Is the Current Bull Market A Bubble? An Empirical Investigation

Soon Hyeok Choi* Robert A. Jarrow†

April, 2024

Click here for the most recent version

Abstract

Evidence of excess volatilities at high asset prices is associated with bubbles. We propose a new asset price bubble testing methodology based on volatility estimates. Examining the current U.S. equity bull market, we find that the S&P 500 and Dow Jones do not exhibit bubbles, but the Nasdaq does. We stress test our methodology with individual stocks and simulation models to build confidence in the procedure. We show that these results are robust to various adjustments for outliers.

Keywords: Asset Price Bubbles, Explosive Volatility, Local Martingale, Equities

JEL Codes: G12, C50, C51, C59

*Saunders College of Business, Rochester Institute of Technology. schoi@saunders.rit.edu.
†SC Johnson College of Business, Cornell University. raj15@cornell.edu.

1 Introduction

Asset price bubbles are characterized by three elements: a deviation from an asset's fundamental value, extended price run-ups, and an eventual crash (Blanchard, 1979; Diba and Grossman, 1988; Jarrow et al., 2010; Brunnermeier and Oehmke, 2013; Fama, 2014; Shiller, 2016; Greenwood et al., 2019). Empirically identifying asset price bubbles is challenging because the traditional methodology requires an explicit estimation of the asset's fundamental value. The identification requires postulating and estimating a stochastic process for the asset's cash flows, risk premium, default-free interest rates, and liquidation value. In the literature, there is a substantial disagreement on the estimation techniques. As a consequence, given the resulting controversy, it is widely believed that one cannot empirically test for the existence of price bubbles.

The purpose of this paper is to show that this common belief is false, and that one can easily test for the existence of price bubbles by using a new methodology that side-steps these difficulties in estimating an asset's fundamental value. This methodology is based on the local martingale theory of bubbles. Therein, a bubble can be identified solely by studying the characteristics of the market price process itself, in particular its volatility (under reasonable hypotheses documented below). If the asset's volatility increases sufficiently fast with the level of the asset's price, then this is a necessary and sufficient condition for the existence of a price bubble.[1] The asset's volatility is easily estimated and this necessary and sufficient condition empirically tested. We apply this methodology to the current U.S. equity market to determine if it is experiencing a price bubble. A question of considerable current interest.

This paper also makes a second contribution to the literature. We refine and extend the statistical methodology contained in Choi and Jarrow (2022) for testing asset price bubbles. The existing methodology has six limitations. The first is that this methodology needs to be extended to include cash flows. Second, the variance estimator needs to be augmented to include non-equal price observation times and price level intervals.[2] Third, although the existing volatility estimator is consistent, it is biased for small sample sizes. Fourth, the hypothesis testing procedure is conservative, being based on upper and lower bounds for the "true" volatility function, which results in regions where there are inconclusive results with respect to the existence of price bubbles. Fifth, the robustness procedure does not include information from the estimated volatility's

[1]This theory is reviewed in Section 2 below.

[2]The detailed reasons for this and subsequent statements are given below in section 3.

sampling distribution. And finally, the standard errors in a regression estimating the upper and lower bounds for the volatility functions are not adjusted for probable heteroskedasticity and autocorrelation.

This paper addresses these six limitations. First, we include cash flows. Second, we allow for non-equal time and price intervals. Third, we create and implement a small sample bias adjustment. Fourth, if the hypothesis testing approach yields an inconclusive result, we employ Bayesian statistics to provide a posterior probability of the asset price exhibiting a bubble. Fifth, we develop a robustness check based on the volatility's sampling distribution. And finally, in the volatility function's upper and lower bound estimations, we adjust the regression estimates' standard errors for heteroskedasticity and autocorrelation.

To validate the bubble testing methodology, we simulate two hypothetical markets, one with the asset price exhibiting a bubble and one without. We simulate 10,000 paths for the risky asset's prices over 3 years in both markets, and apply our bubble testing methodology to see if it correctly identifies the bubble and no-bubble markets. The simulation validates the methodology. For the no-bubble market, using a hypothesis test at the 95% significance level, only 0.04% of the simulated paths are misclassified as bubbles. Exactly as should occur. For the bubble market, using a conservative hypothesis test, 71% are initially classified correctly as bubbles. For those simulated paths in the bubble market that are inconclusive for this hypothesis test, approximately 41% exhibit a posterior probability of a bubble of more than 90%. The remaining simulated paths remain inconclusive, with the exception that 39% exhibit a posterior probability of a bubble of less than 10%. This is to be expected given the hypothesis testing is conservative, and there is error in the price process's path due to simulating a continuous stochastic differential equation with a discretized Euler scheme.

We apply this refined bubble detection technology to the U.S. equity market from March 2023 to March 2024 using daily price data to see if the current bull market is a price bubble. We present three main findings. The S&P 500 and Dow Jones indices do not exhibit price bubbles, while the Nasdaq index does. Various robustness checks are performed on the estimation that confirm the validity of these conclusions.

Second, we provide various case studies to provide anecdotal, but confirmatory evidence, that the method does correctly identify bubbles. We select three stocks that are often alleged to contain bubbles (Bitcoin, Meta, NVIDIA) and three stocks that are not (JP Morgan, Bank of America, Wells Fargo).[3] When we apply our methodology to these securities, Bitcoin and NVIDIA have price bubbles, while Meta does not. For the banks, JP Morgan, Wells Fargo have no bubbles, but Bank of America does.

Finally, we test our methodology on a recent event with respect to Lyft. On February 14th 2024, Lyft announced an erroneous earnings projection that stated a 500 basis point margin instead of a 50 basis point margin. Although the CEO corrected this mistake in less than an hour, the firm's stock surged at least 67% higher based on the incorrect earnings release.[4] Intuitively, if Lyft did not undergo a fundamental change between February 13th and February 14th , this phenomenon reflects a price bubble. We apply our methodology to two time periods: March 1st 2023 to February 13th 2024 (pre-announcement) and March 1st 2023 to March 14th 204 (post-announcement). We show that at a 1% significant level, Lyft was not in a bubble before the announcement, but reflected a bubble after the announcement.

Our paper is related to two literatures. First, it relates to an econometric literature testing for price bubbles (Jarrow et al., 2011a,b; Shiryaev et al., 2016; Phillips et al., 2015; Phillips and Shi, 2020; Jarrow and Kwok, 2021; Choi and Jarrow, 2022). Similar to our paper, these studies primarily focus on estimating the explosive feature of price bubbles (e.g., the Feller test, augmented Dickey-Fuller test). Our method differs from those using the local martingale theory of bubbles in our extrapolation procedure. Our paper also relates to the statistical literature detecting and adjusting for outliers (Grubbs, 1969; Rocke and Woodruff, 1996; Aguinis et al., 2013). Our methodology uses the convex hull of volatility estimates to conservatively approximate the minimum and maximum area under the extrapolated volatility function. As such, it is potentially vulnerable to large and small volatility estimates. We provide a new technique for modifying these outliers based on the sampling distribution to check for this possibility.

This paper is organized as follows. Section 2 provides a summary of the local martingale theory of bubble. Section 3 juxtaposes the existing and new methodologies. Section 4 documents the simulation results, and Section 5 presents the empirical results. Section 6 provides robustness tests, and Section 7 concludes.

---

[3]Although Bitcoin is not a stock, Bitcoin is a cryptocurrency widely documented to have periodic price bubbles (Chaim and Laurini, 2019; Choi and Jarrow, 2022). Therefore, we include it in the set of alleged-bubble stocks.

[4]Rana, Preetika. "Lyft Shares Surge as Strong Earnings Report Offsets Typo Confusion", The Wall Street Journal, 14 February, 2024.

2 The Local Martingale Theory of Bubbles

This section briefly reviews the local martingale theory of bubbles. For a detailed presentation see Jarrow (2018), chapter 3. The local martingale theory of bubbles is based on a continuous time, continuous trading, frictionless and competitive market model over a finite horizon $[0, T]$. Traded are a default-free money market account (mma) and risky assets. Without loss of generality, we assume that there is only one risky asset traded.

Denote the time t market price of the risky asset by $\hat{S}_t$, and assume that it is always non-negative. Let G_t denote the asset's cumulative cash flow at time t, starting with $G_0 = 0$. The cumulative cash flow process is non-decreasing and therefore a finite variation process. Denote the time t value of the money market account (mma) by

$$B_t = e^{\int_0^t r_s ds}$$

where $B_0 = 1$ and r_t is the default-free spot rate of interest.

Starting at time 0, the value of a position in the stock plus reinvested cash flows (in the mma) over $[0, t]$ is

$$\hat{S}_t + \left(\int_0^t \frac{1}{B_s} dG_s \right) B_t.$$

We suppose that the market is arbitrage-free (i.e., the market satisfies No Free Lunch with Vanishing Risk). Hence, by the First Fundamental Theorem of asset pricing, there exists a risk neutral probability $\mathbb{Q}$, equivalent to the statistical probability $\mathbb{P}$, such that the normalized asset's price process plus reinvested cash flows,

$$\frac{\hat{S}_t + \left(\int_0^t \frac{1}{B_s} dG_s \right) B_t}{B_t} = \frac{\hat{S}_t}{B_t} + \int_0^t \frac{1}{B_s} dG_s,$$

is a $\mathbb{Q}$ local-martingale. A local-martingale is a generalization of a martingale. Equivalent means that $\mathbb{Q}$ and $\mathbb{P}$ agree on zero probability events.

The market is not assumed to be complete, hence, by the Second Fundamental Theorem of asset pricing, there could be an infinite number of risk neutral probabilities. If the market is incomplete, we assume that a unique risk neutral probability $\mathbb{Q}$ is chosen by the market, either via an economic equilibrium or via the asset market being embedded in a larger market including traded derivatives that is complete.

The asset's fundamental value at time t, F_t, is defined to be the expected value of the asset's liquidation payoff at time T plus all reinvested cash flows over $[t,T]$, discounted to the present, i.e.

$$F_t := E_t^{\mathbb{Q}}\left(\frac{\hat{S}_T}{B_T} + \int_t^T \frac{1}{B_s} dG_s\right) B_t \tag{1}$$

where $E_t^{\mathbb{Q}}$ denotes the conditional expectation at time t using the risk neutral probabilities $\mathbb{Q}$. The use of the risk neutral probabilities $\mathbb{Q}$ adjusts for risk in computing this present value. This is the classical definition of an asset's fundamental value in the economics literature.

The asset's price bubble is defined to be

$$\beta_t = \hat{S}_t - F_t. \tag{2}$$

Note that the stock price $\hat{S}_t$ is after all cash flows have been paid at time t.

Using the definition of the fundamental value, we can rewrite the normalized bubble's magnitude as

$$\frac{\beta_t}{B_t} = \underbrace{\left(\frac{\hat{S}_t}{B_t} + \int_0^t \frac{1}{B_s} dG_s\right)}_{\substack{\text{normalized asset price} \\ \text{+ reinvested cash flows over } [0,t] \\ \text{(A1)}}} - \underbrace{E_t^{\mathbb{Q}}\left(\frac{\hat{S}_T}{B_T} + \int_0^T \frac{1}{B_s} dG_s\right)}_{\substack{\text{expected normalized liquidation value} \\ \text{+ reinvested cash flows over } [t,T] \\ \text{(A2)}}}.$$

Since (A1) is a $\mathbb{Q}$ local-martingale, it is a non-negative $\mathbb{Q}$ supermartingale. This implies that

$$\frac{\hat{S}_t}{B_t} + \int_0^t \frac{1}{B_s} dG_s \geq E_t^{\mathbb{Q}}\left(\frac{\hat{S}_T}{B_T} + \int_0^T \frac{1}{B_s} dG_s\right).$$

We can further deduce that the normalized asset's price bubbles is non-negative, i.e. $\left(\frac{\beta_t}{B_t}\right) \geq 0$. Given the normalized fundamental value (A2) is a $\mathbb{Q}$ martingale, it follows that a price bubble exists ($\beta > 0$) if and only if the asset's price plus reinvested cash flows (A1) is not a $\mathbb{Q}$ martingale. In this case, we say that (A1) is a *strict* $\mathbb{Q}$ local martingale. This key insight provides the theoretical basis of our bubble detection methodology.

> The statistical methodology for detecting asset price bubbles tests to see if the price process $\left(\frac{\hat{S}_t}{B_t} + \int_0^t \frac{1}{B_s} dG_s\right)$ is a strict $\mathbb{Q}$ local-martingale (bubble) or a $\mathbb{Q}$ martingale (no bubble).

3 The New Statistical Methodology

This paper provides a new statistical methodology for detecting asset price bubbles, extending the previous approaches used by (Jarrow et al., 2011a; Obayashi et al., 2017; Jarrow and Kwok, 2021; Choi and Jarrow, 2022). First, we briefly review its most recent application to detecting bubbles in cryptocurrencies and foreign currencies by Choi and Jarrow (2022). Second, we explain our new refinements and contributions to the bubble detection methodology.

3.1 The Existing Methodology

Choi and Jarrow (2022) study bubbles in assets that do not have cash flows, i.e. $G_t \equiv 0$. To simply the notation, let $S_t := \frac{\hat{S}}{B}$ denote the normalized risky asset's price process.[5] And, to simplify the exposition, we will call S the risky asset's price process, dropping the qualifier "normalized."

The statistical methodology assumes that S_t follows the diffusion process

$$dS_t = \mu(S_t)dt + \sigma(S_t)dW_t \tag{3}$$

where S_0 is a constant and W_t is a standard Brownian motion with $W_0 = 1$.

