Johnson School Research Paper Series #28-2010

How to Detect an Asset Bubble

Robert A. Jarrow—Cornell University
Younes Kchia—Centre de Mathematiques Appliquees, Ecole Polytechnique
Philip Protter—Cornell University

March 2011

How to Detect an Asset Bubble

Robert Jarrow[*] Younes Kchia[†] Philip Protter[‡§]

February 24, 2011

Abstract

After the 2007 credit crisis, financial bubbles have once again emerged as a topic of current concern. An open problem is to determine in real time whether or not a given asset's price process exhibits a bubble. Due to recent progress in the characterization of asset price bubbles using the arbitrage-free martingale pricing technology, we are able to propose a new methodology for answering this question based on the asset's price volatility. We limit ourselves to the special case of a risky asset's price being modeled by a Brownian driven stochastic differential equation. Such models are ubiquitous both in theory and in practice. Our methods use sophisticated volatility estimation techniques combined with the method of reproducing kernel Hilbert spaces. We illustrate these techniques using several stocks from the alleged internet dot-com episode of 1998 - 2001, where price bubbles were widely thought to have existed. Our results support these beliefs.

1 Introduction

This paper is interested in the detection of financial bubbles. The question we address is a timely one. Recently William Dudley, the President of the New York Federal Reserve, in an interview with Planet Money [8] stated "...what I am proposing is that we try to identify bubbles in real time, try to develop tools to address those bubbles, try to use those tools when appropriate to limit the size of those bubbles and, therefore, try to limit the damage when those bubbles burst."

Asset price bubbles have also been recently characterized in frictionless, competitive, and continuous trading economies using the abitrage-free martingale pricing technology underlying option pricing theory (see [22], [23], [2], [9], [17] and [18]). In this classical setting, Jarrow, Protter and Shimbo [17] and [18] show there are three types of asset price bubbles possible. Two of these price bubbles exist only in infinite horizon economies, the third - called type 3 bubbles - exist in finite horizon settings. Consequently, type 3 bubbles are

[*]Johnson Graduate School of Management, Cornell University, Ithaca, NY, 14853 and Kamakura Corporation.

[†]Centre de Mathématiques Appliquées, Ecole Polytechnique, Paris.

[‡]Statistics Department, Columbia University, New York, NY, 10027.

[§]Supported in part by NSF grant DMS-0906995

those most relevant to actual market experiences. For this type of bubble, whether or not a bubble exists amounts to determining if the price process under a risk neutral measure is a martingale or a strict local martingale: if it is a strict local martingale, there is a bubble. The difference between a martingale and a strict local martingale has been recently investigated by several authors ([1], [25] and [20] for instance). However, the distinction is subtle and in the case of a diffusion it amounts to understanding the asymptotic behavior of the asset's price volatility. If the asset's price volatility is large enough, then a bubble exists.

More formally, we model the asset price process by a standard stochastic differential equation driven by a Brownian motion W:

$$dS_t = \sigma(S_t)dW_t + \mu(S_t)dt \tag{1}$$

for all t in $[0, T]$, in some filtered probability space $(\Omega, \mathcal{F}, P, \mathbb{F})$ where $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$. We make the standing assumption that the asset price S is nonnegative. The asset's price volatility $\sigma(S_t)$ is stochastic since it depends on the level of the asset's price. Assuming no arbitrage in the sense of "No Free Lunch Vanishing Risk (NFLVR)," there exists a risk neutral measure (see [5]) under which this SDE simplifies to

$$S_t = S_0 + \int_0^t \sigma(S_s)dW_s. \tag{2}$$

It is well known ([1], [25]) that this process S is a strict local martingale if and only if

$$\int_\alpha^\infty \frac{x}{\sigma^2(x)}dx < \infty \tag{3}$$

for all $\alpha > 0$. Dmitry Kramkov has recently pointed out that this follows quite simply from Feller's test for explosions [21]. This last condition forms the basis of our bubble testing methodology, i.e. type 3 bubbles exist if and only if this integral is finite. This uses the theory of bubbles as presented, for example, in [2]. The intuition behind the distinction between a martingale and a strict local martingale (in the case where the local martingale $S > 0$) derived from the fact that S is always a supermartingale and is a martingale if and only if it has constant expectation. So for a strict local martingale its expectation decreases with time. Thus on average under the risk neutral measure the buy and hold strategy is a losing one. A "typical" path of such a nonnegative continuous local martingale is to shoot up to high values and then quickly decrease to small values and remain at them; and this is also the typical behavior of prices of assets undergoing speculative bubbles.

We can also replace the equation (1) for S with a more general one, for example

$$\begin{aligned} dS_t &= \sigma(S_t)dW_t + \mu(S_t, Y_t)dt \\ dY_t &= s(Y_t)dB_t + g(Y_t)dt \end{aligned} \tag{4}$$

where B is a Brownian motion which is independent of W. This gives a model for S in the context of an incomplete market. This is perhaps the simplest model that implies an incomplete market. An alternative incomplete market model, and one that we do not consider here, introduces a stochastic volatility function as well. In any event, for our purposes, we are inevitably led in both situations (complete or incomplete models) to equation (2), under any risk neutral measure. If the models are complete, we appeal to the NFLVR framework of Cox and Hobson [2], and if the models are incomplete we can use the No Dominance framework of [17],[18], which we consider a better framework for models of bubbles.

