

Stahl, Philip; Blauth, Jérôme

Article — Published Version

Martingale defects in the volatility surface and bubble conditions in the underlying

Review of Derivatives Research

Provided in Cooperation with:

Springer Nature

Suggested Citation: Stahl, Philip; Blauth, Jérôme (2024) : Martingale defects in the volatility surface and bubble conditions in the underlying, Review of Derivatives Research, ISSN 1573-7144, Springer US, New York, NY, Vol. 27, Iss. 1, pp. 85-111, <https://doi.org/10.1007/s11147-023-09200-x>

This Version is available at:

<https://hdl.handle.net/10419/315413>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>



Martingale defects in the volatility surface and bubble conditions in the underlying

Philip Stahl^{1,2,3} · Jérôme Blauth^{2,3}

Accepted: 14 December 2023 / Published online: 19 January 2024
© The Author(s) 2024

Abstract

The martingale theory of bubbles enables testing for asset price bubbles by analyzing option prices. As recently shown by Piiroinen et al. (Asset price bubbles: an option-based indicator, 2018), the SABR model is a strict local martingale when its parameterization implies a positive correlation between stock and option prices. We operationalize this theoretical result and analyze stock price bubbles in 2576 stocks over 26 years. Martingale defect conditions are absorbed quickly by options markets, but identify high proportions in significant and permanent changes in distribution of price returns, option trading activity, short interest in the underlying, and institutional ownership. These results confirm many common assumptions about stock price bubbles. These bubbles are temporally clustered, and tend to occur in periods of positive market development. Martingale defects are rare in market corrections, which indicates that they are a result of overoptimistic speculation.

Keywords Stock price bubbles · Martingale defect · Strict local martingales · SABR model

Mathematics Subject Classification 91B28 · 91B70 · 91B84

1 Introduction

An asset price bubble occurs when the price exceeds the fundamental value of the asset for a sustained period. Unfortunately, testing for bubbles suffers from a joint hypothesis problem: As the fundamental value is, by its nature, unknown, both quantities have to be estimated simultaneously, which severely reduces the usefulness of

✉ Philip Stahl
philip@philipstahl.com

¹ Technische Universität Darmstadt, Darmstadt, Germany

² Deka Investment GmbH, Frankfurt am Main, Germany

³ IQAM Research GmbH, Frankfurt am Main, Germany

these types of analysis (Camerer, 1989). The martingale theory of bubbles as developed by Jarrow (1992), Loewenstein & Willard (2000), Cox & Hobson (2005), Heston et al. (2007), Jarrow et al. (2010), Biagini et al. (2014), and many others, characterizes asset price bubbles in terms of strict local martingales.

To overcome the joint hypothesis issue, Jarrow (2015) suggests specifying an option pricing model, which can be validated separately, and checking whether it implies an underlying price process that is a *true martingale*, and not a *strict local martingale*.¹ Piiroinen et al. (2018) derive a martingale defect indicator for SABR dynamics, fundamentally transforming the task of identifying an asset price bubble into a calibration problem. In this paper, we empirically investigate the connection between stock price bubbles and changes in several non-related variables in a large-scale study of 2576 stocks over approximately 26 years.

We adapt the approach suggested by Piiroinen et al. (2018) to identify martingale defects in the volatility surface of single stocks. We evaluate defect persistence, stock price return distribution, option trading activity, short interest, and institutional ownership on a per-event basis and find strong evidence that martingale defects coincide with permanent changes in these variables. Our findings confirm many fundamental assumptions about stock price bubbles. Furthermore, we examine the temporal clustering of martingale defect events over time and find that they predominantly occur during phases of positive market returns and rarely in phases of sustained negative returns.

Monitoring the volatility surface for martingale defects is a useful tool to detect overoptimistic speculation in stocks. It enables investors to identify individual stocks which may not be rationally priced and adjust their exposure accordingly. It is furthermore an appropriate tool for regulators to improve market monitoring and focus their resources in order to protect retail investors. Across the entire market, it provides a gauge of general investor optimism and a markets propensity to develop bubbles.

The remainder of this paper is structured as follows. To provide background and context to our study, we will begin with a short review of the related literature. Then we examine the details of the martingale theory of bubbles and the implementation by Piiroinen et al. (2018). Section 4 provides an overview of the data, specifics of our calibration procedure, and further implementation details. Our subsequent analysis is two-fold. First, we examine individual bubble events and evaluate changes after observing a martingale defect. Second, we aggregate bubble events across the entire market and investigate their clustering behaviour over time. Section 6 concludes the paper and identifies topics for further exploration.

2 Literature review

The martingale theory of bubbles is fundamentally based on the idea that an asset in bubble conditions is not a fair bet, as market participants are willing to overpay for a potential upside.

¹ Specifically, see section 3.2.3 of Jarrow (2015).

Participating in these bubbles can be rational under certain conditions, such as laid out by Jarrow (1992), who analyzes asset price bubbles as market manipulation by large traders and extends arbitrage pricing theory, which is based on an economy of price taking actors, by allowing for large traders. These purchase large positions in individual stocks, and may corner the market or squeeze holders of short positions to force them to pay an arbitrary price. He derives sufficient conditions under which these strategies are impossible under fair pricing measures. In this study, we find that martingale defects in option prices do indeed coincide with increased short interest.

We now review the literature concerning the martingale theory of bubbles, delving into its fundamental assumptions and constraints. In a seminal paper, Delbaen & Schachermayer (1994) generalize the fundamental theorem of asset pricing and introduce *No Free Lunch with Vanishing Risk* (NFLVR), thus paving the way for the martingale theory of bubbles. Jarrow et al. (2007) review the literature on asset price bubbles in complete markets with infinite trading horizons and conclude that, under NFLVR no arbitrage conditions, the existence of bubbles implies that markets must be incomplete. Cheridito et al. (2007) show that it is necessary and sufficient for stocks and bonds to be undominated trading opportunities for an equivalent local martingale measure to exist. Heston et al. (2007) provide conditions on the prices of options to rule out bubbles in the underlying. Jarrow et al. (2010) extend the NFLVR framework by imposing the *no dominance* (ND) conditions suggested by Merton (1973) and allow an infinite number of local martingale measures to coexist, which represent the fundamental economic regimes. For each trade, the market chooses one of these measures to determine the price. When market fundamentals change, a different measure is chosen. While asset prices remain unchanged, derivatives prices must change, which implies that stock price bubbles can be inferred from option prices.

Jarrow & Protter (2013) investigate perceived positive alpha under incomplete information, leading to the illusion of arbitrage opportunities in asset price bubbles. Biagini et al. (2020) extend Jarrow and Protter (2013) and study the relationship between private information and perceived bubbles formally, with the potential consequence that asset price bubbles are less likely for assets with less hidden information, such as indices, when compared to single stocks.

The utilization of martingale theory in empirical research proves valuable for examining financial bubbles. Jarrow & Protter (2010) examine the implications for derivatives pricing and detection of asset price bubbles. They suggest three different approaches to detect asset price bubbles. The first approach relies on modeling the fundamental value of assets directly, leading to the joint hypothesis issues described by Camerer (1989). The second approach specifies a stochastic process for the underlying which is calibrated against the time series of stock price returns. The third approach chooses an option pricing model, and calibrates it against observed option prices. The advantage of this approach is that by calibrating the model, it is automatically validated. In this paper, we follow the this suggestion by calibrating the SABR model to observed out-of-the-money option prices. Jarrow et al. (2011) employ martingale-based *volatility modelling* to detect bubbles. They estimate a non-parametric volatility function as proposed by Florens-Zmirou (1993) from past

asset prices. They illustrate their method based on four stocks during the dotcom-bubble and find that the bubble conditions they identify overlap with stocks and periods that were previously considered bubbles as well. Obayashi et al. (2016) utilize the first approach proposed by Jarrow et al. (2011) to analyze a large number of stocks and examine the lifetime of bubbles. They find that these stock price bubbles exhibit surprisingly long lifetimes, typically spanning several months or years. As irrational exuberance would typically be short-lived, this points towards the rationality of stock price bubbles. In contrast, we find that bubble conditions in the volatility surface are typically absorbed within a few days. In conjunction with our other findings, this lends some support to the notion that bubbles in option markets are predominantly propelled by irrational speculation.

By imposing boundary conditions on option prices at extreme strikes, Jarrow & Kwok (2021) identify bubble conditions in the S & P 500 Index and develop a profitable momentum trading strategy based on the concept of “riding the bubble to the top”, as suggested by Conlon (2004), indicating that martingale defects in option prices may have some value for forecasting price returns.