There are two key characteristics of this diffusion process that are exploited for the testing of an asset price bubble. The first is that S is a strict $\mathbb{Q}$ local martingale if and only if S is a strict $\mathbb{P}$ local martingale (Jarrow et al., 2022). This implies that we do not need to estimate or determine the risk neutral probability $\mathbb{Q}$ when testing for price bubbles. The second is that the characterization of S being a strict $\mathbb{P}$ local martingale depends solely on the asset's volatility function, $\sigma(x)$. This is evidenced by the following result.

> The normalized price process S is a strict local martingale under $\mathbb{P}$ if and only if
>
> $$\int_{\varepsilon}^{\infty} \frac{x}{\sigma(x)^2} ds \;<\; \infty \qquad \text{for any} \quad \varepsilon > 0. \tag{4}$$

Hence, testing for a price bubble is equivalent to investigating whether the integral in (4) is finite or not. If the integral converges, there is a bubble. If it diverges, then there

[5] In the estimation below, the stock's price is divided by the value of a money market account before performing the estimation methodology.

is no bubble. Note that the integral is finite if the variance function increases at a faster rate than the price implying the bubbles are associated with large return variances at high price levels.

To estimate the volatility function at the level x, the observation period is partitioned into the discrete time steps $t_1 = 0, t_2, t_3, ..., t_n = T$ where n is the total number of price observations over the time interval $[0, T]$. Then, assuming that the time steps are of equal length, $t_{i+1} - t_i = \frac{1}{n}$ units of a year, the estimator at the level x [expression (5), p. 842, (Jarrow et al., 2011b)] of the variance function is given by

$$V_n(x) = \frac{\sum_{i=1}^{n} 1_{\{|S_{t_i}-x|<h_n\}} n(S_{t_{i+1}} - S_{t_i})^2}{\sum_{i=1}^{n} 1_{\{|S_{t_i}-x|<h_n\}}}. \tag{5}$$

where

$$1_{\{|S_{t_i}-x|<h_n\}} = \begin{cases} 1 & if \quad |S_{t_i} - x| < h_n \\ 0 & \quad otherwise \end{cases}$$

and h_n is a constant depending upon the sample size n. Here, the the denominator counts the number of stock prices observed in the price level interval $[x - h_n, x + h_n]$ and the numerator computes the variance estimator for each time step, $(S_{t_{i+1}} - S_{t_i})^2$, prorated per year, where S_{t_i} is in the interval $[x - h_n, x + h_n]$.

This is a consistent estimator with $V_n(x) \to \sigma^2(x)$ if: (i) σ is bounded above and below from zero with three continuous and bounded derivatives, and (ii) $nh_n \to \infty$ and $nh_n^4 \to 0$ as $n \to \infty$. It can be shown (see the appendix) that if $nh_n^3 \to 0$ as $n \to \infty$, then for large N_x^n, the sampling distribution of this estimator is approximately

$$V_n(x) \sim \Phi\left(\sigma^2(x), 2\frac{V_n^2(x)}{N_x^n}\right) \tag{6}$$

where $\Phi(mean, variance)$ is the normal distribution, and $N_x^n := \sum_{i=1}^{n} 1_{\{|S_{t_i}-x|<h_n\}}$ counts the number of observations S_{t_i} across $i = 1, ..., n$ in the interval $[x - h_n, x + h_n]$. Note that this is an asymptotic distribution for the estimator as $N_x^n \to \infty$, when $h_n \to 0$ and $n \to \infty$.

Next, to test for convergence of the integral (4), the volatility function $\sigma(x)$ needs to be estimated over all the asset price levels $x \in [0, \infty)$. Here, the price level range is partitioned into equally spaced subintervals $[0 = x_0, x_1, x_2, \ldots, x_K = max\{S_{t_i} : i = 1, ..., n\}]$ for some $K > 0$ where $x_j - x_{j-1} = 2h_n$ for $j = 1, ..., K$. Because $max\{S_{t_i} : i = 1, ..., n\}$ is finite, this implies that to check for convergence of the interval, one must extrapolate the volatility's behavior from the observed price levels to those that the asset's price has not yet reached.

Choi and Jarrow (2022) developed an extrapolation technique based on bounding the volatility function using two convex hulls of the estimated volatilities.[6] Fix a trading horizon $[0, T]$ where we observe prices $\{S_{t_i}\}_i$ with $i \in \{1, 2, ..., T\}$. We estimate volatilities for each price partition $\{h_j\}_j$ such that we produce a set of price-volatility pairs $\{(S_j, \sigma(S_j))\}_j$.[7] The procedure consists of the following steps.

1. Extrapolation: Select the best power functions $\sigma_k(x) = \alpha_k x^{\beta_k}$ to fit both the lower ($k = l$) and upper ($k = u$) convex hulls. By construction $\sigma_l^2(x) \leq \sigma^2(x) \leq \sigma_u^2(x)$. Then, estimate the regression

$$\ln(\sigma_k(S_j)) = \ln(\alpha_k) + \beta_k \ln(S_j) + \varepsilon_j \tag{7}$$

with $\alpha_k \geq 0$ and ε_j the regression residuals to obtain the estimated regression coefficients $\hat{\beta}_k$ for $k \in \{u, l\}$.

2. Evaluation (Point Estimation): Define the integrals $\mathscr{I}_u := \int_1^\infty \frac{x}{\sigma_l^2(x)} dx$ and $\mathscr{I}_l := \int_1^\infty \frac{x}{\sigma_u^2(x)} dx$. Note the subscripts for the upper and lower integrals have been reversed from the variance functions, so that

$$\mathscr{I}_u > \mathscr{I} > \mathscr{I}_l$$

where $\mathscr{I}$ in given in expression (4).

   (a) First compute the estimate $\hat{\beta}_l$ in order to evaluate the convergence of the upper bound $\mathscr{I}_u$ for the integral $\mathscr{I}$. The upper bound is evaluated using the lower convex hull's approximating function $\sigma_l^2(x)$. If the estimated coefficient $\hat{\beta}_l > 1$, then this implies the lower convex hull's integral $\mathscr{I}_u < \infty$ converges and there is a bubble.

   (b) If the estimate $\hat{\beta}_l$ implies the lower convex hull's integral $\mathscr{I}_u = \infty$ diverges, then this does not guarantee divergence of $\mathscr{I}$. In this case, use the upper convex hull to obtain the estimate $\hat{\beta}_u$ to evaluate divergence of the lower bound $\mathscr{I}_l$ for the integral $\mathscr{I}$.

[6] Jarrow et al. (2011a,b); Chaim and Laurini (2019) extrapolate the volatility function using a Gaussian kernel and Reproducing Kernel Hilbert Spaces.

[7] Note the total number of price partitions h_n depends on the sample size n.

(c) If the estimated coefficient $\hat{\beta}_u \leq 1$, then this implies that the lower integral $\mathscr{I}_l = \infty$ diverges. Thus, $\mathscr{I} = \infty$ diverges, and there is no bubble.

(d) If the point estimate $\hat{\beta}_u$ implies the lower integral $\mathscr{I}_l < \infty$ converges, then the test is inconclusive. This occurs if $\hat{\beta}_u > 1$.

3. Evaluation (Hypothesis Testing): The hypothesis test uses the point estimates of β obtained in the previous section. The following algorithm controls for both Type I and Type II errors.

   (a) Step 1: Test the null hypothesis of no bubble using the point estimate $\hat{\beta}_l$ to evaluate the upper bound $\mathscr{I}_u$ on the true integral at the 0.95 confidence level. Reject the null if $\hat{\beta}_l > 1 + 1.645\hat{\sigma}_l$. If rejected, stop. The conclusion is that a bubble exists. Otherwise due to the fact that this is upper bound and there is potentially a large Type II error, go to step 2.

   (b) Step 2: Test the null hypothesis of a bubble using the point estimate $\hat{\beta}_u$ to evaluate the lower bound $\mathscr{I}_l$ on the true integral at the 0.95 confidence level. Reject the null if $\hat{\beta}_u \leq 1 - 1.645\hat{\sigma}_u$. If rejected, stop. The conclusion is that there is no bubble. Otherwise, go to step 3.

   (c) Step 3: Stop. The testing is inconclusive, because step 1 accepts the hypothesis of no bubble and step 2 accepts the hypothesis of a bubble, both tests having potentially large Type II errors.

For a robustness check, Choi and Jarrow (2022) winsorize the largest and smallest volatility estimates by replacing them with the respective second largest and smallest estimates.

3.2 Limitations

The existing bubble testing methodology has six limitations in applications. First, the estimation methodology was constructed for assets with no cash flows. Clearly, this is violated for many assets of interest.

Second, the existing variance estimator has equally spaced time and price level partitions. For time, the partitions are denoted $\frac{1}{n}$, where n is the number of price observations. Fixing n as the sample size, for an application the observations may be at different time intervals, e.g. transaction times (which are unequal) or daily (weekends issues). For this reason, it is important to adjust the variance estimator to handle these different possibilities. Given the variance estimator is consistent if both $h_n \to 0$ and

$n \to \infty$, it must be the case that $N_x^n \to \infty$ for the estimator to have a small sampling standard error. This process involves a tradeoff between the size and number of price bins we choose. On the one hand, if the size of the price intervals (i.e., $[x - h_n, x + h_n]$) becomes too tight (i.e., $h_n \approx 0$), then each bin will virtually have no prices observed. On the other hand, if the number of observed prices (i.e., N_x^n) is too large (i.e., $M \approx \infty$), then the number of partitions will naturally decrease making it challenging to generate a sufficient pair of price-volatility estimates. Unfortunately, for the asset price partition $[0 = x_0, x_1, x_2, \ldots, x_K = max\{S_{t_i} : i = 1, ..., n\}]$ equally spaced at $2h_n$, the volatility estimates $V_n(x)$ for x close to zero and x close to $max\{S_{t_i} : i = 1, ..., n\}$ will have the smallest values for N_x^n, and consequently the largest sampling standard errors. These potential large sampling errors could significantly impact the extrapolation methodology employed.

Third, since it is an consistent estimator, it is likely that the estimate is biased for small sample sizes. And, the smaller the sample size N_x^n in any price level interval $[x - h_n, x + h_n]$, potentially the larger the bias. As with the first limitation, this is likely to impact the the volatility estimates $S_n(x)$ for x close to zero and x close to $max\{S_{t_i} : i = 1, ..., n\}$, due to the smaller values for N_x^n.

The intuition for this bias can be explained as follows. If $E_{t_i}(S_{t_{i+1}}) > S_{t_i}$ (typical for stocks that require a risk premium) and $\sigma^2(x)$ is increasing in x (necessary for a price bubble), then the estimator will be biased upward. The bias stems from the fact that $1_{\{|S_{t_i} - x| < h_n\}}$ uses the points $S_{t_i} \in [x - h_n, x + h_n]$ close to x to estimate the variance $\sigma^2(x)$ for the single point $\{x\}$. It does this by estimating the variance for a point S_{t_i} in a neighborhood of x, $\sigma^2(S_{t_i})$. On average $\sigma^2(S_{t_i})$ will be close to $\sigma^2(x)$ because $S_{t_i} \in [x - h_n, x + h_n]$; sometimes above, sometimes below x. But, the estimator computes the sample variance by using the next observation $S_{t_{i+1}}$ as well, for which $\sigma^2(S_{t_{i+1}})$ will be larger than $\sigma^2(S_{t_i})$ if $E_{t_i}(S_{t_{i+1}}) > S_{t_i}$ and $\sigma^2(x)$ is increasing. Intuitively, using both $\sigma^2(S_{t_i})$ and $\sigma^2(S_{t_{i+1}})$ to compute the estimator (roughly, their average) will result in an upward bias in the estimate for $\sigma^2(x)$.

Fourth, the testing method's reliance on the upper and lower bounds of the convex hull points renders the testing to be rather conservative. Consequently, there exist inconclusive regions where one cannot assert the presence of price bubbles. This is problematic in practice where one wants to know if an asset price exhibits a bubble. For this inconclusive region, one would like to compute a point estimate for the probability that the asset exhibits a bubble.

Fifth, the existing procedure's choice of replacing the largest and smallest estimated variances by the second largest and smallest estimates is somewhat arbitrary. If the second largest and smallest estimates are materially different from the largest and smallest counter parts, the resulting convex hull might be substantially differently than that from the initial interpolated points. Utilizing the sample distribution statistic improves the robustness of the method.

Finally, the regression residuals may exhibit heteroskedasticity or autocorrelation. Although the estimated price-volatility pairs are cross-sectional data (i.e., an estimate from the snapshot over a fixed trading period), there is a form of ordering preserved. For example, when an asset price increases significantly today, it typically remains elevated for a period of time. The existing method prescribes a convex hull onto this potentially autocorrelated data, and this ordering can cause the residuals to be correlated with respect to the asset price level. Also, the true volatility function's envelopes are only being approximated by a power function. Homoskedasticity widens our estimate's standard errors and autocorrelation produces biased estimates, which affect the existing method's hypothesis testing.

3.3 The New Methodology

There are five extensions to the existing statistical methodology in this paper. First, we allow the risky asset to have cash flows. Second, we extend the methodology to allow unequal time and price level partitions. The purpose of which is to have a minimal sample size N_x^n across price level intervals, so that the sampling standard errors are more uniform across price levels, especially for the smallest and largest partition levels. Third, we introduce a small sample size bias adjustment. Fourth, we develop a point estimate for the probability of a price bubble conditional on obtaining an inconclusive result in hypothesis testing. Fifth, we provide a collection of robustness tests, based on the sampling distribution of the variance estimator. Each of these extensions are discussed next.