Many authors have proposed estimators for the volatility function $\sigma(x)$. D. Florens-Zmirou [6] proposed a non parametric estimator based on the local time of the diffusion process. V. Genon Catalot and J. Jacod [7] propose an estimation procedure for parameterized volatility functions. M. Hoffmann [11] constructs a wavelets based estimator. In the first part of this paper, we recall Florens-Zmirou's results. Since the constraint on the grid step, noted h_n in the sequel, required by Florens-Zmirou's theorem cannot be satisfied due to the limited data available, we propose another local time based estimator, using a smooth kernel, where the condition on h_n is easier to satisfy. Florens-Zmirou also obtained a limit theorem when $\sigma(x)$ is bounded above and away from zero. We were able to relax this assumption and extend the limit theorem by working in an enlarged filtration. We do not reproduce these results herein since they are tangential to this paper, but they are available to the reader upon request.

The main difficulty in using non parametric estimators is that one can estimate $\sigma(x)$ only at points visited by the process. We, therefore, cannot know the tails of the volatility function and determine if the integral in (3) is finite or infinite. In the second part of this paper we propose two methods to deal with this "extrapolation problem."

The first method is based on a comparison theorem. We compare the behavior of parametric and non-parametric estimators of $\sigma(x)$. When the two estimators are statistically similar within the observation interval, we extrapolate into the tails using the parametric form's asymptotic behavior. The second method is based on Reproducing Kernel Hilbert Spaces (*RKHS*) theory. In fact, a roughly analogous problem arises in physical chemistry for potentials whose asymptotic behavior is known (c.f. [12]). In our case, we do not know the asymptotic behavior of the volatility (that is what we are looking for!). To overcome this problem, we introduce a parameterized family (H_m) of RKHS's. Different m's allow us to construct interpolating functions with different asymptotic behaviors. We optimize over m in a sense that will be explained below and identify the reproducing kernel Hilbert space which allows us to construct an interpolating function that extends the non-parametric estimator from the observation interval to the entire real line.

We devote the last section to illustrating these various estimation methodologies. We focus on stocks from the alleged internet dotcom bubble of 1998 - 2001 (see for instance [29] and [24]) for which we could find relevant tick data. We selected four stocks: *Lastminute.com*, *Etoys*, *Infospace* and *Geocities*. The data was obtained from WRDS [30]. We use our methodology to see whether these stocks exhibited price bubbles. The evidence supports the existence of price bubbles. In addition, these four stocks allow us to illustrate the strengths and weaknesses of our testing methodology. We can also develop (and have done so) tests for bubble behavior via an analysis of derivatives (calls and puts). We have not used that here since the data sets from 1998-2001 are limited, and we could not find a source for large enough quantities of good data on derivatives during that period. Nevertheless such techniques might work well in a more contemporary setting.

We note that the estimation is performed in the real (and not the risk neutral) world. However, Florens-Zmirou shows that the estimators we use do not involve the drift, hence without loss of generality, we assume throughout the remainder of the paper that μ is identically null. Therefore, we consider the stochastic differential equation in (2) where the function $\sigma(x)$ is unknown. We define our non parametric estimator of $\sigma(x)$ based on discrete time observations $S_{t_1}, ..., S_{t_n}$, on the finite time interval $[0, T]$. We assume a regular sampling, that is $t_i = \frac{i}{n}T$.

An outline for our paper is as follows. Sections 2 and 3 present the Florens-Zmirou's and smooth kernel volatility estimators on a compact domain. Section 4 extends these estimators to the nonnegative real line. Section 5 illustrates our testing methodology for asset price bubbles, and section 6 concludes.

2 Florens-Zmirou's Estimators

This section reviews Florens-Zmirou's estimators for our subsequent usage. Her estimator is based on the local time of a diffusion and is explained heuristically as follows. The local time is given by

$$\ell_T(x) = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^T 1_{\{|S_s - x| < \epsilon\}} d\langle S, S\rangle_s$$

where $d\langle S, S\rangle_s = \sigma^2(S_s)ds$ so that $\ell_T(x) = \sigma^2(x)L_T(x)$, and

$$L_T(x) = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^T 1_{\{|S_s - x| < \epsilon\}} ds.$$

Hence, the ratio $\frac{\ell_T(x)}{L_T(x)} = \sigma^2(x)$ yields the volatility at x. These limits and integrals can be approximated by the following sums :

$$L_T^n(x) = \frac{T}{2nh_n} \sum_{i=1}^{n} 1_{\{|S_{t_i} - x| < h_n\}}$$
$$\ell_T^n(x) = \frac{T}{2nh_n} \sum_{i=1}^{n} 1_{\{|S_{t_i} - x| < h_n\}} n(S_{t_{i+1}} - S_{t_i})^2$$

where h_n is a sequence of positive real numbers converging to 0 and satisfying some constraints. This allows us to construct an estimator of $\sigma(x)$ given by:

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

Indeed, Florens-Zmriou [6] proves the following theorems.