Piironen et al. (2018) develop the *martingale defect indicator* from the SABR model by analytically deriving an expression to describe the magnitude of the defect. They analyze the bubble risk of SNAP Inc., Twitter Inc., and Square Inc. in 2017 and 2018 and show that bubble conditions can be accurately detected by calibrating a SABR model to observed option prices. Our empirical investigation relies on this indicator due to its straightforward implementation, as well as the extensive research on the SABR model, which makes its calibration a well-documented process.

Fusari et al. (2022) exploit the put-call price differential to identify bubble conditions from option prices. By estimating a specifically developed *generalized stochastic volatility jump diffusion* or G-SVJD model, which admits both martingale and strict local martingale representations, to put and call options separately, they are able to show that bubbles tend to occur in call options, but not in put options. Furthermore, they find that bubbles tend to occur regularly in single stocks, but rarely in indexes.

A different approach is proposed by Biagini et al. (2022), who suggest training a neural network to recognize whether a smile of call option prices is generated by a strict local martingale or a true martingale. The advantage of this approach is its independence from model specifications, at the cost of computational complexity.

Asset price bubbles are examined not solely through the lens of martingale theory but also via alternative approaches. In our empirical investigation, we are able to substantiate some of these findings. Bakshi et al. (2021) analyze VIX futures curves and find that volatility is mean-reverting and VIX futures are in backwardation when disaster risk is elevated. This provides some empirical backing for the approach of Piironen et al. (2018), who identify martingale defects when the correlation between stock price and variance level turns positive.

The concept of *dark matter* in asset pricing models draws a connection to economic components that are difficult to measure directly and quantifies its impact on model stability, which was formalized by Chen et al. (2022). Using a semimartingale-based approach, Bakshi et al. (2022) show how to decompose equity option risk premiums and examine the dynamics of *jumps crossing the strike* and

local time. Intuitively, our approach identifies bubble conditions in the underlying when the correlation of price and volatility is positive, which suggests a possible link with the findings of Bakshi et al. (2022), who show that negative premia for upside equity risk are consistent with the presence of unspanned risks.

Blocher et al. (2021) find that short sellers are regularly forced to exit positions earlier than optimal. In our study, we find that bubble conditions induce increased short selling activity, supporting their short squeeze hypothesis. SEC (2021) examine market structure and trading activity around the GameStop Inc. bubble in January 2021 and find, contrary to common perception at that time, that short covering was not the primary driver of the stock price run-up. Instead, the primary driver for this specific bubble event were overly optimistic, young, inexperienced investors. Observing overoptimistic speculation in the volatility surface further corroborates these findings. Mohrschladt & Schneider (2021) link option prices with internet search interest and find that retail investors contribute to idiosyncratic volatility through irrational trades, which are exploited by sophisticated market participants in the options market.

3 Theoretical background

In this section we describe the market setting and adapt the approach by Piironen et al. (2019) to utilize the SABR model to detect bubble conditions for a large data set.

We fix a finite time horizon $T > 0$ and consider the filtered probability space $(\Omega, \mathcal{F}, (F_t)_{t \geq 0}, \mathbb{Q})$, where \mathcal{F} is the σ -field of measurable subsets of Ω and the filtration $(F_t)_{t \geq 0}$ satisfies the usual conditions (see Protter (2016)). \mathbb{Q} denotes an equivalent local martingale measure and, by possibly embedding $(\Omega, \mathcal{F}, (F_t)_{t \geq 0})$ into a larger complete market, we can assume that \mathbb{Q} is also unique. We consider a stock price process $(S_t)_{t \geq 0}$ on Ω with continuous paths \mathbb{Q} almost surely. Its forward price is then given by

$$F_t = S_t e^{(r-d)t},$$

for some constant risk free rate r and dividend yield d which might include borrow costs. We denote the expectation operator with respect to \mathbb{Q} by \mathbb{E} and, if we want to empathize on the current price $x = S_0 \geq 0$, we put x into the subscript \mathbb{E}_x .

We follow Piironen et al. (2019) by making the following

Definition 1 The stock price S_t is said to admit a bubble on $[0, T]$ with respect to \mathbb{Q} if the discounted process

$$S_t e^{-(r-d)t}$$

is a strict local martingale on $[0, T]$ with respect to \mathbb{Q} . The normalized martingale defect is defined as

$$d_x(T) := 1 - \frac{e^{(d-r)T}}{x} \mathbb{E}_x[S_T].$$

The intuition behind this definition is as follows: the asset is currently in the state of a bubble if the current market value of the asset S_0 exceeds its fundamental value which is the discounted expectation of its future value $e^{-rT} \mathbb{E}[S_T]$.

From the definition of a (local) martingale it is immediate that our martingale defect indicator satisfies $d_x(T) \geq 0$ with equality if and only if our model does not admit a bubble.

It shall be noted that d_x not only tells us whether or not an asset is currently in bubble condition but also, due to its domain $d_x \in [0, 1]$, quantifies the strength of the bubble: the closer d_x is to one, the stronger the indication of a bubble event.

We use the results in Piironen et al. (2019) to calculate an analytical expression for d_x from market data. More precisely, we use available option prices for a given asset, calibrate a SABR model to this data and use the resulting parameters to compute our bubble indicator. This reduces our problem of estimating d_x to the problem of calibrating a volatility smile and leads to a particularly simple representation of d_x in the model parameters. We start by giving a rough overview over the model.

The stochastic volatility model introduced by Hagan et al. (2002), commonly known as *stochastic alpha, beta, rho* or SABR model, is an extension of the CEV model and is determined by the SDEs

$$dF_t = \alpha_t F_t^\beta dZ_t^1 \quad (1)$$

$$d\alpha_t = \nu \alpha_t dZ_t^2, \quad (2)$$

where $\alpha_0 = \alpha$ and

$$\langle dZ_t^1 dZ_t^2 \rangle = \rho t. \quad (3)$$

The elasticity parameter $\beta \in [0, 1]$ controls the general behaviour of the model and is usually not calibrated but chosen in advance. Following (Piironen et al., 2019, Theorem 3.1), we set $\beta = 1$, which is the log-normal case. $\nu \geq 0$ is the volatility-of-volatility and $\rho \in [-1, 1]$ is the correlation between the driving Brownian motions Z^1 and Z^2 .

The advantage of the SABR model in this context is that it admits a strict local martingale representation, unlike many traditionally used option pricing models. The SABR model therefore implies that stock price bubbles can possibly exist. Fusari et al. (2022) examine this issue in further detail, and contribute a more sophisticated approach based on their G-SVJD model.

For notational clarity and without loss of generality, from now on, interest rates and dividend are omitted and assumed to be zero. Piironen et al. (2019) then show that under SABR dynamics the martingale defect takes the particularly simple form

$$A(\alpha, \nu, \rho) := \lim_{T \rightarrow \infty} d_x(T) = 1 - e^{\frac{-2\rho\alpha}{\nu}}. \quad (4)$$

This is Piiroinen et al., (2018, Equation 15).

Since both volatility α and volatility-of-volatility ν must be strictly positive, A can only become positive when the correlation between stock price and volatility ρ is positive as well. Intuitively, this means that options get more expensive as the stock experiences positive returns, hinting at optimistic risk-seeking preferences.

4 Data and calibration

Our analysis covers the constituents of the MSCI USA Investible Market Index, which attempts to cover all large-, mid- and small-cap US stocks.² We analyze daily close prices of all put and call options on all stocks within the universe within a set of restrictions outlined below. In total, 2576 stocks are considered, and our universe is updated monthly to match the index. Constituent information is provided by MSCI. The option data is provided by OptionMetrics, begins on January 1st, 1996, and ends on April 25th, 2022, spanning 6758 trading days. The options dataset consists of daily close prices for approximately 51.3 million option contracts on all analyzed stocks. Stocks for which no options have been traded are excluded from the analysis. The option price data is matched to price data of the underlying price and to linearly interpolated U.S. treasury rates. Dividends are assumed to be a constant yield, based on the last known dividend payment. American options are evaluated using the approach suggested by Cox et al. (1979). Average daily trading volume, outstanding short interest, and institutional holdings data have been provided by Bloomberg.