3.3.1 Cash Flows

The purpose of this subsection is to show that the introduction of cash flows does not impact the estimation methodology. With cash flows, the evolution of the risky asset's price plus reinvested cash flows $\left(S_t + \int_0^t \frac{1}{B_s} dG_s\right)$ is

$$dS_t + \frac{1}{B_t} dG_t = \left[\mu(S_t)dt + \frac{1}{B_t} dG_t\right] + \sigma(S_t) dW_t.$$

As evidenced by this evolution, the volatility of the risky asset price plus reinvested cash flows is identical to that of the risk asset's price process. Because the characterization of a bubble is based on expression (4), which only involves the volatility function $\sigma(x)$, and both S_t and $\left(S_t + \int_0^t \frac{1}{B_s} dG_s\right)$ have the same volatility function, a bubble exists in S_t if and only if it exists in $\left(S_t + \int_0^t \frac{1}{B_s} dG_s\right)$. Hence, we have proven the following result.

> The estimation methodology without cash flows applies to risky assets with cash flows.

3.3.2 Unequal Time Intervals and Price Level Partitions

The existing formula has equal time partitions, denoted $\frac{1}{n}$, where n is the number of price observations. Fixing n as the sample size, in applications, it is important to adjust the variance estimator to handle unequal time intervals. In addition, the existing methodology constructs a price level partitioning with equal price level intervals. As noted earlier, such a partitioning will have unequal sample sizes N_x^n for different price levels x. And, the different sample sizes will result in larger standard errors of the estimates for low sample size intervals. These intervals will impact the shape of the upper and lower convex hulls of the volatility functions. To make the sample sizes N_x^n more uniform across x, we allow the price level partitions to be increasing in the price level x, and we select the size of the intervals to give a lower bound on N_x^n across all x.

The appendix derives the modified variance estimator, given by

$$V_n\left(x_j\right) = \frac{\sum_{i=1}^n 1_{\left\{S_{t_i} \in \left[2\sum_{k=1}^{j-1} h_k, 2\sum_{k=1}^{j} h_k\right]\right\}} \left(S_{t_{i+1}} - S_{t_i}\right)^2 \cdot \frac{1}{[t_{i+1} - t_i]}}{N_{x_j}^n}, \tag{8}$$

where we partition the price axis into m bins in the following fashion. First, fix the size of the first partition $h_1 > 0$. Then, set $h_j = \theta \times j \times h_1$ for $j = 1, ..., m$ where $\theta \in (0, 1)$. The partition is

$$\{0, 2h_1, 2h_1 + 2h_2, 2\sum_{k=1}^{3} h_k, \cdots, 2\sum_{k=1}^{m} h_k\}$$

where m is an even number. This partitioning has the size of the price level interval increasing as the price level increases. For example, if $h_1 = 5$ and $\theta = \frac{1}{2}$, then the partition points are $\{0, 10, 17.5, 27.5, \cdots, 2\sum_{k=1}^{m} h_k\}$. Next, we choose $\{h_1, m, \theta\}$, using an iterative procedure, so that $min\{N_x^n\}$ is large. Finally, we choose $\{h_1, m, \theta\}$ so that $min\{N_x^n\}$ is large. The iterative process continues to obtain the largest possible price footprint in each price interval.

3.3.3 Small Sample Size Bias Adjustment

As explained earlier, if $E_{t_i}(S_{t_{i+1}}) > S_{t_i}$ (typical for stocks that require a risk premium) and $\sigma^2(x)$ is increasing in x (necessary for a price bubble), then the estimator will be upward biased. The proof of this assertion is contained in the appendix. Using a linear approximation for the volatility function, the appendix shows that an unbiased estimator for $\sigma^2(x)$ is

$$\hat{\sigma}^2(x) = \frac{V_n(x)}{1 + \frac{2}{x}\sum_{i=1}^{n} \frac{1_{\{|S_{t_i} - x| < h_n\}}(S_{t_i} - x)}{N_x^n}}. \tag{9}$$

As given, we see that if on average $S_{t_i} > x$, the bias adjustment reduces the estimated variance. That is, the unadjusted variance estimate is biased upward. And, the smaller the price level x, the larger is the variance adjustment.

3.3.4 The Probability of a Bubble given an Inconclusive Result

The hypothesis testing method, based on upper and lower bounds for the volatility function, is conservative. As such, even after all the extensions just discussed, there is a region where the hypothesis testing is inconclusive. To obtain a point estimate of the probability of a bubble in the inconclusive region, we take a Bayesian perspective and view β_k for $k \in \{u, l\}$ as random variables.

We assume that the posterior distribution for these random variables is given by the sampling distributions for the estimates $(\hat{\beta}_k)$ based on the standard errors $(\hat{\sigma}_{\beta_k})$, obtained from the regression of the upper and lower convex hulls in expression (7). Under the assumption, the appendix shows that a point estimate for the probability of a bubble is

$$\widehat{Prob}(bubble) = 1 - \frac{\left[\Phi\left(\frac{1-\hat{\beta}_l}{\hat{\sigma}^2_{\hat{\beta}_l}}\right) + \Phi\left(\frac{1-\hat{\beta}_u}{\hat{\sigma}^2_{\hat{\beta}_u}}\right)\right]}{2} \tag{10}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

3.3.5 Robustness Tests

The convex hull method developed by Choi and Jarrow (2022) winsorizes the maximum and minimum estimated volatility points with their second highest and lowest volatility points, respectively. In the new method, we modify the maximum and minimum estimated volatility points based on the sampling distribution of the standard error, given in expression (6) above. The estimate of the variance's standard error is

$$\sqrt{2\frac{V_n^2(x)}{N_x^n}} = \sqrt{2}\frac{V_n(x)}{\sqrt{N_x^n}},$$

which increases with $V_n(x)$. The robustness test procedure is as follows.

- Replace the maximum value $V_n^*(x)$ with $V_n^*(x) - \kappa\sqrt{2}\frac{V_n^*(x)}{\sqrt{N_x^n}}$ with $\kappa > 0$ a constant, and
- Replace the minimum value $V_n^{\#}(x)$ with $V_n^{\#}(x) + \kappa\sqrt{2}\frac{V_n^{\#}(x)}{\sqrt{N_x^n}}$.
- Given the sampling distribution in the previous section, these κ determine the probability that the variance exceeds the adjusted variance estimator, i.e.

$$P\left\{\sigma^2(x) > V_n^{\#}(x) + \kappa\sqrt{2}\frac{V_n^{\#}(x)}{\sqrt{N_x^n}}\right\} = 1 - \Phi\left\{k\right\}$$

  where Φ is the standard $(0, 1)$ cumulative normal distribution function.

- For various choices of κ (i.e., 0.05,0.10,0.15,0.20), evaluate the bubble test results.

Choosing various levels of κ provides a sensitivity analysis that indicates the impact of the likelihood of the error in the estimated variance on the hypothesis test for a bubble.

3.3.6 Heteroskedasticity and Autocorrelation

When fitting the linear regression in expression (7) to the volatility function's lower and convex hulls, the regression residuals may exhibit heteroskedasticity or autocorrelation. To address this possibility, we use the Newey-West variance estimator (White, 1980; Newey and West, 1987) to generate consistent standard errors.

4 Simulation

To validate the bubble testing methodology, we construct a hypothetical market, where the risky asset price follows a constant elasticity of variance (CEV) process under an equivalent local martingale measure $\mathbb{Q}$, given by

$$\frac{dS(t)}{S(t)} = \alpha S_t^{\beta-1} dW_t \tag{11}$$

where W_t is a standard Brownian motion, initialized at $W_0 = 1$, with $\alpha \geq 0$ and β constants.

We fix $\alpha = 0.3$ per year (typical for a stock market index), and construct two different markets, one with a bubble ($\beta = 1.5$) and one without ($\beta = 0.5$). Using a Euler scheme, we discretize the continuous-time evolution as

$$S(t_{i+1}) = S(t_i) \exp\left[-\frac{1}{2}\left(\alpha S_t^{\beta-1}\right)^2 (t_{i+1} - t_i) + \left(\alpha S_t^{\beta-1}\right) \sqrt{t_{i+1} - t_i} Z_{i+1}\right] \tag{12}$$

where the time interval corresponds to one day, i.e. $(t_{i+1} - t_i) = \frac{1}{365}$.

We simulate a path for the risky asset's price over 3 years (1095 days), and then apply our bubble testing methodology to see if it correctly identifies the bubble and no-bubble markets. Given the randomness in the simulation itself, we perform this exercise 10,000 times for each market.

It is well-known (see Chapter 4, Jaeckel (2002) and Chapter 6, Glasserman (2002)) that the Euler scheme has an approximation error that converges to zero as the time intervals uniformly converge to zero. In our context, this approximation error can be viewed as analogous to the existence of market micro-structure noise (e.g., bid/ask prices) in actual realized market prices. Nonetheless, given our small time step, this discretizing of the price process' path is not large enough to change a simulated bubble market into a no-bubble market and conversely.[8]

Figure 1 (Figure 2) consists of four plots. The first two plots present the simulated price paths and price distribution for $\beta = .5$ and $\beta = 1.5$. First, we note that price

[8]This is not the case, however, with $\beta = 1$, which is on the boundary of the no bubble market. In a simulation, this approximation error could transform the no bubble market into one containing a bubble. Because the purpose of the simulation is to see if our statistical methods can identify a bubble in a controlled experiment, the case $\beta = 1$ eliminates the control. Consequently, we do not simulate a hypothetical market with $\beta = 1$.

paths are more disperse with $\beta = 1.5$. This visually confirms that the CEV process with $\beta = 1.5$ yields more volatility at higher realized price levels. The middle figure plots the number of assets that accept and reject the first (second) test's null hypothesis. The last figure plots the percentage of assets that fall in each posterior bubble probability bucket given the results are inconclusive for $\beta = .5$ and $\beta = 1.5$. When evaluating these results, recall that our hypothesis testing is conservative, with a 95% significance level. Consequently, we would expect to see around 5% of the no bubble markets categorized as bubbles in step 1, and the same for the hypothesis testing of bubble markets in step 2.

Recall that step 1 tests the null hypothesis of no bubble. For $\beta = .5$ (no bubble), only 0.04% of the simulated paths are misclassified as bubbles. Exactly as expected. For $\beta = 1.5$ (bubbles), 71% are classified correctly as bubbles. Remember, however, that this hypothesis test is based on a lower bound for the volatility function, so it is conservative, and it will reject fewer no-bubbles than if the "true" volatility function had been utilized.

For those paths that are not rejected in step 1, considering type II error, step 2 tests the null hypothesis of a bubble. For the 0.96% of the simulations (9544 paths) accepting the null hypothesis of no bubble for $\beta = .5$, 85% reject the hypothesis of a bubble in step 2. This is strong evidence supporting the validity of the testing methodology. We emphasize that this is a conservative test using the upper bound for the volatility function at the 95% significance level, so it will reject fewer bubbles than if the "true" volatility function had been utilized. Next, in step 2, for the 29% of the simulations (2832 paths) for $\beta = 1.5$ that do not reject no-bubble in step 1, only 5.4% reject a the path as having a bubble. This is what we expect since we are using a 95% significance level in this hypothesis test.

Finally, for all those paths that result in an inconclusive determination after both steps 1 and 2, for $\beta = .5$, approximately 45% (39%) exhibit a posterior bubble probability of more than 90% (less than 10%). And, for $\beta = 1.5$, approximately 41% (49%) exhibit a posterior bubble probability of more than 90% (less than 10%). When we add the additional 950 paths for $\beta = 1.5$ that are inconclusive, but with a posterior probability of a bubble greater than 90%, a combined total of 81% of the 10,000 simulation paths for $\beta = 1.5$ exhibit price bubbles. In conjunction, these simulation results provide strong evidence in support of the methodology's ability to identify asset price bubbles.

