25)	(0.24)
Standard-Errors		Date			Date & Firm	
Observations	1,057,387	1,036,409	1,036,409	1,057,387	1,036,409	1,036,409
R^2	0.10650	0.06103	0.05441	0.05792	0.05017	0.04433
Date-Industry fixed effects	✓	✓	✓	✓	✓	✓
Firm fixed effects	✓	✓	✓	✓	✓	✓

Table 22: Difference in Explosiveness around SI Dissemination

The table presents a fixed-effect regression analysis of the predictive power of the change in the short interest ratio on the change in the probability of detecting an explosion, either up (in the last three columns) or down (in the first three columns). This analysis covers at least one of the three subsequent overlapping 10-day intervals following the dissemination of short interest (SI). The dependent variable is the explosiveness indicator over $G = 3$ 10-day intervals, adjusted by the indicator of an explosion lagged 12 days before the dissemination. The control variables and fixed effects used in this analysis are consistent with the controls employed for the return regressions reported in Table 21. Only continuous or statistically significant control variables are reported.

	Others			High SIR Change		
	Before	After	Difference	Before	After	Difference
$\bar{r}_{t,t+l}$ (%)	-0.04	-0.03	0.01	-0.05	-0.33	-0.28
s.d. (%)	0.31	0.31	0.00	0.35	0.35	0.00
$Min(\%)$	-2.50	-2.50	-0.01	-2.75	-2.83	-0.08
$Q_{0.1}$ (%)	0.25	0.25	0.00	0.30	0.30	0.00
$Q_{0.9}$ (%)	-0.25	-0.25	0.00	-0.30	-0.30	0.00
$Max(\%)$	2.54	2.56	0.02	2.81	2.85	0.04
Skew	0.11	0.13	0.01	0.14	0.08	-0.05
Kurt	38.62	38.65	0.03	37.40	40.26	2.86
Explosiveness Up	2.53	2.59	0.07	2.58	2.88	0.31
Explosiveness Down	3.20	3.28	0.08	3.10	3.51	0.42

Table 23: Moments of HF returns around SI dissemination

The table presents the average change in high-frequency stock moments around the SI dissemination date. The first three columns represent the average change for the returns of the stocks in the bottom 9 deciles sorted by the partialled out SIR. The other three columns present results for the returns of the top decile by SIR. The first and the second of the three columns in each group report the average statistics measured over a 10-day interval prior to and after the dissemination of SI data. The last columns report the difference between the values. The first row reports the average log-return over the 10-day interval. The next 7 rows report the averages of standard deviation, minimum, maximum, 10th and 90th percentiles, skewness, and kurtosis estimated based on the intraday log-returns within the 10-day window. The last two rows report the average frequency of detection of explosiveness up and down within the interval estimted using the SADF procedure with a significance level of 1% and lag order $k = 1$. All variables but kurtosis and skewness are reported in percentage terms. The frequency of observations is 5 minutes per day, incorporating nine overnight returns from 16:00 to 9:40 of the next day, resulting in a total of 770 observations.

Dep.Var	Change in the variable							
	$\bar{r}_{t,t+l}$ (%)	s.d.	Max(%)	Min(%)	$Q_{0.9}$ (%)$	$Q_{0.1}$ (%)$	Skew (%)	Kurt (%)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High SI	-0.37***	0.002***	0.03**	-0.08***	-0.0009**	8.7×10^{-5}	-5.7***	281.8***
	(0.08)	(0.0007)	(0.01)	(0.01)	(0.0004)	(0.0003)	(2.2)	(35.5)
Observations	1,058,189	1,058,189	1,058,189	1,058,189	1,058,189	1,058,189	1,058,189	1,058,189
Squared Correlation	0.29843	0.22185	0.12219	0.11655	0.38780	0.40353	0.08207	0.07301
Date-Industry fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Firm fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Table 24: Change of Empirical Moments of HF returns around SI dissemination

The table presents a fixed-effect regression analysis that examines the predictive power of the change in the short interest ratio on the change in the empirical moments of high-frequency returns estimated over 10-day intervals following the dissemination of short interest (SI). The dependent variables represent the difference in the estimates of various return measures specified in the header of the table over a 10-day period before the dissemination date and a 10-day period after the dissemination, excluding the dissemination date itself. The specific return measures are detailed in Table 17. The control variables and fixed effects used are consistent with those employed in other analyses reported in Table 17, with the exception of short-term reversal, which is lagged to just prior to date $d - L$. Standard errors are two-way clustered by Date and Firm. The primary explanatory variable is the indicator for being in the top decile by the partialled out Short Interest Ratio (SIR). The sample period spans from 2003 to 2022, and statistical significance is indicated at the 10%, 5%, and 1% levels by *, **, ***, respectively.