Options with a daily trading volume of less than 100 contracts are excluded for that day. We write $k_i = \log\left(\frac{K_i}{F_T}\right)$ for the log moneyness. Our calibration only considers out-of-the-money options, as they tend to be more liquid than in-the-money options. This entails using call implied volatilities for strikes below the forward price and put implied volatilities for strikes exceeding the forward price. As shown by Fusari et al. (2022), we will not be able to see a bubble solely based on the put options. However, it is worth highlighting that the martingale defect indicator, as it is fundamentally derived from the parameter ρ , effectively models the skew of the volatility smile. In the SABR model, this parameter characterizes the 'difference' between the call and put wings of the smile. In this regard, our approach is consistent with Fusari et al. (2022). We further restrict the calibration to options with absolute log-moneyness $|k| < 0.5$.

This leaves us with a number of pairs $(K_i, \sigma_i)_i$ of strikes and implied volatilities for the remaining call or put options. The remaining contracts are enumerated by running variable $i = 1, \dots, N$, where N is the total number of remaining contracts. If $N < 5$, the stock is excluded from the analysis for that day. On average, each calibration is fitted against 6.12 contracts. This number increases over time as more contracts are traded. Figure 4 in the appendix provides the average number of available contracts per calibration over time.

² MSCI (2017) documents the index' construction.

As the volatility process 2 is not mean reverting, the SABR model is better suited for short expirations (Gatheral, 2006, Ch. 7). Our analysis is therefore two-fold. We calibrate the model against the full option price surface and refer to these results as *full surface*. We repeat the analysis, but restrict the calibration of the model to options with 1 month to expiration. For the *one-month-tenor*, contracts are selected to have a remaining lifetime between 25 and 35 days. We do not use shorter expirations, because the relative volume has shifted within the time frame of our analysis. Options with short maturities became more popular within the time frame of our analysis.³

We calibrate the model to the remaining option mid prices, which we convert to implied volatilities. Our minimization problem is

$$\underset{\alpha, \nu, \rho}{\text{minimize}} \sum_{i=1}^n (\sigma_{\alpha, \nu, \rho}(k_i) - \sigma_i)^2 \quad (5)$$

$$\begin{aligned} \text{s.t.} \\ \alpha \geq 0 \end{aligned} \quad (6)$$

$$\nu > 0 \quad (7)$$

$$\rho \in (-1, 1) \quad (8)$$

where $\sigma_{\alpha, \nu, \rho}(\cdot)$ is the SABR implied volatility function for given parameters α, ν, ρ and fixed time to maturity T .

We implement the suggestion of Le Floc'h & Kennedy (2014), who present an explicit initial guess procedure to generate an initial parameter guess for the minimization problem. This initial guess is the starting point for a Nelder & Mead (1965)-minimization of equation 5. We repeat this procedure for both the 1-month tenor as well as the full surface calibration daily for every stock under consideration. In order to avoid redundancy, we occasionally present only results for the full surface in the text body and move one-month tenor to Appendix E, where appropriate.

To validate the model calibration, we calculate the at-the-money SABR-model implied volatility of the root mean squared error on a 30 day horizon $RMSE_{IV}$. In total, we have 1.4 million calibrations with a mean $RMSE_{IV}$ of 0.97%. Of these, 56722 calibrations (4.05% of all calibrations) show a martingale defect $A(\cdot) \geq 0.05$, with a mean $RMSE_{IV}$ of 1.99%. Appendix A provides full results.

From these calibrated parameters α, ν and ρ , the martingale defect $A(\cdot)$ follows from Eq. 4. Due to the large number of stocks in our study and the computational effort required for Markov Chain Monte Carlo methods, we forgo obtaining a distribution for the martingale defect as suggested by Piiroinen et al. (2019).

For our analysis across a large number of stocks we must convert the martingale defect indicator $A(\alpha, \nu, \rho)$ of Eq. 5 into a binary signal $I(A(\cdot)) \in \{0, 1\}$. A positive signal marks an event in our study, which is specific to a single stock and date,

³ Appendix C shows that the relative proportion of contracts shifts to the short end from 2012 onward.

at which the option price surface for this stock indicates bubble conditions for its underlying.

By design, the defect indicator will be greater when the volatility surface deformation is more pronounced, and very small values indicate a low bubble intensity (Piironen et al., 2018). To reduce the noise created by the faintest bubble conditions, we impose a minimum threshold A_{min} on the indicator.⁴ Since the martingale defect indicator is based on the correlation ρ , which is normalized and thus time independent, we pick a static threshold A_{min} .

Asset price bubbles are in part defined by their persistence. We utilize this property to fortify our analysis against calibration issues which may not have been dealt with by the calibration procedure described in section 4. By requiring $A(\cdot)$ to remain above a threshold for a number of consecutive days p , the total number of events is reduced, but the accuracy of the signal can be improved. Section 5.1 examines defect persistence in more detail. Days where we cannot find a sufficiently accurate calibration are excluded, and break persistence. Our indicator is thus defined as

$$I_t = \begin{cases} 1, & \text{if } A_{t-n} \geq A_{min} \forall n \in [0, \dots, p] \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

To prevent counting the same fundamental event multiple times, event periods must be non-overlapping. Where SABR model is calibrated against the full surface, the indicator is denoted I^{full} , and where it is calibrated against the 1-month tenor of the volatility surface, it is denoted I^{1m} . Figure 1 provides the numbers of events for various thresholds and persistence requirements. Increasing these two requirements reduces the total number of events considered as well as the number of affected companies. We will further examine this in the next section. Calibrating the SABR model against the full surface generates a larger number of bubble events because the number of datapoints for each calibration is larger, and calibration errors are less likely. Consequently, persistent events are less likely to be interrupted by miscalibrations. Since calibrating the SABR model on the full volatility surface tends to smooth out anomalies in the low-DTE part, we expect to see more false negatives (*Type II errors*) compared to the 1-month tenor.

Most events do not persist over multiple days, which will be examined in detail in the next section. While we observe 18526 events with a persistence of 1 day above a threshold of $A_{min} = 0.01$ based on the 1-month tenor, this count reduces to 3505 after 2 days, and to 1339 after 3 days. Based on the full surface, we observe 38798 events with a persistence of 1 day, which reduces to 6731 after two, and 2477 after 3 days. The effect of A_{min} is similar. For $A_{min} = 0.05$, only 9983 events remain for the 1-month tenor, while 21087 events remain on the full surface calibration. The effect is even stronger for $A_{min} = 0.1$, with 6659 and 14602 events remaining before applying a persistence requirement. A threshold of $A_{min} = 0.05$ with a minimum persistence of 2 trading days appears to strike a balance between sensitivity and noise, and are chosen for the remaining analysis. With these requirements, 1518 events

⁴ Appendix A provides summary statistics of all events conditional on the martingale defect level.

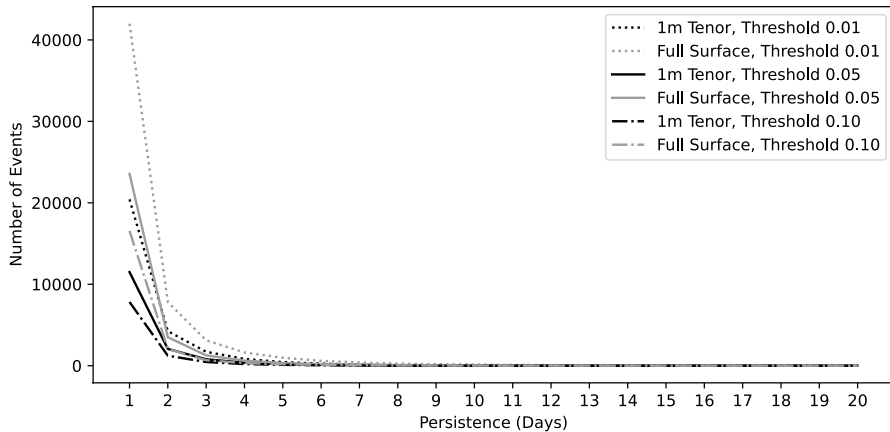


Fig. 1 Relative frequencies of persistence of martingale defect events in trading days. Martingale defects based on SABR calibration of the 1-month tenor as well as the full volatility surface with thresholds $A_{min} = \{0.01, 0.05, 0.1\}$

remain for the 1-month tenor, and 2745 events remain when calibrating against the full surface.

5 Analysis

In order to investigate the relationship between martingale defects derived from the volatility surface and bubble conditions in the underlying, we assess return distribution and trading activity before and after detecting a martingale defect event. Furthermore, we examine the relative occurrence of martingale defects over time.