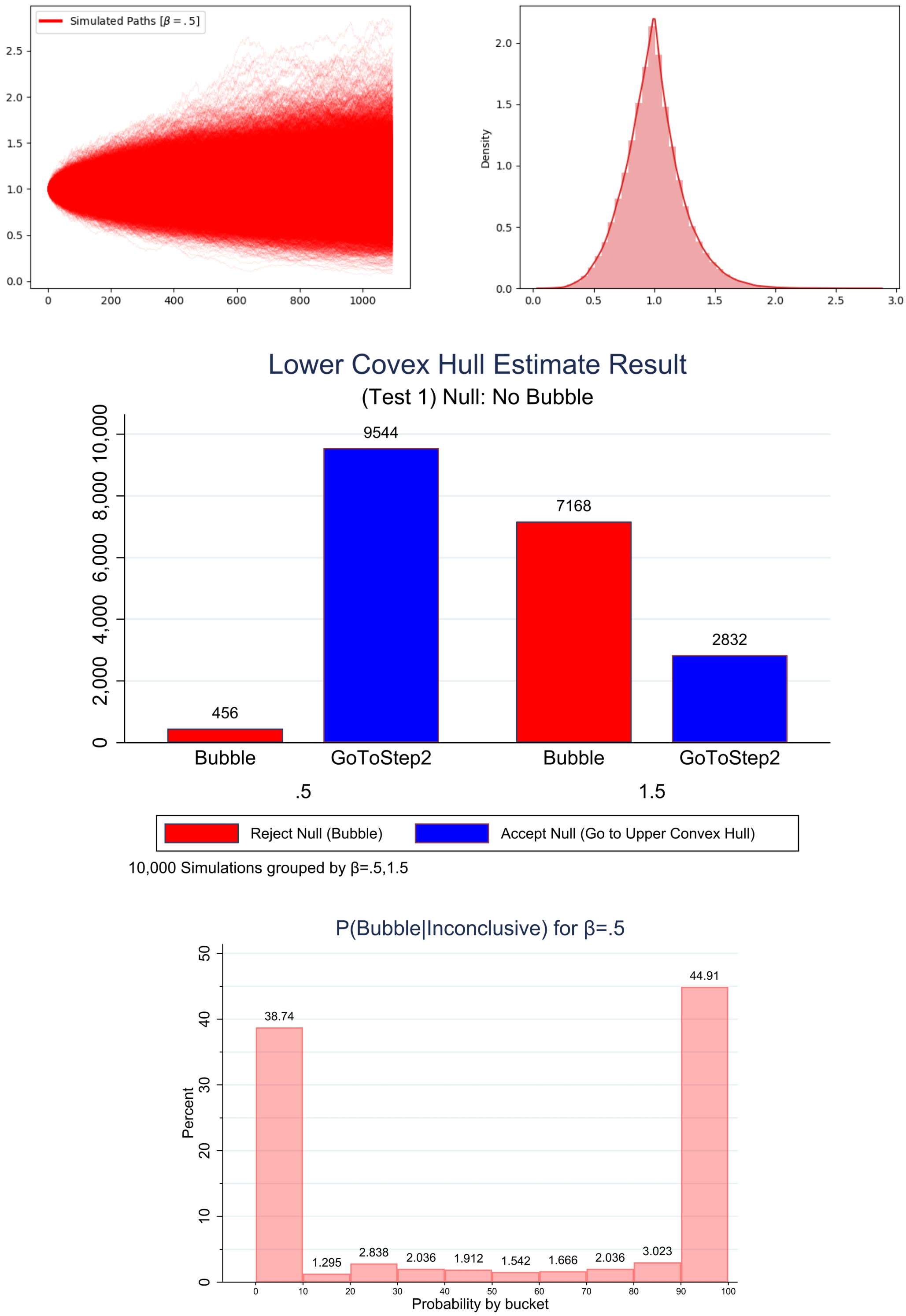

Figure 1. Simulated Paths. The figure graphs the simulated price paths & distributions of a Constant Elasticity of Variance (CEV) process for $\beta \in \{.5, 1.5\}$ with 10,000 iterations. The first two graphs represent test 1 results pertaining to the lower convex hull. The last graph plots the posterior probability distribution of the CEV process with $\beta = .5$ exhibiting a bubble.

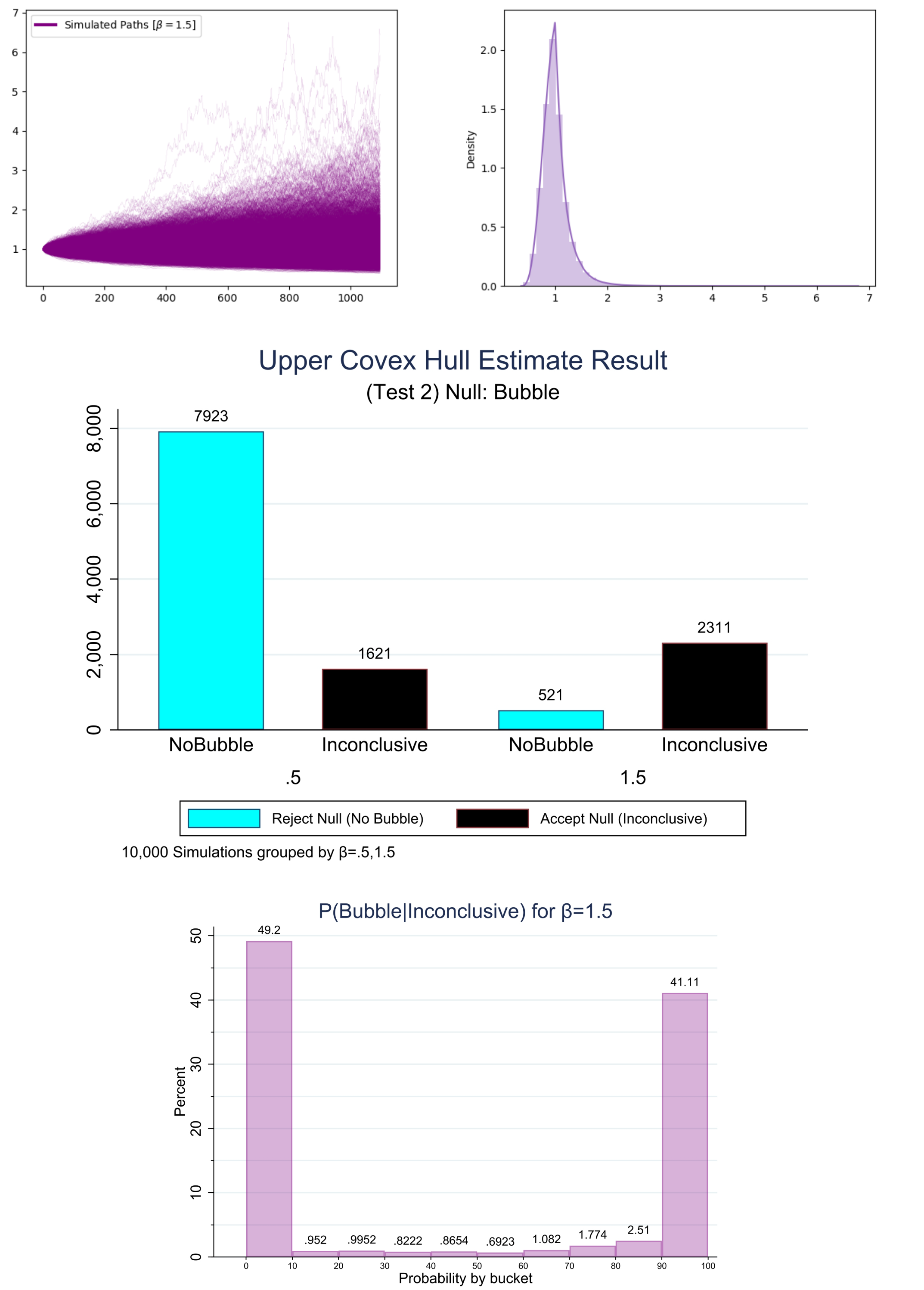

Figure 2. Simulated Paths. The figure graphs the simulated price paths & distributions of a Constant Elasticity of Variance (CEV) process for $\beta \in \{.5, 1.5\}$ with 10,000 iterations. The first two graphs represent test 2 results pertaining to the upper convex hull. The last graph plots the posterior probability distribution of the CEV process with $\beta = 1.5$ exhibiting a bubble.

5 The Empirical Investigation

Our empirical investigation consists of three parts. In the first we examine daily closing prices of three US major indices (S&P 500, Dow Jones Industrial Average, Nasdaq) to see if they exhibit price bubbles. In the second, we examine five individual stocks and one cryptocurrency (case studies) to provide confidence that the procedure works well. Here, we examine three assets that are often alleged as containing bubbles (Bitcoin, Meta, NVIDIA) and three banks stocks that are not (JP Morgan, Bank of America, Wells Fargo).

The third looks at a recent event in Lyft's stock on February 14th 2024. On that day, Lyft announced an erroneous earnings projection stating that the firm expected to increase its margin by 500 basis points; the correct expected margin increase was 50 basis points, which the Lyft CEO corrected less than an hour following the initial release. However, the Lyft stock traded 67% higher than the previous day's closing price in the interim. We hypothesize that this surge in the Lyft stock's price due to an erroneous earnings projection is a bubble. We apply our methodology to the pre-announcement (March 1, 2023 – February 13, 2024) and post-announcement (March 1, 2023 – March 4, 2024) periods to test this hypothesis.

5.1 Data

We collect the data from the LSEG Data & Analytics' Workspace platform. The sample data consists of ten assets: three indices (S&P 500, DJI, Nasdaq), three assets that may exhibit bubbles (Bitcoin, Meta, NVIDIA), three assets that may not (JP Morgan, Bank of America, and Wells Fargo), and Lyft. The sample period is from March 1st 2023 to March 4th 2023 consisting of 264 trading days. Given our statistical methodology assumes a normalized price process (see Section 3.1), we use the Secured Overnight Financing Rate (SOFR) over the sample period to compute the normalized closing daily asset prices. For the money market account, let the time t money market account value be B_t where r_t is the default-free spot rate (per year). In symbols, we compute

$$B_t = e^{\sum_{s=0}^{t-1} r_s \left(\frac{1}{365}\right)},$$

where the normalized asset price is $\frac{S_t}{B_t}$.

TABLE I. Descriptive Statistics of the Regression Sample

The table presents descriptive statistics of the sample's ten assets. The sample consists of the daily closing prices of three major US equity indices and seven stocks over 264 trading days from March 1 2023 to March 4 2024. Bitcoin (BTC) is in 1000 US Dollars. Columns (1)–(10) are based on quoted prices without the money market account normalization. Columns (1A)–(10A) are based on the normalization using the Secured Overnight Financing Rate Data (SOFR) over the sample period.

	SP500	DJI	NASDAQ	BTC	Meta	NVIDIA	JPM	BoA	WellsFargo	Lyft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mean	4353.28	34276.71	13398.33	0.03	299.64	432.16	147.30	29.37	42.61	11.04
Std. Dev	250.74	1575.07	1060.23	0.01	67.69	122.37	13.04	2.24	3.85	1.95
Min	3851.44	31648.80	11127.79	0.01	173.42	226.98	124.63	24.57	36.15	7.93
Max	4949.88	37732.97	15681.86	0.05	484.00	821.19	179.85	34.15	53.77	18.37
N	264	264	264	264	264	264	264	264	264	264
	(1A)	(2A)	(3A)	(4A)	(5A)	(6A)	(7A)	(8A)	(9A)	(10A)
Mean	4272.38	33644.55	13146.41	0.03	293.57	423.19	144.52	28.83	41.82	10.83
Std. Dev	207.98	1278.42	919.20	0.01	63.30	115.68	11.44	2.08	3.47	1.82
Min	3847.12	30898.25	11116.70	0.01	173.42	226.98	124.35	23.99	36.07	7.86
Max	4769.50	36384.40	15110.39	0.05	466.36	791.15	173.27	34.14	51.80	17.72
N	264.00	264.00	264.00	264.00	264.00	264.00	264.00	264.00	264.00	264.00

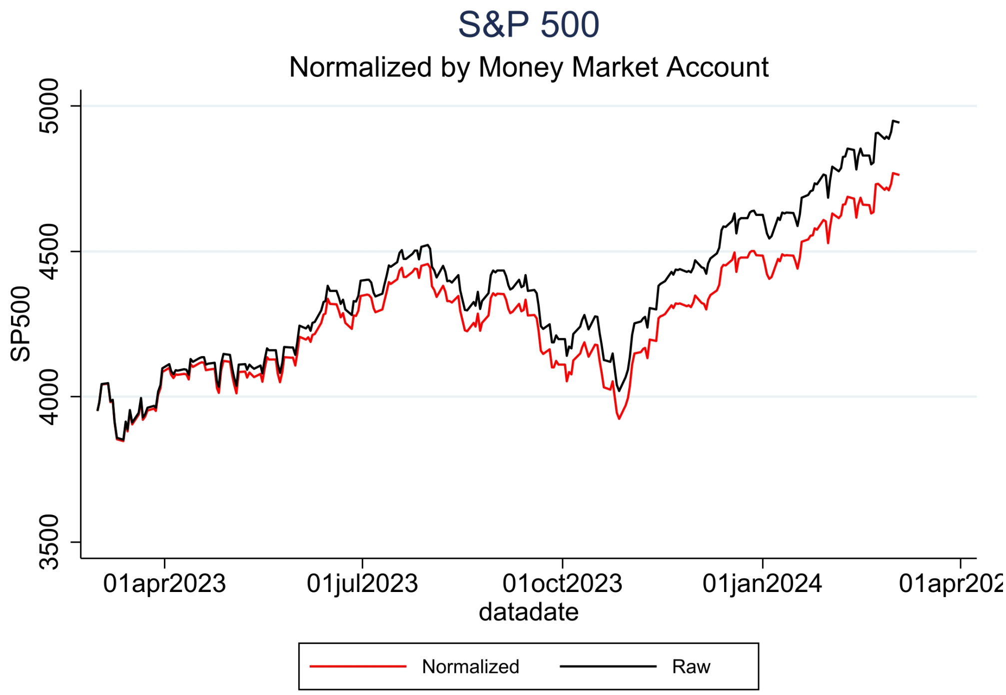

Figure 3. Historical Price of S&P 500 from March 1 2023 to March 4 2024. This figure plots the quoted price (raw) and normalized price (by the money market account) of S&P 500.

Table I provides descriptive statistics of the ten assets' quoted (raw) and normalized prices. Figure 3 plots a juxtaposition of the S&P 500 paths for normalized (red) and raw (black) historical prices from March 1 2023 to March 4 2024. The price process normalized by the cost of borrowing cash overnight (collateralized by Treasury) trends lower as the cost has risen from 4.5% to 5.38% during the sample period.