		Gaussian			Bootstrap				Stoch. Vol.	Actual Data
Expl	RW	Const Cofs	AR	Overnight	No AR	No Drift	Drift	Overnight	Heston	Actual
Correctly Specified										
Down	0.46	0.46	0.72	0.54	0.98	0.98	1.35	1.41	1.22	3.26
Up	0.45	0.44	0.62	0.49	0.98	0.97	1.32	1.36	1.29	3.11
Using more lags										
Down	0.48	0.47	0.75	0.55	1.00	0.99	1.38	1.45	1.21	3.26
Up	0.50	0.49	0.64.	0.51	1.02	1.01	1.36	1.40	1.30	3.11
Misspecified by one lag										
Down	0.41	0.30	0.55	0.41	0.95	0.84	1.15	1.18	1.01	3.26
Up	0.42	0.27	0.48	0.38	0.92	0.82	1.11	1.15	1.07	3.11

Table 25: Alpha of the test based on simulations

This table presents the results of SADF test size across various simulations. The probability of detecting explosion up and down at 1% significance level is expressed in percentage points for each specification. The first four columns pertain to simulations with Gaussian standard errors, where error moments are calibrated to actual TAQ data. *RW* stands for the random walk model with zero drift and autocorrelation. *Const Cofs* attributes constant autocorrelation equal to the median re-estimated autocorrelation coefficients (ϕ-s) in the population. *AR* randomly draws coefficients from the empirical distribution of pre-estimated parameters. The *Overnight* specification incorporates scheduled (overnight) Gaussian jumps drawn based on the pre-estimated overnight average and standard deviation that define the distribution. The intraday residuals are drawn based on intraday observations only. The subsequent four columns pertain to simulations using error terms drawn from the actual distribution of errors. The *overnight* specification involves simulations that incorporate overnight jumps based on the overnight distribution of returns. The *stochastic volatility* specification involves simulations with errors following the Heston stochastic volatility model, which is pre-estimated on TAQ data using GMM. The final column represents the frequency of detecting explosiveness based on actual data. The sample size consists of 100,000 observations per year (from 2004 to 2022), totaling 1,900,000 observations in all. The simulations were performed on non-filtered sample for $k = 2$. The statistics for actual data are based on non filtered data as well.

Figures

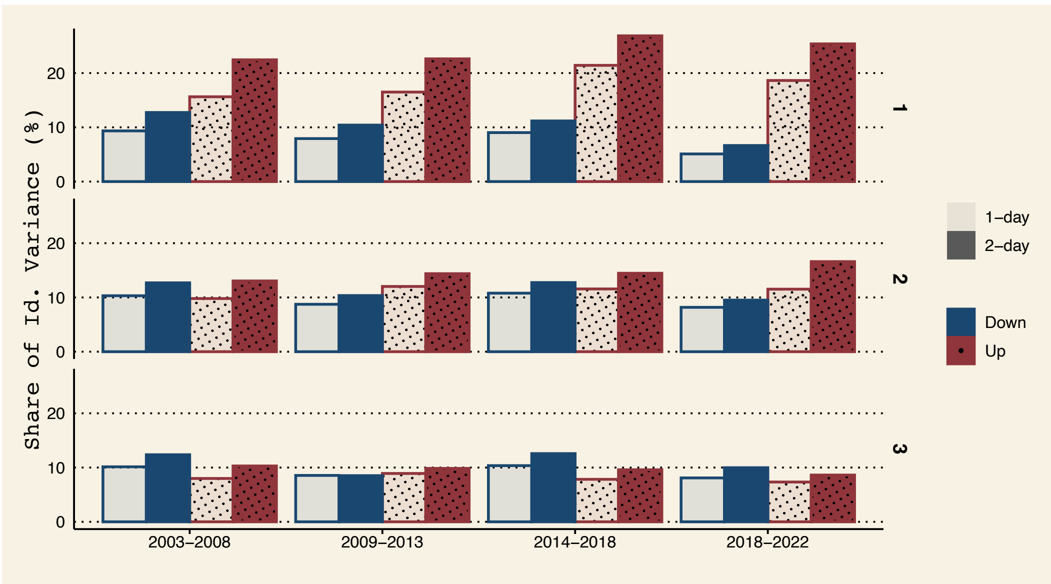

Figure 23: Idiosyncratic Variance Share by explosive events

The figure illustrates the proportion of idiosyncratic variance ascribed to the dates when explosions are detected. Explosions are identified using FF-3 adjusted prices, with a significance level of detection set at 2.5%, and lag order $k = 1$. The stocks are pre-classified into three equal groups based on day-specific size, measured 30 days prior. The shares for the smallest (largest) stocks are reported in Panel 1 (3). As small stocks are more susceptible to exclusion during cleaning procedures, the primary sample is predominantly composed of medium and large stocks. For further details on the sample, refer to Table 13.