For each event, we consider data during a certain number of trading days before and after the event, which we refer to as *event period* t_{event} . To assess whether an effect is persistent, we provide results for multiple event periods

$$t_{event} \in \{21, 63, 84, 105, 126, 189, 252\}.$$

Since bubbles are hard to quantify, we analyze a variety of metrics to establish a tight-knit connection between martingale defect and bubble conditions in the underlying. Since it is the foundation of our analysis, we begin with the reaction of the option market to a martingale defect event. Assessing the persistence of martingale defects in the volatility surface gauges the ability of the options market to arbitrage irregularities away. Next, we analyze whether the distribution of returns changes after observing a defect. Using the number of actively traded option contracts, we find that the martingale defect coincides with increasing option trading activity for this underlying. We also find that outstanding short interest increases. By examining institutional ownership, we find that institutional traders tend to reduce their exposure to stocks in suspected bubble conditions, leaving the participation to retail investors.

These four characteristics are naturally time-varying, and might exhibit time dependent variation that are unrelated to stock price bubbles. To avoid confounding the bubble-induced change with natural variation over time, we study a set of placebo-events and compare the results to identified bubble events. We generate these placebo events on a per-stock basis by shuffling the dates at which martingale defect events have been identified. This way, the number of events per stock remains identical, but the timing of the events is randomized. The events are restricted to the time period where each stock is a constituent of the IMI Index. We repeat the analysis on this set of placebo events and report the results for comparison.

Finally, we investigate bubble events over time across the entire market and find that martingale defects tend to occur more often in good times, and rarely in bad times.

5.1 Defect persistence

The martingale defect fundamentally indicates an irrational deformation of the volatility surface.

Figure 1 shows the frequency of event persistence for all incidences with $A_{min} \in \{0.01, 0.05, 0.1\}$. The majority of events occurs for only 1 day, and frequency drops quickly. The longest event in the 1-month tenor analysis was Sundial Growers Inc., where a defect was indicated for 14 consecutive days in May 2021. In the full surface analysis, Myriad Genetics Inc. indicated a defect for 20 consecutive days in December 2007. PubMatic Inc. also indicated a 20 day defect in June 2021. All three companies are well-known as *meme stocks* on various retail trading investing websites.⁵

The short persistence of the martingale defect hints at the efficiency of options markets to absorb bubble conditions in the underlying into a rational shape of the volatility surface. It does not imply that irrational exuberance in the underlying is not persistent, but rather that options markets can accommodate and return to efficiency quickly. Obayashi et al. (2016) analyze the lifetime of financial bubbles by modelling the distribution of the underlying directly, and find that bubble conditions in the underlying persist on the scale of years.

Appendix D presents more details on the effect of threshold and persistence on the number of observed events, as well as the number of affected companies.

5.2 Change of distribution

To confirm the connection between martingale defect events in the volatility surface and the price process of the underlying, we analyze whether the distribution of historical log returns changes with an event.

To assess whether the martingale defect indicator can reliably identify a change in the distribution of log-returns of the underlying, we employ the two-sample Kolmogorov–Smirnov ($K-S$)-test. The Null-hypothesis is that the distribution of

⁵ For example, see Sun (2021), Smith (2007), Farooque (2022).

log-returns during a period of given length before and after an event is identical. The interpretation of this would be that observing a martingale defect in the volatility surface does not coincide with a changing log-return distribution of the underlying, and is purely an anomaly in the volatility surface.

More specifically, for given underlying and length of event period $t_{event} > 0$ and for each date t_0 where we observe a martingale defect signal, we perform a K–S-test between the two sets $\{r_{t_0-1}, \dots, r_{t_0-t_{event}}\}$ and $\{r_{t_0+1}, \dots, r_{t_0+t_{event}}\}$, where we denote by

$$r_i = \log\left(\frac{S_{t_i}}{S_{t_i-1}}\right) \quad (10)$$

the return on day t_i .

For each event, we compute the K–S test statistic and p-value. Finally, we calculate the 1%-quantile and 5%-quantile of all p -values, which we report for a range of event periods in Table 1. These quantiles represent the proportion of events where the Null hypothesis that distribution of log-returns before and after an event is equal can be rejected with a confidence level of 1% and 5% respectively. By comparing different time frames, we gain insight into the persistence of changes. It should be emphasized that, regardless of the event period length, the martingale defect events are identical. The number of events changes only where event periods overlap or extend beyond the available data.

We observe that, for both the 1%-quantile and 5%-quantile, a longer event period increases the proportion of significantly different return distributions. For an event period of 252 trading days, almost 60% of events reject the null hypothesis of an unchanged log-return distribution with a confidence of 95%. Even though the options market absorbs the martingale defect within a few days, these findings imply that it hints at a permanent change of the underlying's price process for a large

Table 1 Aggregate results of two-sample Kolmogorov–Smirnov tests to compare empirical daily return distributions of the underlying before and after an event

Period t_{event}	One-Month Tenor					Full Surface				
	N	$p \leq 1\%$		$p \leq 5\%$		N	$p \leq 1\%$		$p \leq 5\%$	
63	867	12.31%	(1.29%)	22.99%	(3.28%)	1402	11.40%	(0.94%)	21.65%	(2.18%)
84	860	17.25%	(1.86%)	28.80%	(5.71%)	1393	15.24%	(2.50%)	26.44%	(6.82%)
105	836	21.09%	(2.31%)	33.81%	(8.29%)	1371	19.97%	(2.44%)	31.94%	(10.26%)
126	820	23.93%	(4.29%)	37.00%	(14.54%)	1345	23.56%	(3.50%)	36.61%	(11.71%)
189	793	36.69%	(4.07%)	50.55%	(15.03%)	1310	35.14%	(3.81%)	48.58%	(13.36%)
252	739	47.07%	(10.57%)	59.66%	(22.05%)	1245	47.71%	(8.47%)	59.20%	(20.06%)

Results on the left side are based on the calibration of the SABR model against the 1-month tenor of the volatility surface, while results on the right side are based on the full surface. Event Period t_{event} provides the number of trading days before and after an event under consideration. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p -values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses

proportion of events. A longer event period increases the number of datapoints and reduces uncertainty, therefore longer event periods are inherently more reliable. It is therefore not clear whether the effect is immediate or takes some time to manifest.

In our placebo study, we shuffle the identified event dates for each stock in order to randomize the timing component. The randomized sample shows a highly reduced proportion of significant changes in the return distribution, indicating that this distribution is indeed changing over time but to a generally lesser degree than when only accounting for bubble events.

5.3 Option trading activity

In order to examine the relationship between a martingale defect in the volatility surface and option trading activity, we analyze whether the number of contracts available remains constant before and after an event.

For each day, we calculate the number of traded options contracts for the underlying, while applying the liquidity requirements laid out in Sect. 4. For a pre-specified event window length, we calculate the average number of daily available contracts before and after each event. We calculate results counting contracts on either the full surface or restrict ourselves to the 1-month tenor. This is independent of the tenor selection for the calibration. Over the time period of our sample, the number of daily actively traded options has grown considerably. Appendix B provides a short overview of option trading activity over time. Therefore, the number of active contracts after an event should be expected to be larger than before, just due to the length of the event period. To compensate for this positive trend in the number of available options we adjust the observations downwards by calculating the total growth in options contracts during estimation and event period and distributing it equally across stocks. This implies (falsely) that option trading activity grows at a constant rate during the event period, and is equally distributed amongst securities. The effect of this adjustment is on average 0.0014 contracts for an event period of 10 days, and 0.0354 contracts for an event period of 1 year, and is negligible. The remaining difference can therefore be attributed to the martingale defect event. The number of contracts is consistently higher after an event. We evaluate this effect using a paired *t*-test, i.e. we compare the number of options contracts before and after an event on a per-event basis. Results are reported in Table 2. The resulting *p*-values are consistently very small across event periods, indicating that the increased number of traded contracts is significant. The mean difference grows with longer event windows and does not rebound after a while. This implies that the cause for the increase is not a sudden price movement, but rather a persistent effect created by increased options trading activity for the affected underlying. The observed effect is much stronger when considering all available options, not only those within the 1-month tenor, implying that the market prefers either shorter or longer options for bubble speculation. As the number of daily available contracts has grown continuously over time, the placebo study shows a similar pattern over time, but for a much smaller proportion of events.

5.4 Short interest ratio

Jarrow (1992) develops the martingale defect theory of bubbles to investigate price manipulation by large traders. One of his examples involves cornering the market and squeezing the holders of short positions to pay any price arbitrarily chosen by the large trader.⁶ He provides two reasons why this might happen. First, short traders are unable to observe the large traders purchases, thus not realize that the market is cornered. Second, that the cornering is technical in nature, i.e. that the short position exceeds the floating supply of shares, and the large traders position exceeds the float.