5.2 The Baseline Model

The baseline model is described in Section 3.1. Given are risky asset price observations $\{S_{t_i}\}_i$ over a fixed horizon, $i \in \{1, 2, ..., T\}$, from which we generate a set of price and estimated volatility pairs $\{(S_j, \sigma(S_j))\}_j$ where j corresponds to the j^{th} bin of a price interval.[9] For these price-volatility pairs, we generate upper and lower convex hulls. For each convex hull, we fit the best power functions $\sigma_k(x) = \alpha_k x^{\beta_k}$ where $k = l$ and $k = u$ correspond to the lower and upper convex hull, respectively. We perform the ordinary least squares regression:

$$\ln(\sigma_k(S_j)) = \ln(\alpha_k) + \beta_k \ln(S_j) + \varepsilon_j. \tag{13}$$

First, we evaluate whether $\hat{\beta}_l$ exceeds 1+1.645$\hat{\sigma}_l$. If $\hat{\beta}_l$ exceeds this threshold, then we reject the null hypothesis (no bubble) and conclude that the asset has a bubble. Otherwise, we evaluate whether $\hat{\beta}_u$ is less than 1-1.645$\hat{\sigma}_u$. If $\hat{\beta}_u$ is less then this threshold, we reject the null hypothesis (bubble) and conclude the asset does not have a price bubble. Otherwise, our hypothesis testing is inconclusive. In this last case, we compute the posterior probability of a bubble given an inconclusive result.

5.3 The Results

This section provides the results for all the assets selected for investigation.

5.3.1 The Baseline Model

We apply the methodology to three major U.S. equity indices: the S&P 500, the Dow Jones Industrial Average, and the Nasdaq. Table II documents our findings. Columns (1), (3), and (5) report the coefficient estimates, 95% confidence interval thresholds, results of the hypothesis tests, and the posterior probability of a bubble for the S&P 500,

[9]Recall that for each price interval, we obtain a pair of price and volatility estimate.

Dow Jones, and Nasdaq. The odd columns provide the numbers for the lower convex hulls and the even columns report the numbers for the upper convex hulls.

TABLE II. Baseline Regression Results: Index

The table reports the coefficient estimates of the regression in (13) for the lower (β_l) and upper (β_u) convex hulls of three major US equity indices from March 2023 to March 2024 using daily closing prices. Test 1 evaluates the null hypothesis of no bubble. Test 2 evaluates the null hypothesis of a bubble. Inconclusive means both tests accept the null hypothesis. In the inconclusive case, the probability of a bubble is given using a Bayes' posterior distribution. All standard errors are computed with the Newey–West adjustment. $^{*}p$ <0.1, $^{**}p$ <0.05, $^{***}p$ <0.01.

	US Stock Market					
	S&P 500		DJI		NASDAQ	
	(1)	(2)	(3)	(4)	(5)	(6)
	$\hat{\beta}_l$	$\hat{\beta}_u$	$\hat{\beta}_l$	$\hat{\beta}_u$	$\hat{\beta}_l$	$\hat{\beta}_u$
Bubble Coefficient	-4.709***	-2.634*	-2.082***	1.672	2.316***	1.963***
	(0.306)	(1.564)	(0.237)	(2.730)	(0.240)	(0.822)
95 CI Threshold	1.503	-1.573	1.389	-3.491	1.395	-0.352
(Test 1) H_0: No Bubble	Accept		Accept		Reject	
(Test 2) H_0: Bubble		Reject		Accept		
Result	Inconclusive	No Bubble	Inconclusive	Inconclusive	Bubble	NA
P(Bubble|Inconclusive)	NA		0.35%		NA	
R^2	0.973	0.265	0.893	0.071	0.937	0.489
N	264	264	264	264	264	264

For the S&P 500 and Dow Jones indices, we accept the hypothesis of no bubble (test 1). For the Nasdaq, we reject the hypothesis of no bubble at the 1% level, indicating that it exhibits a price bubble.

Next, for the the S&P 500 and Dow Jones indices, we investigate test 2. The S&P 500 rejects the hypothesis of a bubble at the 10% significance level, indicating it does not exhibit a bubble. For the Dow Jones index, we cannot reject the hypothesis of a bubble. Hence, the Dow Jones index yields an inconclusive result. Here, we compute the the posterior probability of bubble, which is 0.35%. Hence, the Dow Jones index is unlikely to exhibit a price bubble.

5.3.2 The Case Studies

Here, we apply the methodology to two sets of assets: (i) those alleged to have bubbles and (ii) those alleged to not. The first set consists of Bitcoin, Meta, and NVIDIA, the second set consists of JP Morgan, Bank of America, and Wells Fargo.

TABLE III. Baseline Regression Results: Alleged Bubble Assets

The table reports the coefficient estimates of the regression in (13) for the lower (β_l) and upper (β_u) convex hulls of two U.S. stocks and Bitcoin from March 2023 to March 2024 using daily closing prices. Test 1 evaluates the null hypothesis of no bubble. Test 2 evaluates the null hypothesis of a bubble. Inconclusive means both tests accept the null hypothesis. In the inconclusive case, the probability of a bubble is given using a Bayes' posterior distribution. All standard errors are computed with the Newey–West adjustment. $^{*}p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$.

	Alleged Bubbles					
	Bitcoin		Meta		NVIDIA	
	(1)	(2)	(3)	(4)	(5)	(6)
	$\hat{\beta}_l$	$\hat{\beta}_u$	$\hat{\beta}_l$	$\hat{\beta}_u$	$\hat{\beta}_l$	$\hat{\beta}_u$
Bubble Estimate	1.360***	1.659***	0.361***	0.366*	1.284***	1.235***
	(0.533)	(0.080)	(0.109)	(0.238)	(0.140)	(0.253)
95 CI Threshold	1.877	0.868	1.179	0.609	1.230	0.584
(Test 1) H_0: No Bubble	Accept		Accept		Reject	
(Test 2) H_0: Bubble		Accept		Reject		
Result	Inconclusive	Inconclusive	Inconclusive	No Bubble	Bubble	NA
P(Bubble|Inconclusive)	100%		NA		NA	
R^2	0.503	0.986	0.642	0.283	0.932	0.806
N	264	264	264	264	264	264

TABLE IV. Baseline Regression Results: Alleged No Bubbles

The table reports the coefficient estimates of the regression in (13) for the lower (β_l) and upper (β_u) convex hulls of three U.S. stocks from March 2023 to March 2024 using daily closing prices. Test 1 evaluates the null hypothesis of no bubble. Test 2 evaluates the null hypothesis of a bubble. Inconclusive means both tests accept the null hypothesis. In the inconclusive case, the probability of a bubble is given using a Bayes' posterior distribution. All standard errors are computed with the Newey–West adjustment. $^{*}p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$.

	Alleged No Bubble Assets					
	JP Morgan		Bank of America		Wells Fargo	
	(1)	(2)	(3)	(4)	(5)	(6)
	$\hat{\beta}_l$	$\hat{\beta}_u$	$\hat{\beta}_l$	$\hat{\beta}_u$	$\hat{\beta}_l$	$\hat{\beta}_u$
Bubble Estimate	-2.473***	-2.594***	1.692***	1.741***	-1.001***	-0.850*
	(0.387)	(0.225)	(0.364)	(0.597)	(0.199)	(0.672)
95 CI Threshold	1.636	0.630	1.599	0.018	1.328	-0.106
(Test 1) H_0: No Bubble	Accept		Reject		Accept	
(Test 2) H_0: Bubble		Reject				Reject
Result	Inconclusive	No Bubble	Bubble	NA	Inconclusive	No Bubble
P(Bubble|Inconclusive)	NA		NA		NA	
R^2	0.876	0.959	0.660	0.573	0.791	0.224
N	264	264	264	264	264	264

Table III and Table IV document our findings. For Bitcoin and Meta, test 1 accepts the hypothesis of no bubble, while NVIDIA rejects the hypothesis of no bubble at the 1% significance level. Hence, we conclude NVIDIA has a bubble.

For test 2, we cannot reject the hypothesis of a bubble for Bitcoin, although it is rejected for Meta at the 10% significance level. Hence, this implies Meta has no bubble, and for Bitcoin the testing is inconclusive. Then, we compute the posterior probability of a bubble for Bitcoin, which is 100%. Hence, we conclude that Bitcoin exhibits a bubble.

For the bank stocks, we cannot reject the hypothesis of no bubble for both JP Morgan and Wells Fargo. However, we reject the hypothesis of no bubble for the Bank of America at the 1% significance level. Hence, we conclude that Bank of America has a bubble. Applying test 2 to JP Morgan and Wells Fargo, we reject the hypothesis of a bubble at the 1% and 10% significant levels, respectively, indicating that these stocks do not exhibit bubbles.

5.4 Ex-post Analysis of Lyft

On February 14$^{\text{th}}$ 2024, Lyft issued an earnings projection stating that its margins would increase by 500 basis points. Less than an hour after the release, the Lyft CEO corrected the typo and stated that the projected estimate is 50 basis points. In the interim, the company's stock surged as much as 67% based on the erroneous initial report. From the perspective of the local martingale theory, the Lyft episode provides an ideal ex-post laboratory to test our new methodology.

We perform two regressions. First, we apply the new methodology to Lyft stock prices from March 1$^{\text{st}}$ 2023 to February 13$^{\text{th}}$ 2024, a day before the erroneous earnings projection. Second, we include the post-earnings-announcement date which is after February 14$^{\text{th}}$ 2024. Intuitively, if there was no fundamental change to Lyft as a firm except the incorrect earnings statement, we hypothesize that our refined technology can detect a post-announcement bubble.

Table V documents our findings. In the pre-announcement period, test 1 cannot reject the hypothesis of no bubble. And, test 2 cannot reject the hypothesis of a bubble. Hence, these tests are inconclusive. Then, computing the posterior probability of a bubble, we see that it is 0%. Hence, we conclude that there is no bubble in the pre-announcement period. In the post-announcement period, test 1 rejects the hypothesis of no bubble at the 1% significance level, yielding the conclusion that there was a bubble. In conjunction, we see that the bubble testing methodology confirms our hypothesis

TABLE V. Ex-post Analysis: Lyft

The table reports the coefficient estimates of the regression in (13) for the lower (β_l) and upper (β_u) convex hulls of Lyft stock. The test examines two periods: (i) before the erroneous earnings projection statement date (March 1, 2023 – Feb 13, 2024) and (ii) after the CEO correction date (March 1, 2023 – March 4, 2024). Test 1 evaluates the null hypothesis of no bubble. Test 2 evaluates the null hypothesis of a bubble. Inconclusive means both tests accept the null hypothesis. In the inconclusive case, the probability of a bubble is given using a Bayes' posterior distribution. All standard errors are computed with the Newey–West adjustment. $^{*}p$ <0.1, $^{**}p$ <0.05, $^{***}p$ <0.01.

	Lyft			
	Pre-Announcement		Post-Annoucement	
	March 1, 2023 - Feb 13, 2024		March 1, 2023 - March 4, 2024	
	(1)	(2)	(3)	(4)
	$\hat{\beta}_l$	$\hat{\beta}_u$	$\hat{\beta}_l$	$\hat{\beta}_u$
Bubble Estimate	0.947***	0.727***	1.590***	1.530***
	(0.002)	(0.368)	(0.199)	(0.337)
95 CI Threshold	1.003	0.394	1.327	0.445
(Test 1) H_0: No Bubble	Accept		Reject	
(Test 2) H_0: Bubble		Accept		
Result	Inconclusive	Inconclusive	Bubble	NA
P(Bubble|Inconclusive)	0%		NA	
R^2	0.999	0.401	0.878	0.785
N	250	250	264	264

about the Lyft stock price surge after the false earnings announcement is a bubble.

6 Robustness Tests

We perform a robustness check on the baseline results by applying an error adjustment to the maximum and minimum estimated volatilities. We investigate whether our results are sensitive to changes in these two extreme volatility estimates. In this procedure, we reduce the largest volatility estimate by κ times the standard error, and we increase the smallest volatility estimate by κ times the standard error. The larger the κ, the larger the probability distribution of the estimated volatility that exceeds the adjusted volatility estimate.

Figure 4 demonstrates the robustness check procedure for κ values .05, .15, .1, and .2 applied to the lower convex hull of the S&P 500 price-volatility pairs for the sample period. As the values increase, it visually shows the largest and smallest volatility estimates are adjusted downward and upward, respectively, based on the standard error

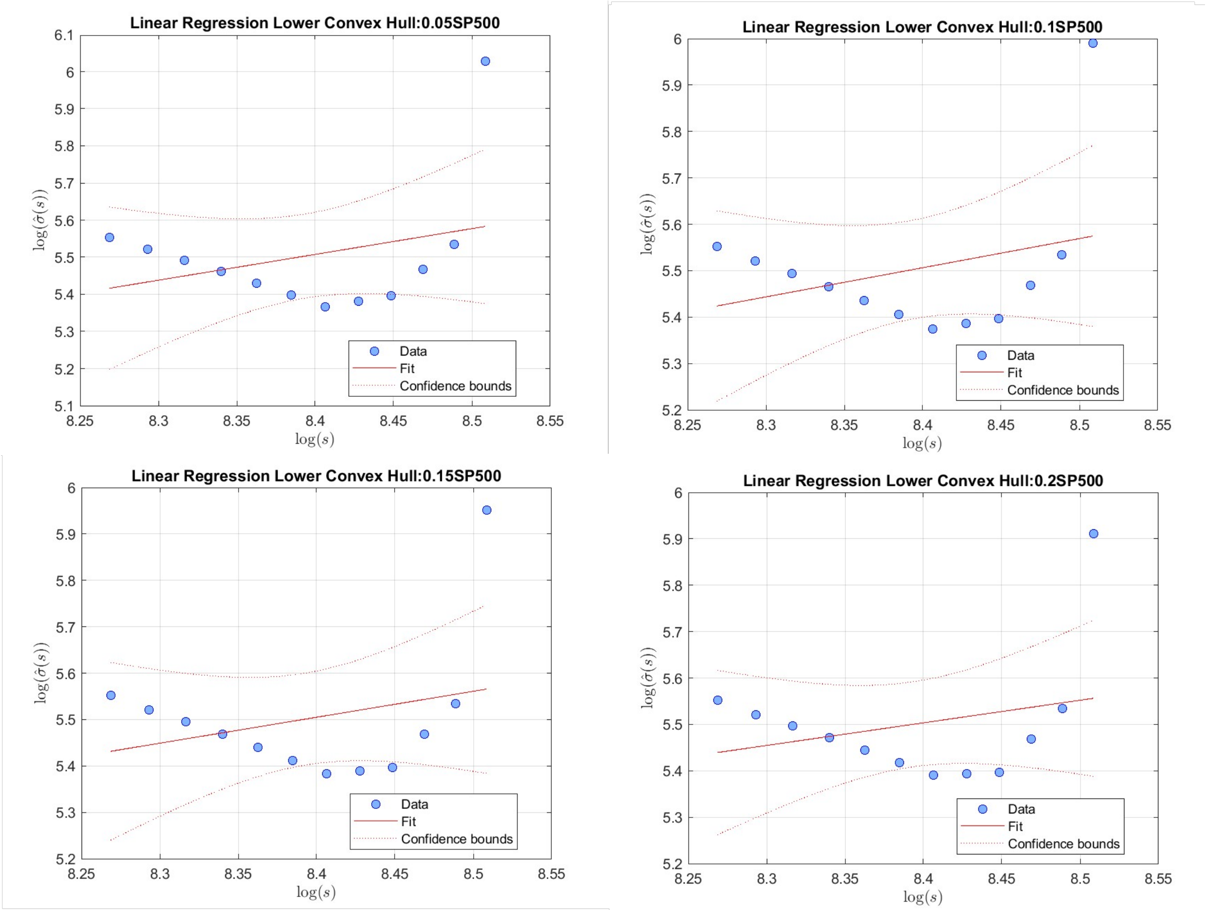

Figure 4. Outlier Correction for S&P 500. The figure plots the adjustment of the maximum and minimum estimated volatilities for $\kappa \in \{.05, .15, .1, .2\}$

distribution. Intuitively, if a subject asset exhibits a price bubble, then adjusting the highest volatility estimate potentially alters the fitted line of the interpolated convex hull to be less explosive.