Theorem 1 *If σ is bounded above and below from zero, has three continuous and bounded derivatives, and if $(h_n)_{n\geq 1}$ satisfies $nh_n \to \infty$ and $nh_n^4 \to 0$ then $S_n(x)$ is a consistent estimator of $\sigma^2(x)$.*

The proof of this theorem is based on the expansion of the transition density. The choice of a sequence h_n converging to 0 and satisfying $nh_n \to \infty$ and $nh_n^4 \to 0$ allows one to show that $L_T^n(x)$ and $\ell_T^n(x)$ converge in $L^2(dQ)$ to $L_T(x)$ and $\sigma^2(x)L_T(x)$, respectively. Hence $S_n(x)$ is a consistent estimator of $\sigma^2(x)$, for any x that has been visited by the diffusion.

We also have the following limit theorem, useful to obtain confidence intervals for the estimator $S_n(x)$ of $\sigma(x)$.

Theorem 2 *If moreover $nh_n^3 \to 0$ then $\sqrt{N_x^n}(\frac{S_n(x)}{\sigma^2(x)} - 1)$ converges in distribution to $\sqrt{2}Z$ where Z is a standard normal random variable and $N_x^n = \sum_{i=1}^n 1_{\{|S_{t_i}-x|<h_n\}}$.*

The aim of the first part of this paper is to construct another estimator based on the local time of the diffusion but using a smooth kernel. Theorem 1 will remain true under the constraint $nh_n^2 \to \infty$. This is important for the purpose of this paper: requiring that h_n satisfies the conditions of Theorem 1 would provide useless estimators (not smooth enough to work with in practice) due to the limited data available to us. When testing our procedure, we will always provide these two estimators although theoretically we cannot be sure that the estimator of Florens-Zmirou converges due to the requirement that $h_n = n^{-\frac{1}{4}}$, which needs more data than we have.

3 A Smooth Kernel Estimator

We introduce a smooth kernel estimator to relax the condition on h_n to $nh_n^2 \to \infty$. In practice we often do not have enough data and the convergence condition $nh_n^4 \to 0$ is too restrictive. The key theorem of this section proves the convergence in probability of our sequence of smooth kernel estimators $S_n(x)$ to $\sigma^2(x)$.

For this section and without loss of generality, we assume that $T = 1$ and $t_i = \frac{i}{n}$. We consider again the discrete observation $S^{(n)} = (S_0, S_{\frac{1}{n}}, ..., S_1)$ defined through the stochastic differential equation (2). We assume that $\sigma(x)$ is bounded above and away from zero, C^3, and with bounded derivatives. These assumptions guarantee the existence of a unique strong solution which does not explode. We denote by Q the law on the space of continuous functions equipped with the canonical filtration $(\mathcal{F}_t)_{t\in[0,1]}$ and under which the canonical process $(S_t, 0 \leq t \leq 1)$ is a solution to the previous SDE.

Note that from a statistical point of view, it is more natural to work with weak solutions. The smoothness assumption and the boundedness of $\sigma(x)$ and its derivatives are required to obtain some estimates. We also consider a compact interval $\mathcal{D}$, which represents the observation interval, i.e. the domain on which the estimation is performed. We emphasize the fact that we are able to estimate $\sigma(x)$ only for those points that have been visited by the diffusion.

The idea underlying the smooth kernel estimator is to replace the kernel $K(x) = \frac{1}{2}1_{\{|x|\leq 1\}}$ by a smooth kernel ϕ, which is a C^6 positive function with compact support and such that $\int_{\mathbb{R}_+} \phi = 1$. We are interested in some L^p (p is stated later) convergence of the following quantities:

$$V_n^x = \frac{1}{nh_n}\sum_{i=0}^{n-1}\phi(\frac{S_{\frac{i}{n}} - x}{h_n})n(S_{\frac{i+1}{n}} - S_{\frac{i}{n}})^2 \quad (6)$$

$$L_n^x = \frac{1}{nh_n}\sum_{i=0}^{n-1}\phi(\frac{S_{\frac{i}{n}} - x}{h_n}) \quad (7)$$

to $\sigma^2(x)L^x$ and L^x respectively, where h_n satisfies $nh_n^2 \to \infty$.

Our convergence theorem requires the use of various lemmas. The first lemma involves the convergence of L_n^x and it follows from Proposition 3 of Hoffmann [11, p. 468].

Lemma 1 *Assume ϕ given in (7) is taken to be C^3. For each $\gamma \geq 2$, there exists a constant C such that*

$$\sup_{x\in\mathcal{D}} E(|L_n^x - L^x|^\gamma) \leq C(h_n^{\frac{\gamma}{2}} + (\frac{1}{nh_n^2})^\gamma).$$

Hence the L^p convergence of L_n^x to L^x is guaranteed, for all $p \geq 2$ and all $x \in \mathcal{D}$, and the L^p convergence for all $p > 0$ follows by Cauchy Schwarz. Our next lemma involves the L^1 convergence of V_n^x.