To investigate this hypothesis, we analyze the short interest in a stock relative to its available supply before and after observing a bubble event. As a proxy for freely available shares we use *Average Daily Trading Volume (ADTV)* as reported by the exchange. We divide outstanding short interest by ADTV, which is commonly referred to as *days-to-cover ratio*, as it measures the number of days would take short sellers buying the entire trading volume to cover their short positions.

For each stock, we retrieve daily outstanding short interest and ADTV, and calculate the days-to-cover ratio. Stocks where either is missing are removed from the sample. Individual missing values are filled forward. For each event, we calculate *t*-statistics of the days-to-cover ratio before and after the event over a certain number of trading days in the same fashion as above. We repeat this analysis for multiple event periods from 21 to 252 days. The results of the *t*-tests are aggregated.

Table 3 provides results based on the events generated by the I^{lm} . The mean *t*-statistic is positive for all event periods, with the largest value being 2.185 for the shortest event period, and ranging between 0.647 and 0.998 for event periods longer than 3 months. This implies that the short interest rises after a martingale defect is observed in the volatility smile. It rises, on average, by roughly one standard deviation of the days-to-cover ratio. For an event period of 21 days, 75.2% of events have a *p*-value of less than 5%, and 66.7% have a *p*-value of less than 1%. With longer event periods, and more daily observations per event, the proportion of significant test results rises. The longest event period shows the largest proportion of highly significant results, with 88.0% of events have a *p*-value of less than 5%, and 83.7% have a *p*-value of less than 1%. Appendix E provides similar results based on I^{lm} in Table 9.

After a martingale defect in the volatility surface, the market appears to be increasing short positions. For a very large proportion of events, this is highly significant. The effect is largest in the short term and levels off after a few months, but remains elevated for at least a year. This seems to confirm that the martingale defect can reveal bubble conditions in the underlying, and that some traders have similar perceptions. Similar to the previous results, the placebo study shows a similar pattern of growth with longer event periods, but with significantly reduced proportions. While short interest seems to be on an upward trend long event periods, martingale defects provide a clear indication of distinct increases in short interest.

⁶ Specifically, example 2 in Jarrow (1992).

Table 2 Aggregate results for t -tests of the number of actively traded option contracts before and after a martingale defect event, based on the calibration of the SABR model against the 1-month tenor of the volatility surface

Period t_{event}	Full Surface			One-month Tenor			
	N	t -statistics	p -values	N	t -statistics	p -values	$p \leq 1\%$
		Mean	$p \leq 5\%$		Mean	$p \leq 5\%$	$p \leq 1\%$
21	780	0.16	27.77% (8.08%)	670	0.36	13.75% (3.58%)	6.36% (1.15%)
63	800	0.47	41.52% (15.34%)	724	0.33	26.07% (6.79%)	16.03% (2.69%)
84	797	0.57	43.26% (25.34%)	702	0.42	25.00% (14.16%)	13.84% (8.32%)
105	775	0.82	46.47% (25.27%)	692	0.54	28.14% (12.36%)	17.84% (8.29%)
126	755	0.94	49.15% (26.18%)	673	0.62	33.17% (17.17%)	21.83% (9.97%)
189	735	1.39	53.41% (30.48%)	660	0.90	36.99% (16.57%)	28.16% (12.08%)
252	677	2.03	55.21% (32.16%)	611	1.12	41.81% (20.06%)	31.94% (14.40%)

Event Period t_{event} provides the number of trading days before and after an event under consideration. Results under *Full Surface* count all available options, while results under *One-month Tenor* count only contracts within the 1-month tenor. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. Mean t -statistic reports the mean of the t -statistic of all events. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p -values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses

Table 3 Aggregate results for t -tests of outstanding short interest divided by ADTV before and after a martingale defect event, based on the calibration of the SABR model against the 1-month tenor of the volatility surface

Period t_{event}	N	t -statistics	p -values			
			Mean	$p \leq 5\%$	$p \leq 1\%$	
21	683	1.454	75.3%	(51.7%)	66.3%	(47.0%)
63	670	1.331	79.6%	(59.4%)	74.0%	(61.0%)
84	661	1.069	80.8%	(57.9%)	75.2%	(55.7%)
105	635	0.614	81.6%	(70.6%)	76.9%	(62.1%)
126	623	0.260	82.7%	(74.8%)	78.2%	(63.9%)
189	594	0.336	85.2%	(70.9%)	80.1%	(67.1%)
252	538	0.528	86.6%	(78.2%)	82.1%	(74.2%)

Event Period t_{event} provides the number of trading days before and after an event under consideration. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. Mean t -statistic reports the mean of the t -statistic of all events. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p -values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses

5.5 Institutional ownership

In recent years, retail traders have shown increased interest in option markets (Deshpande et al., 2020). We investigate whether martingale defects spike the interest of retail investors by analyzing the percentage of shares held by institutions before and after an event.

Our data is matched to institutional holdings data provided by Bloomberg, which includes the holdings of institutions of type 13F, US and International Mutual Funds, US Insurance Companies, and aggregate institutional stake holdings.⁷ This data is available on a weekly basis since March 2010. Stocks where either is missing are removed from the sample. Individual missing values are filled forward. For each event, we calculate t -statistics of institutional ownership before and after the event over a certain number of trading days. We repeat this analysis for multiple event periods from 21 to 252 days. The results of the t -tests are aggregated.

Table 10 provides results based on the events generated by the I^{full} . The mean t -statistic is negative for all event periods, however, no clear trend is apparent. It ranges from -0.102 to -0.479 . In a similar pattern as before, with longer event periods, and more daily observations per event, the proportion of significant test results rises. The longest event period shows the largest proportion of highly significant results, with 64.7% of events have a p -value of less than 5%, and 63.1% have a p -value of less than 1%. Appendix E provides similar results based on I^{lm} in Table 4.

We find strong evidence that after a martingale defect in the volatility surface the percentage of institutional ownership tends to be lower than before. Our analysis does not reveal whether this is due to institutional investors reducing their position

⁷ For further information, see Bloomberg Terminal FLDS DS211.

in response to bubble conditions, or due to increased retail trader demand, and is limited to ownership of the stock, not the options. In addition, the placebo study reveals an overarching trend of slowly increasing non-institutional market participation. Proportions of significant changes are lower throughout, with a larger gap for shorter event periods, indicating that martingale defects mark distinct short-term changes in institutional ownership even under long-term trends.

5.6 Bubbles as market-wide phenomenon

We will now analyze whether the occurrence of martingale defects between stocks are related to each other by examining temporal clustering of events. We show that bubbles happen predominantly in good times, and occurrence of bubbles falls sharply when the overall market corrects.

For our entire stock universe, we count daily bubble events. During the time span of our analysis, option trading activity has grown.⁸ The number of observed events grows in line with this trend. This does not indicate that more assets are in bubble conditions, rather that these conditions are now tradeable in the options market. Figure 2 reports the total count of daily events as a fraction of daily total liquid contracts with the intention of compensating this development for our analysis. For the 1-month tenor, we use the total number of traded contracts within that tenor as denominator. Since the number of active contracts fluctuates daily, we use 252 trading day rolling averages for normalization. As a proxy for overall market performance, we use the S &P 500 Index.

Relative occurrence of bubbles based on both the full surface and the 1-month tenor follow a very similar pattern. Over time, the relative number of bubble events fluctuates between zero and 0.25% of daily liquid options. From 1997 to 2000, the SPX Index had positive performance, and the relative occurrence of martingale defects was highest in the entire sample. In the recession of 2001, the relative number of martingale defects was lower and less clustered, showing waning optimism. After a short rebound, the number of bubble events increases and is more clustered than during the drawdown. A further market correction until June 2003 disappoints the optimism, and a relatively bubble-free period commences until January 2004. As market performance is positive, the relative number of bubbles increases steadily, reaching a peak at the market top in 2007. During the recession, the number of bubble events falls quickly, and there are no bubble events between November 2008 and February 2009. From 2009 to 2012, SPX Index performance is positive, and the number of martingale defects grows. As a reaction to the 2012 correction, the number of events collapses, but continues to be elevated until 2016, after which it slightly levels off. At the beginning of 2020, the number of bubble events reaches 0.5%, which it has not reached since 2014. During the Covid-19 market correction, the number of events drops off sharply but rises quickly afterwards and remains elevated throughout the pandemic.

Overall, we observe that long periods of positive returns lead to a rise in martingale defect events in single stock volatility surfaces. By construction, our indicator

⁸ Appendix B examines the number of daily liquid option contracts over time.