We re-estimate the regressions for all the assets after applying the robustness adjustments. Table VI, Table VII, Table VIII, and Table IX document our findings. Column (1) provides the adjustments we make in κ to treat the outlier points of the convex hulls. Column (2) provides the probability that the variance is greater than the adjusted estimator. Columns (4) and (5) provide the test results of the null hypotheses. Column (6) documents the results. Column (7) provides the posterior probability of a bubble given an inconclusive result from the hypothesis testing.

TABLE VI. Robustness Results: Baseline Index

The table presents an executive summary of the baseline regression results when the maximum and minimum volatility estimates are adjusted downward and upward respectively based on the standard error distribution method developed in Section 3.3.5. The adjustment parameter κ take values .05, .1, .15, and .2. The κ determines the probability that the variance exceeds the adjusted variance estimator denoted as $P(\Phi(0,1) > |\kappa|)$. Test 1 evaluates the null hypothesis of no bubble. Test 2 evaluates the null hypothesis of a bubble. Inconclusive means both tests accept the null hypothesis. In the inconclusive case, the probability of a bubble is given using a Bayes' posterior distribution.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
κ	$P(\Phi(0,1) > \lvert\kappa\rvert)$	Asset	(Test 1) H_0: No Bubble	(Test 2) H_0: Bubble	Result	P(Bubble|Inconclusive)
0.05	0.52	SP500	Accept	Reject	No Bubble	NA
0.05	0.52	DJI	Accept	Accept	Inconclusive	0.00%
0.05	0.52	NASDAQ	Reject	Accept	Bubble	NA
0.1	0.54	SP500	Accept	Reject	No Bubble	NA
0.1	0.54	DJI	Accept	Accept	Inconclusive	0.00%
0.1	0.54	NASDAQ	Reject	Accept	Bubble	NA
0.15	0.56	SP500	Accept	Reject	No Bubble	NA
0.15	0.56	DJI	Accept	Accept	Inconclusive	0.00%
0.15	0.56	NASDAQ	Reject	Accept	Bubble	NA
0.2	0.58	SP500	Accept	Reject	No Bubble	NA
0.2	0.58	DJI	Accept	Accept	Inconclusive	0.01%
0.2	0.58	NASDAQ	Reject	Accept	Bubble	NA

TABLE VII. Robustness Results: Allegedly Bubble

The table presents an executive summary of the bubbly asset regression results when the maximum and minimum volatility estimates are adjusted downward and upward respectively based on the standard error distribution method developed in Section 3.3.5. The adjustment parameter κ take values .05, .1, .15, and .2. The κ determines the probability that the variance exceeds the adjusted variance estimator denoted as $P(\Phi(0,1) > |\kappa|)$. Test 1 evaluates the null hypothesis of no bubble. Test 2 evaluates the null hypothesis of a bubble. Inconclusive means both tests accept the null hypothesis. In the inconclusive case, the probability of a bubble is given using a Bayes' posterior distribution.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
κ	$P(\Phi(0,1) > \lvert\kappa\rvert)$	Asset	(Test 1) H_0: No Bubble	(Test 2) H_0: Bubble	Result	P(Bubble|Inconclusive)
0.05	0.52	BTC	Accept	Accept	Inconclusive	100.00%
0.05	0.52	Meta	Accept	Reject	No Bubble	NA
0.05	0.52	NVIDIA	Reject	Accept	Bubble	NA
0.1	0.54	BTC	Accept	Accept	Inconclusive	100.00%
0.1	0.54	Meta	Accept	Reject	No Bubble	NA
0.1	0.54	NVIDIA	Reject	Accept	Bubble	NA
0.15	0.56	BTC	Accept	Accept	Inconclusive	100.00%
0.15	0.56	Meta	Accept	Reject	No Bubble	NA
0.15	0.56	NVIDIA	Reject	Accept	Bubble	NA
0.2	0.58	BTC	Accept	Accept	Inconclusive	100.00%
0.2	0.58	Meta	Accept	Reject	No Bubble	NA
0.2	0.58	NVIDIA	Reject	Accept	Bubble	NA

TABLE VIII. Robustness Results: Allegedly No Bubble

The table presents an executive summary of the non-bubbly asset regression results when the maximum and minimum volatility estimates are adjusted downward and upward respectively based on the standard error distribution method developed in Section 3.3.5. The adjustment parameter κ take values .05, .1, .15, and .2. The κ determines the probability that the variance exceeds the adjusted variance estimator denoted as $P(\Phi(0,1) > |\kappa|)$. Test 1 evaluates the null hypothesis of no bubble. Test 2 evaluates the null hypothesis of a bubble. Inconclusive means both tests accept the null hypothesis. In the inconclusive case, the probability of a bubble is given using a Bayes' posterior distribution.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
κ	$P(\Phi(0,1) > \|\kappa\|)$	Asset	(Test 1) H_0: No Bubble	(Test 2) H_0: Bubble	Result	P(Bubble|Inconclusive)
0.05	0.52	JPM	Accept	Reject	No Bubble	NA
0.05	0.52	BoA	Reject	Accept	Bubble	NA
0.05	0.52	WF	Accept	Reject	No Bubble	NA
0.1	0.54	JPM	Accept	Reject	No Bubble	NA
0.1	0.54	BoA	Accept	Accept	Inconclusive	99.00%
0.1	0.54	WF	Accept	Reject	No Bubble	NA
0.15	0.56	JPM	Accept	Reject	No Bubble	NA
0.15	0.56	BoA	Accept	Accept	Inconclusive	98.62%
0.15	0.56	WF	Accept	Reject	No Bubble	NA
0.2	0.58	JPM	Accept	Reject	No Bubble	NA
0.2	0.58	BoA	Accept	Accept	Inconclusive	98.15%
0.2	0.58	WF	Accept	Reject	No Bubble	NA

TABLE IX. Robustness Results: Lyft

The table presents an executive summary of the Lyft regression results (post- and pre-announcement) when the maximum and minimum volatility estimates are adjusted downward and upward respectively based on the standard error distribution method developed in Section 3.3.5. The adjustment parameter κ take values .05, .1, .15, and .2. The κ determines the probability that the variance exceeds the adjusted variance estimator denoted as $P(\Phi(0,1) > |\kappa|)$. Test 1 evaluates the null hypothesis of no bubble. Test 2 evaluates the null hypothesis of a bubble. Inconclusive means both tests accept the null hypothesis. In the inconclusive case, the probability of a bubble is given using a Bayes' posterior distribution.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Period	κ	$P(\Phi(0,1) > \|\kappa\|)$	(Test 1) H_0: No Bubble	(Test 2) H_0: Bubble	Result	P(Bubble|Inconclusive)
Post-announcement	0.05	0.52	Reject	Accept	Bubble	NA
	0.1	0.54	Reject	Accept	Bubble	NA
	0.15	0.56	Accept	Accept	Bubble	NA
	0.2	0.58	Accept	Accept	Bubble	NA
Pre-Annoucement	0.05	0.52	Accept	Accept	Inconclusive	0.00%
	0.1	0.54	Accept	Accept	Inconclusive	0.00%
	0.15	0.56	Accept	Accept	Inconclusive	0.00%
	0.2	0.58	Accept	Accept	Inconclusive	0.00%

These analyses indicate that our main regression results are robust even after adjusting for the outliers. In the baseline results, S&P 500 and Dow Jones do not have bubbles and Nasdaq does for all values of κ. In the case studies, Bitcoin's conditional probability of a bubble remains 100% for all κ values. Both Meta and NVIDIA have the identical hypothesis test results implying that Meta does not have a bubble and NVIDIA has a bubble. The results for JP Morgan and Wells Fargo remain without a bubble for all values of κ. Bank of America accepts the null hypothesis of no-bubble at $\kappa = .05$. However, it enters the inconclusive region where its posterior probability of a bubble remains close to 100% for the κ values .1, .15, and .2. These checks document that our initial results are robust to the adjustments of the largest and smallest volatility estimates.

7 Conclusion

In this paper, we provide five refinements to the existing bubble testing methodology based on the local martingale theory of bubbles. First, we enrich the existing method to encompass a large class of assets with cash flows. Second, we allow unequal time intervals and price level partitions to accommodate the reality of transaction times and trading days being uneven. Third, we identify the upward bias issue of the variance estimator and rectify the bias by making a small sample size adjustment. Fourth, we increase our understanding of the inconclusive results by taking a Bayesian view to compute the point estimate for the posterior probability of a bubble given an inconclusive result. Finally, we address the potential presence of heteroskedasticity and autocorrelation persistent in asset price data and the convex hull approach by implementing the Newey-West adjustments.

We apply the enhanced econometric procedure to the U.S. equity market. We show that the Nasdaq index, certain technology stocks, and Bitcoin exhibit bubbles. Conversely, we show that the S&P 500 and Dow Jones Industrial indices and certain bank stocks do not. In the ex-post analysis of Lyft stock, we find that it exhibits a bubble after the erroneous earnings announcement. In the pre-announcement results, Lyft's posterior probability of a bubble given inconclusiveness is zero.

To stress test our new methodology, we simulate a path of a constant elasticity of variance (CEV) price process in a market with a bubble and another market without a bubble over three years. In 10,000 simulated paths, we show that only a small fraction of the no-bubble market is misclassified as bubbles while the converse is true in the bubble-market. These results provide strong evidence in support of the methodology's ability to identify asset price bubbles.

Contrary to the extant literature's divergent views, the main implication of our paper is that a consistent test for asset price bubbles is empirically viable and testable based on the local martingale theory of bubbles. The refinements proposed in this paper can be enhanced and applied further to more diverse and general asset classes in subsequent research.

References

Aguinis, Herman, Ryan K. Gottfredson, and Harry Joo. 2013. "Best-Practice Recommendations for Defining, Identifying, and Handling Outliers." *Organizational Research Methods* 16 270 – 301.

Blanchard, Olivier. 1979. "Speculative bubbles, crashes and rational expectations." *Economics Letters* 3 (4): 387–389.

Brunnermeier, Markus K., and Martin Oehmke. 2013. "Chapter 18 - Bubbles, Financial Crises, and Systemic Risk." Volume 2. of Handbook of the Economics of Finance 1221–1288, Elsevier.

Chaim, Pedro, and Márcio P. Laurini. 2019. "Is Bitcoin a bubble?" *Physica A: Statistical Mechanics and its Applications* 517 222 – 232.

Choi, Soon Hyeok, and Robert A. Jarrow. 2022. "Applying The Local Martingale Theory of Bubbles to Cryptocurrencies." *International Journal of Theoretical and Applied Finance* 25 (03): 2250013.

Diba, Behzad T., and Herschel I. Grossman. 1988. "Explosive Rational Bubbles in Stock Prices?" *The American Economic Review* 78 (3): 520–530.

Fama, Eugene F. 2014. "Two Pillars of Asset Pricing." *American Economic Review* 104 (6): 1467–85.

Glasserman, Paul. 2002. *Monte Carlo Methods in Financial Engineering*. Springer.

Greenwood, Robin, Andrei Shleifer, and Yang You. 2019. "Bubbles for Fama." *Journal of Financial Economics* 131 (1): 20–43.

Grubbs, Frank E. 1969. "Procedures for Detecting Outlying Observations in Samples." *Technometrics* 11 (1): 1–21.

Jaeckel, Peter. 2002. *Monte Carlo Methods in Finance*. Wiley.

Jarrow, Robert. 2018. *Continuous-Time Asset Pricing Theory*. Springer International Publishing, , 1st edition.

Jarrow, Robert A., and Simon S. Kwok. 2021. "Inferring financial bubbles from option data." *Journal of Applied Econometrics* 36 (7): 1013–1046.