Lemma 2 *For each $x \in \mathcal{D}$, V_n^x converges in L^1 to $\sigma^2(x)L^x$.*

In order to prove this lemma, we write $V_n^x - \sigma^2(x)L^x = A_n(x) + B_n(x)$ where:

$$A_n(x) = \frac{1}{h_n}\int_0^1 \phi(\frac{X_s - x}{h_n})\sigma^2(X_s)ds - \sigma^2(x)L^x$$

$$B_n(x) = V_n^x - \frac{1}{h_n}\int_0^1 \phi(\frac{X_s - x}{h_n})\sigma^2(X_s)ds$$

We study each of those two terms separately. Let x be fixed in $\mathcal{D}$, where $\mathcal{D}$ is the domain over which the estimation is performed. Since x is fixed, we omit it from now on and prove that A_n and B_n converge in L^1 to 0.

Lemma 3 (Study of A_n) *For each $\gamma \geq 2$, there exists $c > 0$ such that $E|A_n|^\gamma \leq ch_n^{\frac{\gamma}{2}}$.*

Proof. Let $l = (l_1^x)$ be the local time of the diffusion at time $t = 1$. First, we use the occupation time formula,

$$\begin{aligned} A_n &= \frac{1}{h_n}\int_0^1 \phi(\frac{X_s - x}{h_n})\sigma^2(X_s)ds - \sigma^2(x)L^x \\ &= \frac{1}{h_n}\int_{\mathbb{R}^+} \phi(\frac{y-x}{h_n})l^y dy - \frac{1}{h_n}\int_{\mathbb{R}^+} \phi(\frac{y-x}{h_n})l^x dy = \frac{1}{h_n}\int_{\mathbb{R}^+} \phi(\frac{y-x}{h_n})(l^y - l^x)dy. \end{aligned}$$

Applying Jensen's lemma to the integral, a straightforward change of variables and taking expectations give:

$$E|A_n|^\gamma \leq \frac{1}{h_n^\gamma} E(\int_{\mathbb{R}^+} |l^{zh_n+x} - l^x|^\gamma \phi^\gamma(z) h_n^\gamma dz).$$

The following inequalities follow from an application of Fubini's theorem and the Hölder property of the local time paths of a continuous local martingale. (This is a well known and classic result, given for example by Revuz and Yor [27], p. 227.)

$$E|A_n|^\gamma \leq \int_{\mathbb{R}^+} \phi^\gamma(z) E(|l^{zh_n+x} - l^x|^\gamma) dz \leq h_n^{\frac{\gamma}{2}} \int_{\mathbb{R}^+} |z|^{\frac{\gamma}{2}} \phi^\gamma(z) dz$$

Since ϕ has compact support, we can take $c = \int_{\mathbb{R}^+} |z|^{\frac{\gamma}{2}} \phi^\gamma(z) dz$. Then $E|A_n|^\gamma \leq c h_n^{\frac{\gamma}{2}}$ and the lemma is proved. ■

We now focus on the study of B_n, which we write $-B_n = C_n + D_n$ where:

$$C_n = \frac{1}{h_n} \sum_{i=0}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} (\phi(\frac{S_s - x}{h_n}) \sigma^2(S_s) - \phi(\frac{S_{\frac{i}{n}} - x}{h_n}) \sigma^2(S_{\frac{i}{n}})) ds$$

$$D_n = \frac{1}{h_n} \sum_{i=0}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \phi(\frac{S_{\frac{i}{n}} - x}{h_n}) (\sigma^2(S_{\frac{i}{n}}) - n(S_{\frac{i+1}{n}} - S_{\frac{i}{n}})^2) ds$$

We need the following lemma borrowed from Genon-Catalot and Jacod, which we recall in the one dimensional setting. We refer to [7] for a proof. We define $X_i^n = \sqrt{n}\sigma(S_{\frac{i}{n}})(W_{\frac{i+1}{n}} - W_{\frac{i}{n}})$ and $Y_i^n = \sqrt{n} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma(S_s) dW_s$.

Lemma 4 (Genon-Catalot and Jacod) *Let $g \in C^2$. Assume there exists $\gamma > 0$ such that for all x, $|g(x)| + |g^{'}(x)| + |g^{''}(x)| \leq \gamma(1 + |x|^\gamma)$. Then there exists a constant C such that*

$$E\Big(\big(g(X_i^n) - g(Y_i^n)\big)^2 | \mathcal{F}_{\frac{i}{n}}\Big) \leq \frac{C}{n}.$$

If g is an even function, $\big|E(g(X_i^n) - g(Y_i^n) | \mathcal{F}_{\frac{i}{n}})\big| \leq \frac{C}{n}$.

We can now study the convergence of D_n.

Lemma 5 (Study of D_n) *There exists $C > 0$, such that $E|D_n| \leq \frac{C}{n} E\Big(\frac{1}{nh_n} \sum_{i=0}^{n-1} \phi(\frac{S_{\frac{i}{n}} - x}{h_n})\Big)$ and D_n converges to 0 in L^1.*

Proof. Recall that $D_n = \frac{1}{h_n} \sum_{i=0}^{n-1} \phi(\frac{S_{\frac{i}{n}} - x}{h_n}) (\frac{\sigma^2(S_{\frac{i}{n}})}{n} - \big(\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma(S_s) dW_s\big)^2)$. Write $g(x) = x^2$. The following inequalities are straightforward.