Table 4 Aggregate results for t -tests of the percentage of institutional ownership before and after a martingale defect event, based on the calibration of the SABR model against the 1-month tenor of the volatility surface

Period t_{event}	N	t -statistics	p -values			
		Mean	$p \leq 5\%$		$p \leq 1\%$	
21	880	-0.585	62.0%	(43.1%)	54.5%	(34.2%)
63	867	-1.031	68.1%	(49.1%)	64.8%	(43.8%)
84	860	-0.193	67.4%	(48.6%)	65.3%	(49.4%)
105	836	-0.178	67.3%	(57.4%)	65.4%	(54.0%)
126	820	-0.671	67.4%	(58.5%)	65.9%	(56.4%)
189	793	-0.326	66.5%	(57.4%)	64.5%	(53.5%)
252	739	-0.088	66.9%	(59.6%)	65.3%	(56.8%)

Event Period t_{event} provides the number of trading days before and after an event under consideration. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. Mean t -statistic reports the mean of the t -statistic of all events. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p -values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses

is elevated when—in the parameters of the SABR model—stock price and volatility become positively correlated. Our observation implies that single stock options are used by the market to express strong optimism, and that option prices are higher than rational. These results support the hypothesis that asset price bubbles are an expression of overly optimistic expectations.

From 2016 on, the number of actively traded contracts begins to rise at a larger rate than before. The number of martingale defect events in relation to the total number of contracts is only very slightly lower. This implies that the multitude of single stocks with active option markets now absorb the gambling needs of the market. The bottom panel of Fig. 2 compares relative occurrence of bubble events of the 1-month tenor to those of the entire surface. A negative value means that bubble events are more prevalent in the one-month tenor, while a positive value means the opposite. While monthly contracts appear to be preferred for trading bubble events for the majority of our sample, trading activity shifted with the introduction of weekly and bi-weekly options. The expansion of options markets into shorter expirations and smaller companies does not appear to have created more bubble events, but redistributed them among a larger number of smaller securities.

6 Conclusion

This paper provides a large scale study of martingale defects in the volatility surface and changes in stock price dynamics and trading reaction. We find that martingale defects tend to coincide with other bubble characteristics.

We operationalize the detection of stock price bubbles by calibrating a SABR model to observed option prices, simplifying the approach suggested by Piironen

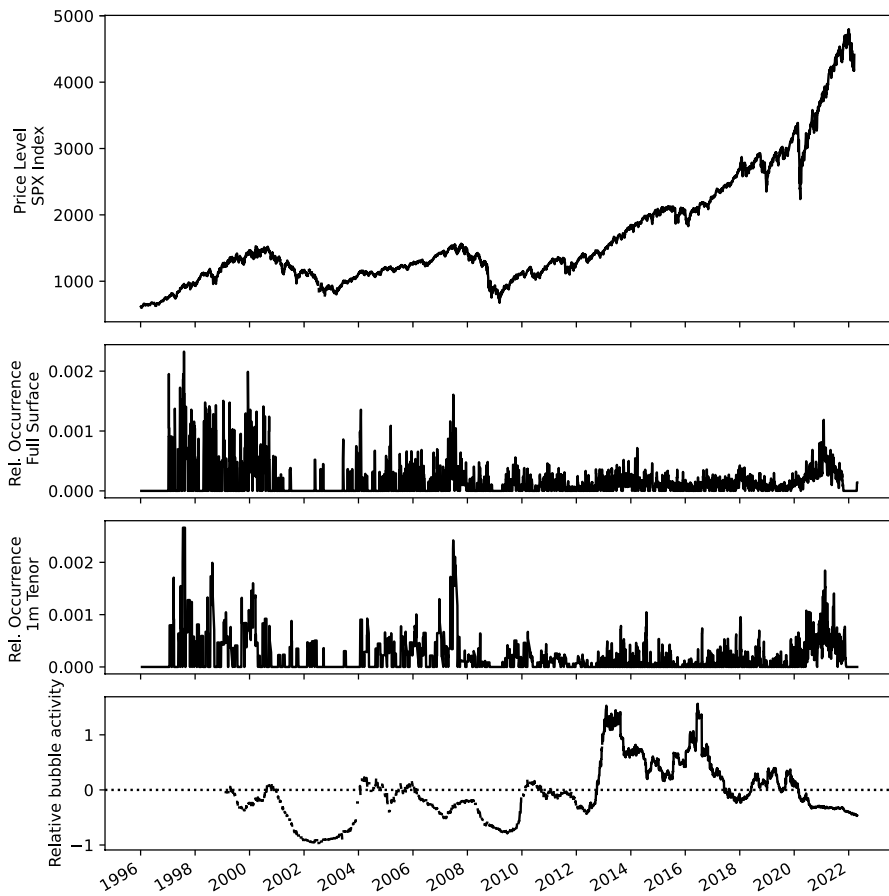


Fig. 2 Number of daily bubble events divided by the number of daily actively traded options contracts for the entire stock universe. Top panel shows the performance of the S & P 500 Index. Second panel shows the relative occurrence of bubble events based on the full surface. Third panel shows the relative occurrence of bubbles based on the 1-month tenor. Bottom panel compares relative occurrences from panel two and three, smoothed by 252 trading days

et al. (2018). Using a large stock and option price dataset, we calculate a daily bubble indicator for the constituents of the MSCI IMI Index.

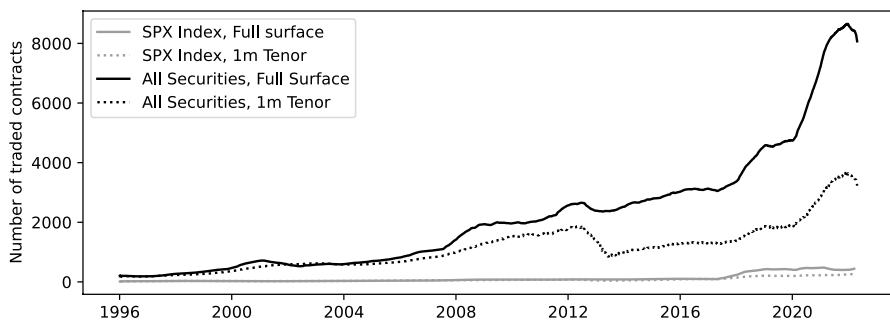
The volatility surface admits bubble conditions regularly, but rarely for longer than 3 days, which implies that the options market is generally efficient absorbing bubble conditions.

For each identified martingale defect event, we analyze changes of the return distribution, option trading activity, short interest, and institutional ownership. For all four factors, we find that a large proportion changes significantly after bubble conditions are observed. These effects appear to be permanent. By comparing our results with those obtained from a set of event dates that have been randomly shuffled, we affirm the reliability of martingale defects in identifying bubble conditions. This

Table 5 Descriptive statistics of calibration results.

Count	Total	Full sample	$A(\cdot) > 0$	$A(\cdot) \geq 0.01$	$A(\cdot) \geq 0.05$	$A(\cdot) \geq 0.1$
		1,402,230	159,309	109,414	56,722	39,270
	% of total		11.36%	7.80%	4.05%	2.80%
$RSME_{IV}$	Mean	0.0097	0.0192	0.0182	0.0199	0.0219
	Min	0.0000	0.0000	0.0000	0.0000	0.0000
	25% quantile	0.0000	0.0000	0.0000	0.0000	0.0017
	50% quantile	0.0027	0.0080	0.0072	0.0094	0.0116
	75% quantile	0.0102	0.0255	0.0238	0.0259	0.0280
	Max	0.2121	0.2121	0.2121	0.2121	0.2121
Defect $A(\cdot)$	Mean	0.0172	0.1511	0.2181	0.3978	0.5432
	Min	0.0000	0.0000	0.0100	0.0500	0.1000
	25% quantile	0.0000	0.0071	0.0228	0.0857	0.1743
	50% quantile	0.0000	0.0245	0.0535	0.1886	0.4281
	75% quantile	0.0000	0.0973	0.2038	0.8419	1.0000
	Max	1.0000	1.0000	1.0000	1.0000	1.0000

The first two rows provide the number of valid calibrations in the sample, as well as their proportions. The second set of rows provides statistics of calibration $RSME_{IV}$. The third set of rows provides statistics of the resulting martingale defect indicator $A(\cdot)$. The columns provide statistic conditional on whether $A(\cdot)$ exceeds a given threshold [0.0, 0.01, 0.05, 0.1]

**Fig. 3** Daily active options contracts for the SPX Index as well as our entire universe

shows that our methodology successfully avoids any potential confounding of stock price bubbles with naturally changing characteristics.