Jarrow, Robert A., Philip Protter, and Kazuhiro Shimbo. 2010. "ASSET PRICE BUBBLES IN INCOMPLETE MARKETS*." *Mathematical Finance* 20 (2): 145–185.

Jarrow, Robert, Younes Kchia, and Philip Protter. 2011b. "Is There a Bubble in LinkedIn's Stock Price?" *The Journal of Portfolio Management* 38 (1): 125–130.

Jarrow, Robert, Younes Kchia, and Phillip Protter. 2011a. "How to Detect an Asset Bubble." *SIAM Journal on Financial Mathematics* 2(1) 839–865.

Jarrow, Robert, Philip Protter, and Jaime San Martin. 2022. "Asset price bubbles: Invariance theorems."

Newey, Whitney K., and Kenneth D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (3): 703–708.

Obayashi, Yoshiki, Philip Protter, and Shihao Yang. 2017. "The Lifetime of a Financial Bubble." *Mathematics and Financial Economics* 11 (1): 45–62.

Phillips, Peter C. B., Shuping Shi, and Jun Yu. 2015. "TESTING FOR MULTIPLE BUBBLES: LIMIT THEORY OF REAL-TIME DETECTORS." *International Economic Review* 56 (4): 1079–1134.

Phillips, Peter C.B., and Shuping Shi. 2020. "Chapter 2 - Real time monitoring of asset markets: Bubbles and crises." In *Financial, Macro and Micro Econometrics Using R*, edited by Vinod, Hrishikesh D., and C.R. Rao Volume 42. of Handbook of Statistics 61–80, Elsevier.

Rocke, David M., and David L. Woodruff. 1996. "Identification of Outliers in Multivariate Data." *Journal of the American Statistical Association* 91 (435): 1047–1061.

Shiller, Robert J. 2016. *Irrational Exuberance*. Princeton: Princeton University Press.

Shiryaev, A. N., M. V. Zhitlukhin, and W. T. Ziemba. 2016. "When to Sell Apple and the NASDAQ? Trading Bubbles with a Stochastic Disorder Model." In *GREAT INVESTMENT IDEAS*, Chap. 11 231–248, World Scientific Publishing Co. Pte. Ltd..

White, Halbert. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48 (4): 817–838.

Tables & Figures

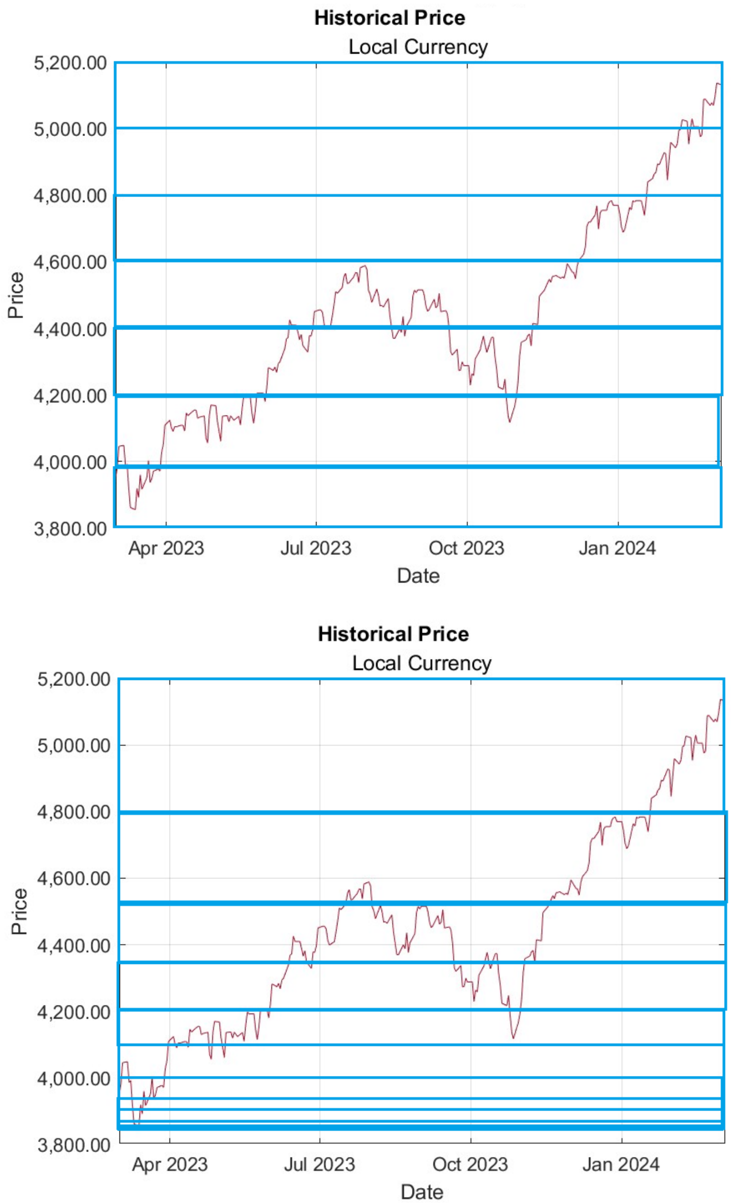

Figure 5. Partitioning Price Intervals as a Function of the Asset Price. The figure juxtaposes the existing method's even partitioning (above) against our new methodology's uneven partitioning (below) taking unequal trading days and transaction times into account.

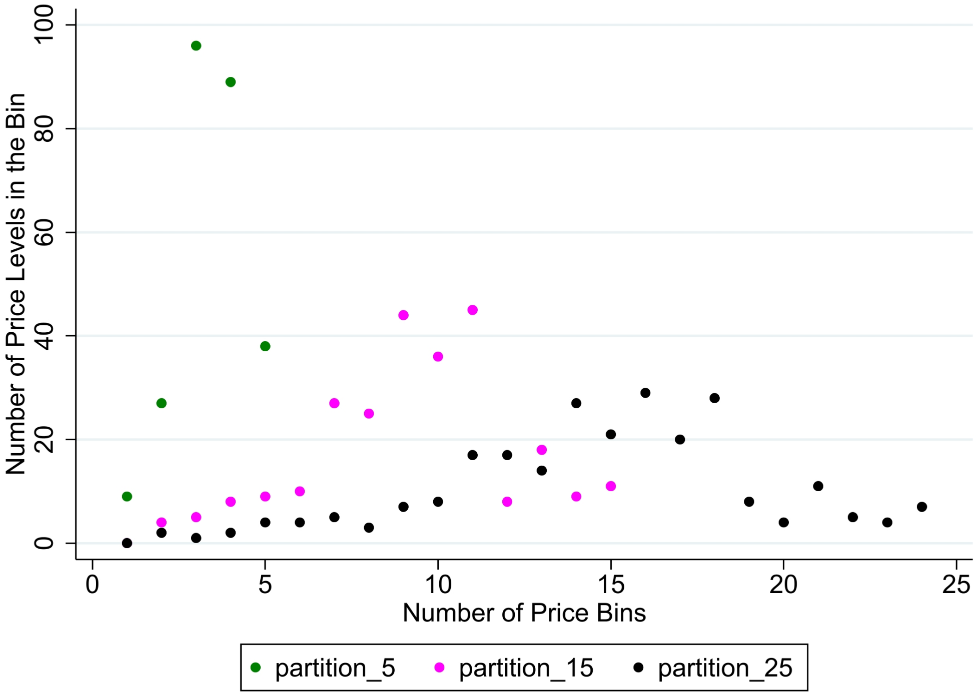

Figure 6. Number of Observed Prices by Un-equispaced Partition Bins. The figure plots the number of price levels (i.e., price footprint) we observe by the number of implemented partitions. The green, pink, and black dots correspond to the partitioning of five, 15, and 25 bins of S&P 500 over the sample period (March 2023 – March 2024).

A Appendix

A.1 Unequal Time Intervals Between Price Observations

Fixing n as the sample size, the time observations may be at transaction times, daily, or weekly. For this reason, it is important to adjust the variance estimator accordingly. For arbitrary discrete intervals $[t_i, t_{i-1}]$, the conditional variance is

$$var_{t_i}\left(S_{t_{i+1}} - S_{t_i}\right) = \sigma^2(S_{t_i})[t_{i+1} - t_i].$$

As an approximation, the estimator is:

$$var_{t_i}\left(S_{t_{i+1}} - S_{t_i}\right) \approx (S_{t_{i+1}} - S_{t_i})^2.$$

Combined, this implies

$$\sigma^2(S_{t_i}) \approx \frac{(S_{t_{i+1}} - S_{t_i})^2}{[t_{i+1} - t_i]}.$$

With multiple observations in small intervals around x, our estimator becomes

$$V_n(x) \approx \frac{\sum_{i=1}^{n} 1_{\{|S_{t_i} - x| < h_n\}} (S_{t_{i+1}} - S_{t_i})^2 \cdot \frac{1}{[t_{i+1} - t_i]}}{\sum_{i=1}^{n} 1_{\{|S_{t_i} - x| < h_n\}}} \tag{14}$$

Using daily observations, we have $[t_{i+1} - t_i] = \frac{1}{365}$.

A.2 The Standard Error

The distribution of the estimator is (Theorem 2, p. 843, Jarrow et al. (2011a)):

$$\sqrt{N_x^n}\left(\frac{V_n(x)}{\sigma^2(x)} - 1\right) \sim \sqrt{2}\Phi(0, 1) \tag{15}$$

where $N_x^n = \sum_{i=1}^{n} 1_{\{|S_{t_i} - x| < h_n\}}$ counts the number of observations $i = 1, ..., n$ in the interval $[x - h_n, x + h_n]$. Hence,

$$V_n(x) \sim \Phi\left(\sigma^2(x), 2\frac{\sigma^4(x)}{N_x^n}\right).$$

Replace $\sigma^4(x)$ with $S_n^2(x)$, to obtain

$$V_n(x) \sim \Phi\left(\sigma^2(x), 2\frac{S_n^2(x)}{N_x^n}\right).$$

The standard error (sample standard deviation) of $V_n(x)$ is estimated as:

$$\sqrt{2\frac{V_n^2(x)}{N_x^n}} = \sqrt{2}\frac{V_n(x)}{\sqrt{N_x^n}}.$$

Note that the standard error increases with $V_n(x)$.

A.3 Robustness Test

This robustness test adjusts the maximum and minimum observed variance estimates, which typically have the largest standard errors. The procedure is as follows.

- Replace the maximum value $V_n^*(x)$ with $V_n^*(x) - \kappa\sqrt{2}\frac{V_n^*(x)}{\sqrt{N_x^n}}$ with $\kappa > 0$ a constant, and
- replace the minimum value $V_n^\#(x)$ with $V_n^\#(x) + \kappa\sqrt{2}\frac{V_n^\#(x)}{\sqrt{N_x^n}}$.
- Given the sampling distribution in the previous section, these κ determine the probability that the variance exceeds the adjusted variance estimator, i.e.

$$P\left\{\sigma^2(x) > V_n^\#(x) + \kappa\sqrt{2}\frac{V_n^\#(x)}{\sqrt{N_x^n}}\right\} = 1 - \Phi\left\{k\right\}$$

  where Φ is the standard $(0, 1)$ cumulative normal distribution function.
- We try various choices of $\kappa > 0$, i.e $\kappa = 0.05, 0.10, 0.15, 0.20$.

A.4 Unequal Price Levels

A problem with the large size of the standard error bands is that if h_n is small, the number of elements in $\{|S_t - x| < h_n\}$, i.e. $N_x^n = \sum_{i=1}^n 1_{\{|S_{t_i} - x| < h_n\}}$ is small. Our estimator is consistent as $h_n \to 0$ and $n \to \infty$, implying $N_x^n \to \infty$. In a finite sample, as h_n gets smaller, N_x^n gets smaller as well. So, there is a trade-off with the size of h_n and the size of N_x^n.

A.4.1 Equal Price Level Partitions

We give the details of the equal price level partition so that the unequal price level partition case is more easily understood.

Consider the graph where the x - axis is time and the y - axis is the stock price level. On the x - axis we have the times $t_1 = 0, t_2, t_3, ..., t_n = T$ where n is the total number of price observations over the time interval $[0, T]$.

Partition the y - axis in equal units of $2h_n$ where h_n is in dollars. Here, the y - axis now has the partition points at

$$\{0, 2h_n, 4h_n, 6h_n, \cdots, mh_n\}$$

where m is an even number. The partitioned intervals are

$$\{[0, 2h_n], [2h_n, 4h_n], \ldots [(j-1)h_n, jh_n], \ldots, [(m-2)h_n, mh_n]\}.$$

The midpoint of these intervals are the stock price levels

$$\{x_2 = h_n, x_4 = 3h_n, ..., x_m = (m-1)h_n\}$$

used in the variance estimators. For example, $x_2 = h_n \in [0, 2h_n]$. Note to keep the notation simple, we index the stock price levels $j = 2, 4, ..., m$ with the same index as the upper value of the partitions on the y - axis.

For t_i, $i = 1, ..., n-1$ compute the estimate of the sample variance using $(S_{t_i+1} - S_{t_i})^2$. Note that this looks forward to the next time period to compute the value.