$$\begin{aligned} E|D_n| &\leq \frac{1}{h_n} \sum_{i=0}^{n-1} E\Big(\phi(\frac{S_{\frac{i}{n}} - x}{h_n}) \big| E\big(\frac{\sigma^2(S_{\frac{i}{n}})}{n} | \mathcal{F}_{\frac{i}{n}}\big) - E\big((\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma(S_s) dW_s)^2 | \mathcal{F}_{\frac{i}{n}}\big) \big| \Big) \\ &\leq \frac{1}{h_n} \sum_{i=0}^{n-1} E\Big(\phi(\frac{S_{\frac{i}{n}} - x}{h_n}) \big| E(g(X_i^n) - g(Y_i^n) | \mathcal{F}_{\frac{i}{n}}) \big| \Big) \frac{1}{n}. \end{aligned}$$

Since g is assumed to be an even function, Lemma 4 ensures the existence of a constant C such that:

$$\left|E\left(g(X_i^n) - g(Y_i^n)|\mathcal{F}_{\frac{i}{n}}\right)\right| \leq \frac{C}{n}.$$

Hence, $E|D_n| \leq \frac{C}{n}E\left(\frac{1}{nh_n}\sum_{i=0}^{n-1}\phi(\frac{S_{\frac{i}{n}}-x}{h_n})\right)$. Using Lemma 1, the sum converges to the local time of the diffusion in x and $E|D_n| \to 0$. ■

In order to study C_n, we introduce the function $f(y) = \phi(\frac{y-x}{h_n})\sigma^2(y)$ which is C^3 by assumption. We use a third order Taylor expansion and get: for all s in $[\frac{i}{n}, \frac{i+1}{n}[$, there exists $\Xi_{s,\frac{i}{n}}$ such that

$$f(S_s) = f(S_{\frac{i}{n}}) + (S_s - S_{\frac{i}{n}})f'(S_{\frac{i}{n}}) + \frac{(S_s - S_{\frac{i}{n}})^2}{2}f''(S_{\frac{i}{n}}) + \frac{(S_s - S_{\frac{i}{n}})^3}{6}f^{(3)}(\Xi_{s,\frac{i}{n}}).$$

We plug this into C_n and obtain $C_n = C_n^1 + C_n^2 + C_n^3 + C_n^4$ where

$$\begin{aligned}
C_n^1 = \frac{1}{h_n}\sum_{i=0}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}} & \Bigg((S_s - S_{\frac{i}{n}})\phi(\frac{S_{\frac{i}{n}}-x}{h_n})(\sigma^2)'(S_{\frac{i}{n}}) + \frac{1}{2}(S_s - S_{\frac{i}{n}})^2\Big(\frac{2}{h_n}\phi'(\frac{S_{\frac{i}{n}}-x}{h_n})(\sigma^2)'(S_{\frac{i}{n}}) \\
&+ \phi(\frac{S_{\frac{i}{n}}-x}{h_n})(\sigma^2)''(S_{\frac{i}{n}})\Big) + \frac{1}{6}(S_s - S_{\frac{i}{n}})^3\Big(\frac{3}{h_n^2}\phi''(\frac{\Xi_{s,\frac{i}{n}}-x}{h_n})(\sigma^2)'(\Xi_{s,\frac{i}{n}}) \\
&+ \frac{3}{h_n}\phi'(\frac{\Xi_{s,\frac{i}{n}}-x}{h_n})(\sigma^2)''(\Xi_{s,\frac{i}{n}}) + \phi(\frac{\Xi_{s,\frac{i}{n}}-x}{h_n})(\sigma^2)^{(3)}(\Xi_{s,\frac{i}{n}})\Big)\Bigg)ds
\end{aligned}$$

and

$$\begin{aligned}
C_n^2 &= \frac{1}{h_n^2}\sum_{i=0}^{n-1}\phi'(\frac{S_{\frac{i}{n}}-x}{h_n})\sigma^2(S_{\frac{i}{n}})\int_{\frac{i}{n}}^{\frac{i+1}{n}}(S_s - S_{\frac{i}{n}})ds \\
C_n^3 &= \frac{1}{h_n^3}\sum_{i=0}^{n-1}\phi''(\frac{S_{\frac{i}{n}}-x}{h_n})\sigma^2(S_{\frac{i}{n}})\int_{\frac{i}{n}}^{\frac{i+1}{n}}\frac{(S_s - S_{\frac{i}{n}})^2}{2}ds \\
C_n^4 &= \frac{1}{h_n^4}\sum_{i=0}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}}\phi^{(3)}(\frac{\Xi_{s,\frac{i}{n}}-x}{h_n})\sigma^2(\Xi_{s,\frac{i}{n}})\frac{(S_s - S_{\frac{i}{n}})^3}{6}ds
\end{aligned}$$

We prove in the following lemma that C_n^1 converges in L^1 to 0 when nh_n^2 tends to infinity. The idea is to bound C_n^1 by first bounding σ, ϕ and their three first derivatives and then using the Burkholder-Davis-Gundy inequalities (hereafter referred to simply as BDG) to obtain estimates of powers of $S_s - S_{\frac{i}{n}}$.