The empirical distribution of returns is significantly different for almost 60% of events after 1 year. Option trading activity, measured by actively traded contracts of any maturity, remains significantly elevated a year after an event for about 55% of events. Short interest remains significantly increased after a year for 88% of events. Institutional ownership decreases significantly within the year after 67% of events. Our results become more robust as the event period increases, and all examined effects appear to be permanent.

Our market-wide analysis of bubble events over time reveals that martingale defects in the volatility surface tend to occur in periods of positive market returns.



Fig. 4 Average daily number of contracts available for calibration. For visual clarity, the plot shows 1 year rolling averages

Table 6 Tenor classifications

Tenor	1 week	2 weeks	1 month	3 months	6 months
DTE	2 – 7 days	12 – 16 days	25 – 35 days	80 – 100 days	165 – 195 days

Market corrections lead to an immediate collapse of the number of martingale defects across all stocks. As option contract availability increased over time, martingale defects are distributed across more stocks, but the relative occurrence of defects remains somewhat constant. The implication is that overoptimistic speculation is a constant element of markets, and that better availability of options can help distribute this irrational force across a larger number of underlyings.

The advantages of the presented simplified implementation of the martingale defect indicator make it a promising tool for future research. Of particular interest might be the correlation between the analyzed factors, the distribution of option trading activity between retail traders and institutional investors to isolate the propensity to take more risk than intended, and the particular effect of options with a lifetime of less than 2 weeks. Furthermore, our results might be of interest to regulators to help balance option availability, since an overabundance of options for small stocks might permit easier market cornering and thus manipulation, negatively affecting market stability and trust in institutions.

A Calibration results

Table 5 provides descriptive statistics of calibration results. The large sample size leads to a small number of calibrations with above-desirable $RSME_{IV}$ figures, but for the vast majority of observations these results indicate adequate fit. Comparing calibration results for the full sample with those observations where the resulting $A(\cdot)$ exceeds a set of thresholds confirms that the calibration achieves acceptable fit most of the time even under martingale defect conditions.

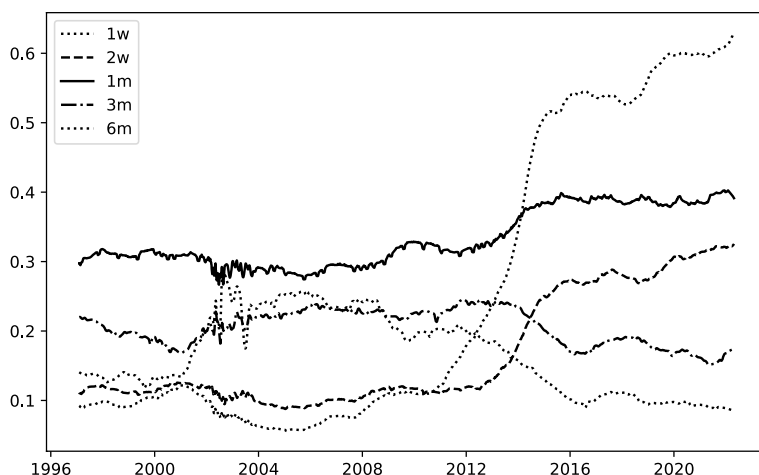


Fig. 5 Proportions of traded option contracts by tenor. For visual clarity, the plot shows 1 year rolling averages

B Total trading activity over time

Within the timespan of our dataset, option trading activity has increased. The number of underlyings with actively traded options as well as the number of options per underlying have both grown. This development propagates into our results, since more underlyings with active option trading implies a higher number of possible events. To measure and compensate for this trend, we use the total number of actively traded option contracts as proxy for market-wide option trading activity. We employ the same liquidity requirements as laid out in Sect. 4. Figure 3 provides the total number of daily active contracts for the SPX Index as well as the complete securities universe. As activity is fluctuating strongly from day to day, we use 252 trading day rolling averages. Figure 4 provides the average number of contracts available for the calibration of the SABR model.

C Relative trading activity over time by maturity

Appendix B examines the total number of actively traded contracts. In this section, we will examine the proportional trading activity by maturity, as far as it is relevant to our analysis. We separate contracts by their remaining lifetime in days. These slices are commonly called *tenors*. Table 6 provides a list of tenor definitions.

Figure 5 provides an overview of the development of available option contracts within a tenor. It shows the number of contracts within a range of days-to-maturity relative to the total number of contracts available. We apply the liquidity

Table 7 Number of events and number of unique affected companies for different persistence and threshold configurations

		By Persistence					
By Threshold		1 Month tenor			Full surface		
		1 day	2 days	3 days	1 day	2 days	3 days
0.01	Events	18526	3505	1339	38798	6731	2477
	Companies	1591	820	475	1923	1193	717
0.05	Events	9983	1518	529	21087	2745	856
	Companies	1391	527	249	1806	806	382
0.10	Events	6659	823	274	14602	1519	390
	Companies	1203	353	149	1674	561	217

The left half provides results based on options with approximately one month to expiration, the right half provides results based on the entire volatility surface. The upper row provides the number of events in total, the lower row shows across how many companies these events are spread out. A persistence of 1 day means that the indicator must be above threshold for one day, i.e. there is no persistence requirement

requirements laid out in Sect. 4. As the periods are not exactly adjacent, the sum of the relative proportions may be less than 100%. Each day, all available option contracts are sorted by days to expiration, and their proportions are compared over time. The proportion of available options with a lifetime of less than 1 month begins to increase in 2012. The proportion of options with a lifetime of approximately 1 month remains very roughly constant, while the proportion of longer options declines from 2012 onward. After 2012, the proportion of available contracts has increased strongly. The proportion of the one-month tenor is relatively stable, but has increased as well. Overall, the market appears to have embraced the availability of low-DTE contracts. This fundamental shift in the dataset motivates us to provide our findings based on the 1 month tenor as well as the full surface.

D Persistence

Table 7 provides details on the results discussed in Sect. 4. Furthermore, the number of affected stocks is disclosed.

E Results for full surface

Table 8 complements table 2 and provides results based on the full surface.

Table 9 complements Table 3 and provides results based on the full surface.

Table 10 complements Table 4 and provides results based on the full surface.

Table 8 Aggregate results for t -tests of the number of actively traded option contracts before and after a martingale defect event, based on the calibration of the SABR model against the full volatility surface

Period t_{event}	Full surface			One-month tenor				
	N	t -statistics	p -values	N	t -statistics	p -values		
		$p \leq 5\%$			$p \leq 5\%$			
		Mean	$p \leq 1\%$		Mean	$p \leq 1\%$		
21	1149	0.19	22.55% (7.13%)	12.19% (3.17%)	780	0.38	10.01% (2.16%)	4.58% (0.86%)
63	1196	0.36	32.41% (9.81%)	23.34% (5.60%)	953	0.18	20.63% (5.81%)	12.13% (2.76%)
84	1195	0.45	33.86% (17.50%)	24.66% (12.73%)	930	0.32	19.19% (9.85%)	10.50% (6.67%)
105	1168	0.67	35.82% (19.14%)	28.28% (13.44%)	944	0.44	23.08% (10.18%)	14.73% (6.19%)
126	1179	0.79	39.96% (22.60%)	31.25% (16.67%)	956	0.49	25.74% (14.27%)	17.34% (9.92%)
189	1152	0.95	42.70% (23.72%)	34.07% (17.96%)	957	0.61	28.42% (14.42%)	20.40% (9.56%)
252	1090	1.35	45.45% (29.85%)	37.25% (24.20%)	912	0.68	31.13% (20.06%)	22.04% (13.84%)

Event Period t_{event} provides the number of trading days before and after an event under consideration. Results under *Full Surface* count all available options, while results under *One-month Tenor* count only contracts within the 1-month tenor. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. Mean t -statistic reports the mean of the t -statistic of all events. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p -values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses

Table 9 Aggregate results for t -tests of outstanding short interest divided by ADTV before and after a martingale defect event, based on the calibration of the SABR model against the full volatility surface

Period t_{event}	N	t -statistics	p -values			
			Mean	$p \leq 5\%$	$p \leq 1\%$	
21	1060	2.185	75.2%	(46.5%)	66.7%	(38.5%)
63	1042	1.250	77.9%	(64.4%)	71.8%	(64.2%)
84	1033	0.979	80.8%	(73.7%)	75.4%	(63.1%)
105	1010	0.867	81.0%	(60.9%)	75.7%	(48.1%)
126	985	0.647	83.1%	(79.5%)	77.9%	(66.5%)
189	950	0.879	86.3%	(73.8%)	81.5%	(63.5%)
252	884	0.998	88.0%	(75.1%)	83.7%	(63.2%)