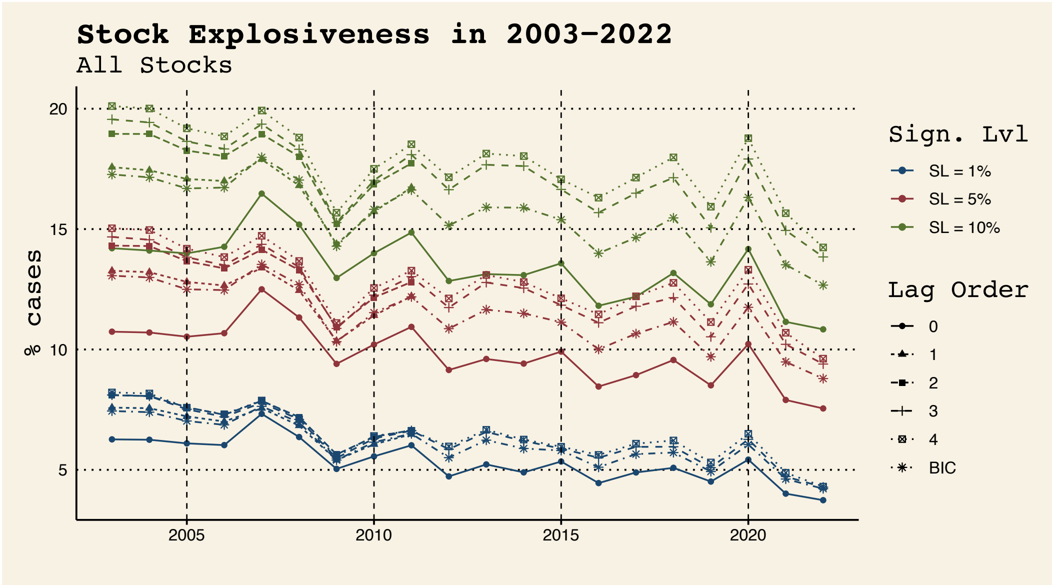

Figure 24: Rate of Detection

This figure illustrates the occurrence frequency of individual stock explosions during random 10-day intervals, organized by the year of observation. Various colors and linetypes represent distinct hyperparameters, including significance level and lag order, employed for the detection process. The BIC lag-order specifically reflects detection outcomes based on a model with k selected by Baysessina Information Criterion.

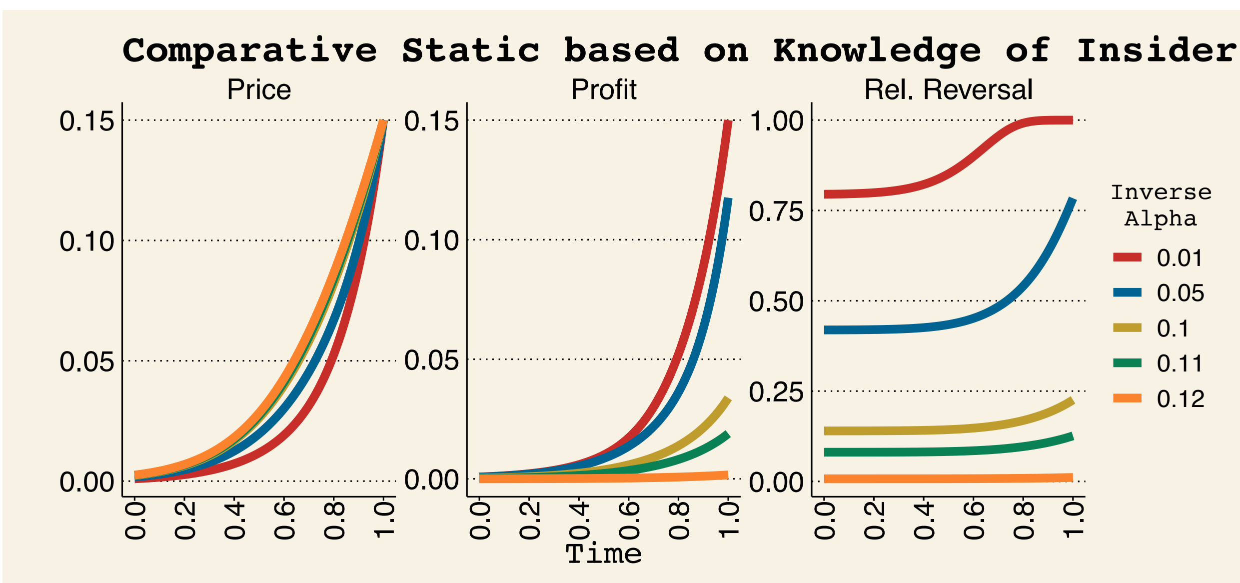

Figure 25: Price, $p(t)$, profit, $\pi(t)$, and reversal(t) depending on α^{-1}

This figure illustrates the sensitivity of price, profit, and relative reversal to the knowledge possessed by insiders, which is captured by α^{-1}. The other parameters replicate the suggested parameters from the numerical example in 8.3.