Lemma 6 (Study of C^1) *Assume that $nh_n^2 \to \infty$. Then C_n^1 converges in L^1 to 0.*

Proof. Since ϕ and σ and their derivatives are bounded, there exist nonnegative constants

$(c_i)_{1\le i\le 6}$ such that

$$E|C_n^1| \le E\Big(\sum_{i=0}^{n-1}\frac{c_1}{h_n}\int_{\frac{i}{n}}^{\frac{i+1}{n}}|S_s-S_{\frac{i}{n}}|ds+(\frac{c_2}{h_n}+\frac{c_3}{h_n^2})\int_{\frac{i}{n}}^{\frac{i+1}{n}}|S_s-S_{\frac{i}{n}}|^2ds$$
$$+(\frac{c_4}{h_n}+\frac{c_5}{h_n^2}+\frac{c_6}{h_n^3})\int_{\frac{i}{n}}^{\frac{i+1}{n}}|S_s-S_{\frac{i}{n}}|^3ds\Big)$$

It follows clearly that

$$E|C_n^1| \le \sum_{i=0}^{n-1}\frac{c_1}{h_n}\int_{\frac{i}{n}}^{\frac{i+1}{n}}E(\sup_{s\in[\frac{i}{n},\frac{i+1}{n}]}|S_s-S_{\frac{i}{n}}|)ds+(\frac{c_2}{h_n}+\frac{c_3}{h_n^2})\int_{\frac{i}{n}}^{\frac{i+1}{n}}E(\sup_{s\in[\frac{i}{n},\frac{i+1}{n}]}|S_s-S_{\frac{i}{n}}|^2)ds$$
$$+(\frac{c_4}{h_n}+\frac{c_5}{h_n^2}+\frac{c_6}{h_n^3})\int_{\frac{i}{n}}^{\frac{i+1}{n}}E(\sup_{s\in[\frac{i}{n},\frac{i+1}{n}]}|S_s-S_{\frac{i}{n}}|^3)ds$$

We apply now BDG inequalities for continuous local martingales. For each $1\le p\le 3$, there exist nonnegative constants C_p such that

$$E\Big(\sup_{s\in[\frac{i}{n},\frac{i+1}{n}]}|S_s-S_{\frac{i}{n}}|^p\Big)\le E\Big(\Big(\int_{\frac{i}{n}}^{\frac{i+1}{n}}\sigma^2(S_s)ds\Big)^{\frac{p}{2}}\Big)\le\frac{C_p}{n^{\frac{p}{2}}}.$$

Integrating, summing, and taking the expectation, we finally obtain

$$\begin{aligned}E|C_n^1| &\le \frac{c_1}{h_n}\frac{C_1}{\sqrt{n}}+(\frac{c_2}{h_n}+\frac{c_3}{h_n^2})\frac{C_2}{n}+(\frac{c_4}{h_n}+\frac{c_5}{h_n^2}+\frac{c_6}{h_n^3})\frac{C_3}{n\sqrt{n}}\\ &\le \frac{c_1C_1}{\sqrt{nh_n^2}}+\frac{C_2}{nh_n^2}(c_2h_n+c_3)+\frac{C_3}{nh_n^2}(\frac{c_4h_n}{\sqrt{n}}+\frac{c_5}{\sqrt{n}}+\frac{c_6}{\sqrt{nh_n^2}})\end{aligned}$$

Since $nh_n^2\to\infty$ and $h_n\to 0$, it follows clearly that C_n^1 converges in L^1 to 0. ■

We turn now to the study of C_n^2.

Lemma 7 (Study of C_n^2) *If $nh_n^2\to\infty$, then C_n^2 converges in L^1 to 0.*

Proof. First, $E|C_n^2| \le \frac{1}{h_n^2}||\sigma^2||_\infty E\Big(\sum_{i=0}^{n-1}|\phi'|(\frac{S_{\frac{i}{n}}-x}{h_n})\Big|E\Big(\int_{\frac{i}{n}}^{\frac{i+1}{n}}|S_s-S_{\frac{i}{n}}|ds\big|\mathcal{F}_{\frac{i}{n}}\Big)\Big|\Big)$. It follows from an application of a BDG inequality that there exists a constant M_1 such that

$$E|C_n^2|\le\frac{M_1}{\sqrt{n}h_n}E\Big(\frac{1}{nh_n}\sum_{i=0}^{n-1}|\phi'(\frac{S_{\frac{i}{n}}-x}{h_n})|\Big).$$

Now using the kernel $\frac{1}{a}|\phi'|$, where $a=\int|\phi'|(x)dx$, the quantity $\frac{1}{nh_n}\sum_{i=0}^{n-1}|\phi'|(\frac{S_{\frac{i}{n}}-x}{h_n})$ converges in L^1 to aL^x. Hence $E|C_n^2|$ converges to zero as soon as $nh_n^2\to\infty$, which proves the lemma. ■

We can provide another proof of this result following the proof in [11], page 477. We obtain that for each $\gamma\ge 2$ there exists a constant C such that

$$E|C_n^2|^\gamma\le C\frac{1}{n^\gamma}h_n^{-\frac{3}{2}}$$

It follows that $E|C_n^2| \leq \sqrt{E(|C_n^2|^2)} \leq \frac{\sqrt{C}}{nh_n^2} h_n^{\frac{5}{4}}$ which converges to zero. The remaining terms to estimate are C_n^3 and C_n^4. Note that in the expansion of C_n, C_n^3 is the most important term.