Event Period t_{event} provides the number of trading days before and after an event under consideration. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. Mean t -statistic reports the mean of the t -statistic of all events. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p -values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses

Table 10 Aggregate results for t -tests of the percentage of institutional ownership before and after a martingale defect event, based on the calibration of the SABR model against the full volatility surface

Period t_{event}	N	t -statistics	p -values			
			Mean	$p \leq 5\%$	$p \leq 1\%$	
21	1060	-0.264	59.0%	(43.1%)	52.9%	(34.2%)
63	1042	-0.479	65.1%	(49.1%)	62.9%	(43.8%)
84	1033	-0.176	64.5%	(48.6%)	62.0%	(49.4%)
105	1010	-0.319	63.8%	(57.4%)	61.4%	(54.0%)
126	985	-0.244	64.0%	(58.5%)	61.7%	(56.4%)
189	950	-0.308	65.4%	(57.4%)	63.3%	(53.5%)
252	884	-0.102	64.7%	(59.6%)	63.1%	(56.8%)

Event Period t_{event} provides the number of trading days before and after an event under consideration. N refers to the number of events after discarding stocks with insufficient data and overlapping event periods. Mean t -statistic reports the mean of the t -statistic of all events. $p \leq 5\%$ and $p \leq 1\%$ report the quantile of p -values in the sample. The corresponding results of the placebo study with shuffled event dates are shown in parentheses

Acknowledgements We thank the participants of the IQAM Research Seminar for insightful comments and discussion. We owe gratitude to Deka Investment for continuous support of this project and providing the data. Furthermore, we are indebted to Gurdip Bakshi, the editor, and two anonymous referees for their detailed reviews of this paper and for suggesting many improvements. Any remaining errors are our own.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended

use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bakshi, G. S., Crosby, J., Gao, X., & Xue, J. (2021). Volatility Uncertainty and VIX Futures Contango. *Fox School of Business Research Paper, SMU Cox School of Business Research Paper*, No. 21-19. ISSN 1556-5068. <https://doi.org/10.2139/ssrn.3930703>.
- Bakshi, G., Crosby, J., & Gao, X. (2022). Dark matter in (Volatility and) equity option risk premiums. *Operations Research*, 70(6), 3108–3124.
- Biagini, F., Gonon, L., Mazzon, A., & Meyer-Brandis, T. (2022). Detecting asset price bubbles using deep learning.
- Biagini, F., Mazzon, A., & Perkkio, A.-P. (2020). *Optional projection under equivalent local martingale measures*.
- Biagini, F., Föllmer, H., & Nedelcu, S. (2014). Shifting martingale measures and the birth of a bubble as a submartingale. *Finance and Stochastics*, 18(2), 297–326. <https://doi.org/10.1007/s00780-013-0221-8>
- Blocher, J., Dong, X., Ringgenberg, M. C., & Savor, P. G. (2021). Short Covering.
- Camerer, C. (1989). Bubbles and fads in asset prices. *Journal of Economic Surveys*, 3(1), 3–41. <https://doi.org/10.1111/j.1467-6419.1989.tb00056.x>
- Chen, H., Dou, W. W., & Kogan, L. (2022). Measuring 'Dark Matter' in asset pricing models.
- Cheridito, P., Filipovic, D., & Kimmel, R. L. (2007). Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics*, 83, 123–170. <https://doi.org/10.2139/ssrn.524724>
- Conlon, J. R. (2004). Simple finite horizon bubbles robust to higher order knowledge. *Econometrica*, 72(3), 927–93. <https://doi.org/10.1111/j.1468-0262.2004.00516.x>
- Cox, A. M. G., & Hobson, D. G. (2005). Local martingales, bubbles and option prices. *Finance and Stochastics*, 9(4), 477–492. <https://doi.org/10.1007/s00780-005-0162-y>
- Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3), 229–263. [https://doi.org/10.1016/0304-405X\(79\)90015-1](https://doi.org/10.1016/0304-405X(79)90015-1)
- Delbaen, F., & Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(1), 463–520. <https://doi.org/10.1007/BF01450498>
- Farooque, M. (2022). 7 Tech Stocks to Buy That Could 10X by 2025.
- Florens-Zmirou, D. (1993). On estimating the diffusion coefficient from discrete observations. *Journal of Applied Probability*, 30(4), 790–804. <https://doi.org/10.2307/3214513>
- Fusari, N., Jarrow, R., & Lamichhane, S. (2022). Testing for asset price bubbles using options data.
- Gatheral, J. (2006). *The volatility surface: A practitioner's guide*. Wiley.
- Hagan, P. S., Kumar, D., Lesniewski, A. S., & Woodward, D. E. (2002). *Managing Smile Risk*. *Wilmott Magazine*, 1, 84–108.
- Heston, S. L., Loewenstein, M., & Willard, G. A. (2007). Options and bubbles. *Review of Financial Studies*, 20(2), 359–390. <https://doi.org/10.1093/rfs/hhl005>
- Jarrow, R. A. (1992). Market manipulation, bubbles, corners, and short squeezes. *The Journal of Financial and Quantitative Analysis*, 27(3), 311. <https://doi.org/10.2307/2331322>
- Jarrow, R. A. (2015). Asset price bubbles. *Annual Review of Financial Economics*, 7, 201–218.
- Jarrow, R. A., Kchia, Y., & Protter, P. (2011). How to detect an asset bubble. *SIAM Journal on Financial Mathematics*, 2(1), 839–865.
- Jarrow, R., & Kwok, S. (2021). Inferring financial bubbles from option data. *Journal of Applied Econometrics*, 36(7), 1013–1046.
- Jarrow, R. A., & Protter, P. (2010). *The martingale theory of bubbles: implications for the valuation of derivatives and detecting bubbles*. Johnson School Research Paper Series.
- Jarrow, R., & Protter, P. (2013). Positive alphas, abnormal performance, and illusory arbitrage. *Mathematical Finance*, 23(1), 39–56. <https://doi.org/10.1111/j.1467-9965.2011.00489.x>
- Jarrow, R. A., Protter, P., & Shimbo, K. (2007). Asset price bubbles in complete markets. *Advances in Mathematical Finance*. https://doi.org/10.1007/978-0-8176-4545-8_7

- Jarrow, R. A., Protter, P., & Shimbo, K. (2010). Asset price bubbles in incomplete markets. *Mathematical Finance*, 20(2), 145–185. <https://doi.org/10.1111/j.1467-9965.2010.00394.x>
- Le Floc'h, Fabien, & Kennedy, Gary J. (2014). *Explicit SABR calibration through simple expansions*.
- Loewenstein, M., & Willard, G. A. (2000). Local martingales, arbitrage, and viability. *Economic Theory*, 16(1), 135–161. <https://doi.org/10.1007/s001990050330>
- Maneesh, S.D., Arnab S., Brian P., Si G., & Elias K.(2020). *Barclays US equity derivatives strategy: Impact of retail options trading*. Technical report.
- Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1), 141. <https://doi.org/10.2307/3003143>
- Mohrschladt, H., & Schneider, J. C. (2021). Idiosyncratic volatility, option-based measures of informed trading, and investor attention. *Review of Derivatives Research*, 24(3), 197–220. <https://doi.org/10.1007/s11147-021-09175-7>
- MSCI. (2017). *MSCI USA IMI sector indexes*. Technical report, MSCI.
- Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, 7(4), 308–313. <https://doi.org/10.1093/comjnl/7.4.308>
- Obayashi, Y., Protter, P., & Yang, S. (2016). The lifetime of a financial bubble.
- Piironen, P., Roininen, L., & Simon, M. (2019). *Brexit risk implied by the sabr martingale defect in the EUR-GBP smile*. Working Paper.
- Piironen, P., Roininen, L., Schoden, T., & Simon, M.. (2018). Asset price bubbles: An option-based indicator.
- Protter, P. (2016). Mathematical models of bubbles. *Quantitative Finance Letters*, 4(1), 10–13. <https://doi.org/10.1080/21649502.2015.1165863>
- SEC. (2021). staff report on equity and options market structure conditions in early 2021. Technical report, Securities Exchange Commission.
- Smith, A. (2007). *Experimental drug could be first to stall Alzheimers*. <https://money.cnn.com/2007/08/06/news/companies/myriad/index.htm>.
- Sun, Z. (2021). is sundial growers a buy? <https://www.fool.com/investing/2021/05/14/is-sundial-growers-a-buy/>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.