Consider the stock price level x_j. Compute the sample variance as

$$
\begin{aligned}
V_n(x_j) &= \frac{\sum_{i=1}^{n} 1_{\{|S_{t_i}-x_j|<h_n\}} (S_{t_{i+1}} - S_{t_i})^2 \cdot \frac{1}{[t_{i+1}-t_i]}}{\sum_{i=1}^{n} 1_{\{|S_{t_i}-x_j|<h_n\}}} \\
&= \frac{\sum_{i=1}^{n} 1_{\{S_{t_i}\in[(j-1)h_n, jh_n]\}} (S_{t_{i+1}} - S_{t_i})^2 \cdot \frac{1}{[t_{i+1}-t_i]}}{\sum_{i=1}^{n} 1_{\{S_{t_i}\in[(j-1)h_n, jh_n]\}}} \\
&= \frac{\sum_{i=1}^{n} 1_{\{S_{t_i}\in[(j-1)h_n, jh_n]\}} (S_{t_{i+1}} - S_{t_i})^2 \cdot \frac{1}{[t_{i+1}-t_i]}}{N_{x_j}^n} \qquad (16)
\end{aligned}
$$

Note that the numerator is the sum of the $(S_{t_{i+1}} - S_{t_i})^2 \cdot \frac{1}{[t_{i+1}-t_i]}$ for those times t_i where the stock price is in the jth partition of the y - axis, i.e, $S_{t_i} \in [(j-1)h_n, jh_n]$. The denominator is the sum of the times where the stock price is in the jth partition of the y - axis, i.e. $S_{t_i} \in [(j-1)h_n, jh_n]$.

The choice of "h_n" should be done so that $N_{x_j}^n$ is a large number for most x_j for $j = 2, 4, ..., m$. We want $N_{x_j}^n$ large so that the sampling distribution is approximately correct, i.e. near the asymptotic result.

A.4.2 Increasing h_n in Stock Price Level

Because the variance $V_n(x_j)$ is increasing in x_j, we can make $h_n = h_n(x)$ increasing in x. To keep the notation simple, we will drop the n subscript in the notation for the partition.

We partition the price axis into m bins in the following fashion. First, set the size of the first partition $h_1 > 0$. Then, set $h_j = \theta \times j \times h_1$ for $j = 1, ..., m$ where $\theta \in (0, 1)$. The partition is

$$\{0, 2h_1, 2h_1 + 2h_2, 2\sum_{k=1}^{3} h_k, \cdots, 2\sum_{k=1}^{m} h_k\}$$

where m is an even number. The price level intervals are

$$\left\{[0, 2h_1], [2h_1, 2h_1 + 2h_2], \cdots, \left[2\sum_{k=1}^{m-1} h_k, 2\sum_{k=1}^{m} h_k\right]\right\}.$$

In the variance estimator, x_j is the midpoint of the jth interval $\left[2\sum_{k=1}^{j-1} h_k, 2\sum_{k=1}^{j} h_k\right]$, computed as

$$x_j = \frac{2\sum_{k=1}^{j-1} h_k + 2\sum_{k=1}^{j} h_k}{2} = 2\sum_{k=1}^{(j-1)} h_k + h_j.$$

The variance estimator is

$$V_n\left(x_j\right) = \frac{\sum_{i=1}^{n} 1_{\left\{S_{t_i} \in \left[2\sum_{k=1}^{j-1} h_k, 2\sum_{k=1}^{j} h_k\right]\right\}} \left(S_{t_{i+1}} - S_{t_i}\right)^2 \cdot \frac{1}{\left[t_{i+1} - t_i\right]}}{N_{x_j}^n}. \tag{17}$$

We note that the estimator will still be consistent as long as for each price level partition k, $nh_k \to \infty$ and $nh_k^4 \to 0$ as $n \to \infty$ as required by the hypothesis of the theorem.

A.5 Small Sample Size Bias Adjustment

Our variance estimator is consistent, but may contain a small sample bias. To obtain a correction for any small sample bias, we add the following assumption.

Assumption: (*Linear approximation*)

$$\sigma(x) \approx \phi + \eta x$$

for constants ϕ, η. These do not need to be positive, the result does not depend on the sign of these constants. In addition, we require $\sigma(x) \to 0$ as $x \to 0$, which implies $\phi = 0$, i.e.

$$\sigma(x) \approx \eta x.$$

(*Conservative Assumption*)

This is a conservative assumption because we are assuming that as an approximation, the process is geometric Brownian motion, which has no bubbles. This implies the assumed volatility function increases in x more slowly than that needed for the existence of an asset price bubble. This ends the remark.

We now prove under this assumption that our variance estimator is biased, and we derive a small sample bias adjustment.

Proof. Under the linear approximation assumption,

$$\frac{d\left(\sigma^2(x)\right)}{dx} = 2\eta^2 x = 2\frac{\sigma^2(x)}{x}.$$

Hence,

$$\begin{aligned}\sigma^2(S_{t_i}) &= \sigma^2(x) + \frac{d\left(\sigma^2(x)\right)}{dx}\left(S_{t_i} - x\right) + \varepsilon \\ &\approx \sigma^2(x) + \frac{2\sigma^2(x)}{x}\left(S_{t_i} - x\right).\end{aligned}$$

To see the bias, we first replace $\sigma^2(S_{t_i})$ in this expression with our estimator for a single observation, $\frac{(S_{t_{i+1}} - S_{t_i})^2}{t_{i+1} - t_i}$. This gives

$$\frac{\left(S_{t_{i+1}} - S_{t_i}\right)^2}{t_{i+1} - t_i} \approx \sigma^2(x) + \frac{2\sigma^2(x)}{x}\left(S_{t_i} - x\right).$$

Summing across $i = 1, ..., n$ and multiplying by

$$\frac{1_{\{|S_{t_i}-x|<h_n\}}}{N_x^n}$$

gives our estimator on the left side:

$$V_n(x) = \sum_{i=1}^{n} \frac{1_{\{|S_{t_i}-x|<h_n\}}}{N_x^n} \frac{(S_{t_{i+1}} - S_{t_i})^2}{[t_{i+1} - t_i]} = \sigma^2(x) + \frac{2\sigma^2(x)}{x} \sum_{i=1}^{n} \frac{1_{\{|S_{t_i}-x|<h_n\}}(S_{t_i} - x)}{N_x^n} \tag{18}$$

where

$$\sum_{i=1}^{n} \frac{1_{\{|S_{t_i}-x|<h_n\}}}{N_x^n} = 1.$$

This shows the following result:

Under the linear approximation assumption, the variance estimator is biased. If on average $S_{t_i} > x$, then the estimator will be biased upward.

This will typically be the case, as discussed earlier, if $E_{t_i}(S_{t_{i+1}}) > S_{t_i}$ and $\sigma^2(x)$ is increasing in x.

A.5.1 The Bias Correction

Using expression (18), an approximately unbiased estimator for $\sigma^2(x)$ can be shown to be the following.

$$\hat{\sigma}^2(x) = \frac{V_n(x)}{1 + \frac{2}{x}\sum_{i=1}^{n} \frac{1_{\{|S_{t_i}-x|<h_n\}}(S_{t_i}-x)}{N_x^n}} \tag{19}$$

Proof:

$$E_t\left[\frac{V_n(x)}{1 + \frac{2}{x}\sum_{i=1}^{n} \frac{1_{\{|S_{t_i}-x|<h_n\}}(S_{t_i}-x)}{N_x^n}}\right]$$

$$= E_t\left[\frac{\sigma^2(x)\left(1 + \frac{2}{x}\sum_{i=1}^{n} \frac{1_{\{|S_{t_i}-x|<h_n\}}(S_{t_i}-x)}{N_x^n}\right)}{1 + \frac{2}{x}\sum_{i=1}^{n} \frac{1_{\{|S_{t_i}-x|<h_n\}}(S_{t_i}-x)}{N_x^n}}\right] = \sigma^2(x).$$

□

A.6 Inconclusive Region and the Probability of Default

The hypothesis testing method, based on upper and lower bounds for the volatility function, is conservative. As such, there is a region where the hypothesis testing is inconclusive. In this inconclusive region, this section computes a lower bound on the probability of default.

Recall that the lower convex hull and the upper convex hull functions satisfy the inequalities $\sigma_l^2(x) \leq \sigma^2(x) \leq \sigma_u^2(x)$ by construction where $x \in [0, \infty)$. Given the integrals: $\mathscr{I} := \int_1^\infty \frac{x}{\sigma^2(x)} dx$,

$\mathscr{I}_u := \int_1^\infty \frac{x}{\sigma_l^2(x)}dx$, and $\mathscr{I}_l := \int_1^\infty \frac{x}{\sigma_u^2(x)}dx$, this implies that

$$\mathscr{I}_u > \mathscr{I} > \mathscr{I}_l.$$

This is because the estimates are in the denominator of the integrals. Therefore,[10]

$$\mathscr{I}_u < \infty \Rightarrow \mathscr{I} < \infty \Rightarrow \mathscr{I}_l < \infty. \tag{20}$$

We note that

$$\mathscr{I} < \infty \Leftrightarrow bubble.$$

Next assume that the upper and lower convex hulls satisfy

$$\sigma_k(x) := \alpha_k x^{\beta_k}, \tag{21}$$

where $\alpha_k \geq 0$ and $k \in \{u, l\}$. Substitution into the integrals yields:

$$\mathscr{I}_k < \infty \Leftrightarrow \beta_k > 1$$

for $k \in \{u, l\}$.

Given the hypothesis testing provides no conclusion regarding whether the price process exhibits a bubble, we take the Bayesian perspective that the variance functions, integrals, and betas are random variables on a probability space $(\Psi, \mathscr{G}, \mathscr{P})$ where $\mathscr{P}$ is the posterior distribution for these random variables given the price observations $x_i \in [0, \infty) : i = 1, ..., n$.

Given this interpretation, letting $\psi \in \Psi$, then for a given $x \in [0, \infty)$, $\sigma(x)_\psi : \Psi \to \mathbb{R}$ and $\sigma_k(x)_\psi : \Psi \to \mathbb{R}$ are random functions, and $\{\mathscr{I}(\psi), \mathscr{I}_k(\psi), \beta(\psi), \beta_k(\psi)\}$ are random variables for $k \in \{u, l\}$. Because we are not interested in the parameter α_k, we assume it is a constant for $k \in \{u, l\}$.

On this probability space, we have the following events: $bubble = \{\psi \in \Psi : \mathscr{I} < \infty\}$ and $\{\psi \in \Psi : \mathscr{I}_k < \infty\}$ for $k \in \{u, l\}$. Note that these events are exhaustive and mutually exclusive, i.e. $\{\psi \in \Psi : \mathscr{I} < \infty\} \cup \{\psi \in \Psi : \mathscr{I} = \infty\} = \Psi$, and the same for $\mathscr{I}_k$, $k \in \{u, l\}$.

Expression (20) gives[11]

$$\{\psi \in \Psi : \mathscr{I}_u < \infty\} \subseteq \{\psi \in \Psi : \mathscr{I} < \infty\} \subseteq \{\psi \in \Psi : \mathscr{I}_l < \infty\},$$

which implies that

$$\mathscr{P}\{\beta_l > 1\} = \mathscr{P}\{\mathscr{I}_u < \infty\} \leq \mathscr{P}\{\mathscr{I} < \infty\} = \mathscr{P}\{bubble\}$$

and

$$\mathscr{P}\{bubble\} \leq \mathscr{P}\{\mathscr{I}_u < \infty\} = \mathscr{P}\{\beta_u > 1\}.$$

[10] In the other direction, $\mathscr{I}_l = \infty \Rightarrow \mathscr{I} = \infty \Rightarrow \mathscr{I}_k = \infty$.

[11] Similarly, the complements satisfy

$$\{\omega \in \Omega : \mathscr{I}_l = \infty\} \subset \{\omega \in \Omega : \mathscr{I} = \infty\} \subset \{\omega \in \Omega : \mathscr{I}_u = \infty\}.$$

The probability distribution (law) for β_k, $k \in \{u, l\}$ is denoted

$$Prob(\beta_k \leq x) \coloneqq \mathscr{P}\{\beta_k^{-1}((-\infty, x])\}$$

for $(-\infty, x] \in \mathscr{B}(\mathbb{R})$, the Borel σ-algebra. Then,

$$1 - Prob(\beta_l \leq 1) = \mathscr{P}\{\beta_l > 1\} \leq \mathscr{P}\{bubble\}$$

and

$$\mathscr{P}\{bubble\} \leq \mathscr{P}\{\beta_u > 1\} = 1 - Prob(\beta_u \leq 1).$$

We add the following assumption, motivated by the point estimation for β_k obtained from the regression analysis and its sampling distribution.

Assumption (Normal Distribution)

$$Prob(\beta_k \leq x) = \Phi\left(\hat{\beta}_k, \hat{\sigma}^2_{\hat{\beta}_k}\right)$$

with mean $E^{\mathscr{P}}(\beta_k) = \hat{\beta}_k$ and variance $var^{\mathscr{P}}(\beta_k) = \hat{\sigma}^2_{\hat{\beta}_k}$.

Under this assumption,

$$Prob\left(\beta_l > 1\right) = 1 - \mathit{\Phi}\left(\frac{1 - \hat{\beta}_l}{\hat{\sigma}^2_{\hat{\beta}_l}}\right) \leq \mathscr{P}\{bubble\}$$

and

$$\mathscr{P}\{bubble\} \leq 1 - \mathit{\Phi}\left(\frac{1 - \hat{\beta}_u}{\hat{\sigma}^2_{\hat{\beta}_u}}\right) = Prob\left(\beta_u > 1\right).$$

Given diffuse priors across the upper and lower bounds on the probability of a bubble, our point estimate is

$$\hat{\mathscr{P}}\{bubble\} = 1 - \frac{\left[\mathit{\Phi}\left(\frac{1-\hat{\beta}_l}{\hat{\sigma}^2_{\hat{\beta}_l}}\right) + \mathit{\Phi}\left(\frac{1-\hat{\beta}_u}{\hat{\sigma}^2_{\hat{\beta}_u}}\right)\right]}{2}.$$