Lemma 8 (Study of C_n^3 and C_n^4) *If $nh_n^2 \to \infty$, then C_n^3 and C_n^4 converge to zero in L^1.*

Proof. It is straigthforward to obtain the estimate

$$E|C_3^n| \leq \frac{M}{h_n^3} E\Big(\sum_{i=0}^{n-1} |\phi''|(\frac{S_{\frac{i}{n}} - x}{h_n}) \int_{\frac{i}{n}}^{\frac{i+1}{n}} \frac{1}{2} E\big((S_s - S_{\frac{i}{n}})^2 \mid \mathcal{F}_{\frac{i}{n}}\big)\Big)$$

Now a BDG inequality implies the existence of a constant C such that $E((S_s - S_{\frac{i}{n}})^2 \mid \mathcal{F}_{\frac{i}{n}}) \leq C(s - \frac{i}{n})$. Hence

$$E|C_3^n| \leq \frac{MC}{2nh_n^2} E\big(\frac{1}{nh_n} \sum_{i=0}^{n-1} |\phi''|(\frac{S_{\frac{i}{n}} - x}{h_n})\big)$$

which converges to zero when $nh_n^2 \to \infty$. The same techniques as in this lemma and the previous one can be applied to prove the convergence of C_n^4 in L^1 to 0. ∎

We have a stronger result that the one stated in the lemma above. Under the assumption that $nh_n^2 \to \infty$, both C_n^3 and C_n^4 converge to zero in L^γ, for all $\gamma > 2$. Define

$$\tilde{C}_n^3 = \frac{1}{h_n^3} \sum_{i=0}^{n-1} \phi''(\frac{S_{\frac{i}{n}} - x}{h_n}) \int_{\frac{i}{n}}^{\frac{i+1}{n}} \frac{(S_s - S_{\frac{i}{n}})^2}{2} ds$$

Hoffman's result in [11] guarantees that $\tilde{C}_n^3$ converges in L^γ to 0 as soon as $nh_n^2 \to \infty$, for all $\gamma > 2$. Since σ and its derivatives are assumed to be bounded, it is not hard to see that we have the same result for C_n^3.

Putting all these lemmas together proves that C_n converges in L^1 to 0, and thus B_n converges in L^1 to 0. We have then proved that L_n^x converges to L^x in L^p, for all $p > 0$ and that $V_n^x - \sigma^2(x)L^x$ converges in L^1 to 0, which ends the proof of Lemma 2. The following theorem is now straightforward.

Theorem 3 *If $nh_n^2 \to \infty$ then $S_n^x = \frac{V_n^x}{L_n^x}$ converges in probability to $\sigma^2(x)$ and provides a consistent estimator of $\sigma^2(x)$.*

Remark 4 *After finding and proving this theorem, we learned to our chagrin, from Jean Jacod, that he had not only already considered this exact problem more than 10 years ago, but that he has also established similar results which are both more general and more effective. See [14] and [15]. In particular he is able to take $h_n = \frac{1}{\sqrt{n}}$ and he also obtains a rate of convergence and an associated Central Limit Theorem. We have decided nevertheless to retain our estimator and its proof presented here, since it is the one we used to process the data and it seems to works well for our purposes. But we wish to signal for future related work that there are more powerful (if perhaps slightly more complicated) similar estimators available. These remarks also apply for parts of Section 4.*

4 Unbounded Volatility Function Estimators

The previous two estimators for the volatility function $\sigma(x)$ are over a compact domain representing the observation interval. In this section, for the SDE (2), we relax this boundedness assumption on the volatility function $\sigma(x)$. Herein, we now assume that $\sigma > 0$ on $I =]0, \infty[$, it is identically null elsewhere and satisfies $\frac{1}{\sigma^2} \in L^1_{loc}(I)$.

This is the Engelbert Schmidt condition (see, e.g., [4] or [19]) under which the SDE has a unique weak solution S that does not explode to ∞. We let P be the law of the solution on the canonical space $\Omega = C([0,T], \mathbb{R})$ equipped with the canonical filtration $(\mathcal{F}_t)_{t\in[0,T]}$ and the canonical process $S = (S_t)_{t\in[0,T]}$. We also assume that σ is C^3 bounded and with bounded derivatives on every compact set. We add in passing that these hypotheses imply the existence of a strong solution, as well. Let $\tau_0(S)$ be the first time S hits zero. The following theorem provides straightforward but useful extensions of Theorem 1 and Lemma 2.

Theorem 5 *Suppose $\sigma(x)$ has three continuous derivatives. Assume that $nh_n^4 \to 0$ and $nh_n \to \infty$. Then conditional on $\{\tau_0(S) > T\}$, $S_n(x)$ given in (5) converges in probability to $\sigma^2(x)$. The same holds for our smooth kernel estimator under the constraint $nh_n^2 \to \infty$.*

Proof. Let $T_q = \inf\{t, S_t \geq q\}$ and $\tau_p = \inf\left\{t, S_t \leq \frac{1}{p}\right\}$. Then $\lim_{p\to\infty} \tau_p = \tau_0(S)$ and $\lim_{q\to\infty} T_q = \infty$ since S does not explode to ∞. We can take $\sigma_{p,q}$ to be a function bounded above and below away from zero with three bounded derivatives such that $\sigma_{p,q}(x) = \sigma(x)$ for all $\frac{1}{p} \leq x \leq q$. Let $(S_t^{p,q})_{t\in[0,T]}$ be the unique strong solution to the SDE $dS_t^{p,q} = \sigma_{p,q}(S_t^{p,q})dW_t$. Introduce now $S_n^{p,q}(x)$, the estimator computed on the basis of $(S_t^{p,q})_{t\in[0,T]}$ as in (5) or using our smooth kernel estimator. Then under the suitable constraints on the sequence $(h_n)_{n\geq 1}$, $S_n^{p,q}(x)$ converges in probability to $\sigma^2_{p,q}(x)$. Moreover $S_n^{p,q}(x) = S_n(x)$ if $T < T_q \wedge \tau_p$. Then obviously $S_n(x)$ converges in probability to $\sigma^2(x)$, in restriction to the set $\{T < \tau_0(S)\}$. ■

We can extend Theorem 2 by working in the filtration $\mathcal{G}_t = \sigma(S_s, s \leq t) \vee \sigma(\tau_p)$ where $\tau_p = inf\left\{t, S_t \in]\frac{1}{p}, p[\right\}$ and more interestingly in $\mathcal{G}_t^{'} = \sigma(S_s, s \leq t) \vee U$ where $U = 1_{\{\tau_p > T\}}$ and T is the time horizon. Consider for instance the initial enlargement $\mathcal{G}_t = \mathcal{F}_t \vee \sigma(\tau_p)$ where $\mathcal{F}_t = \sigma(S_s, s \leq t)$ and τ_p is the first exit time of S from $]\frac{1}{p}, p[$. τ_p is an $\mathcal{F}$ stopping time. Consider the filtered probability space $(\Omega, \mathcal{G}, Q, \mathbb{G})$ where $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$. Under some technical assumptions, and if the sequence h_n satisfies $h_n \to 0$, $nh_n \to \infty$ and $nh_n^3 \to 0$ then we can prove that $1_{\{\tau_p > t\}}\sqrt{N_a^n}(\frac{S_n(a)}{\sigma^2(a)} - 1)$ converges in distribution to $\sqrt{2}Z$ where Z is a standard normal random variable and $N_a^n = \sum_{i=1}^n 1_{\{|S_{\frac{i}{n}} - a| < h_n\}}$. The exact statement of the theorem and its proof can be provided to the interested reader upon request. We do not provide these here because the proof's technicalities use results from the initial enlargement of filtrations [13] (alternatively, see [26]) and subsequently require derivations of many estimates. We focus herein on the methodology for bubble detection and the illustrative tests using dotcom company stocks.

Remark 6 (In practice) *Note that this limit theorem can be applied if during the time* *interval $[0,T]$ the process does not hit 0. The limit theorem also provides us with a confidence interval for the volatility estimator.*

5 Detecting Bubbles

This section illustrates the use of the previous volatility function estimators for detecting asset price bubbles during the alleged dotcom internet stock price bubble episode of 1998 - 2001 (see [24]). As mentioned in the introduction, asset price bubbles have been characterized in frictionless, competitive, and continuous trading models using the abitrage-free martingale pricing technology (see [22], [23], [2], [9], [17] and [18]). As shown in Jarrow, Protter and Shimbo [17] and [18], for finite horizon economies a bubble exists if the price process under a risk neutral measure is a strict local martingale, and not a martingale. The difference between a martingale and a strict local martingale has been recently investigated by several authors. The following theorem has been proved using different techniques. A proof based on explosion time techniques is provided in [1]. Separating time techniques are used to prove a more general result in [25]. Kotani gives a PDE based proof in [20]. As mentioned previously, D. Kramkov has recently pointed out that this follows quite simply from Feller's test for explosions [21].

Theorem 7 *S is a martingale (has no price bubbles) if and only if $\int_\epsilon^\infty \frac{x}{\sigma^2(x)} ds = \infty$, for each $\epsilon > 0$.*

This theorem forms the basis for our bubbles testing methodology. Unfortunately, we are immediately faced with an "extrapolation problem." To see this problem, we note that the volatility function estimators presented in the previous sections provide estimates for $\sigma(x)$ only on a finite interval - those x that have been visited by the process. Given the available stock price data, we can not observe the tails of the volatility function necessary to check for the divergence of the integral in Theorem 7. To check for divergence, we must extrapolate from the observed domain of $\sigma(x)$ to the entire nonnegative real line.

We propose two extrapolation methods to overcome this problem. The first method is to use a parametric estimator as in [7], and a comparison theorem to conclude when the parametric and non parametric volatility estimators are similar. If similar, we extrapolate into the tail using the parametric form's asymptotic behavior. The second method is to use Reproducing Kernel Hilbert Spaces theory to extrapolate the volatility function in the 'best' possible way. We will quantify what we mean by 'best' below.

5.1 Method 1: Parametric Estimation

The appeal of using a parametric form for the volatility function $\sigma(x)$ is that we know the tails once the parameters have been estimated. For this estimation we choose a class of volatility functions large enough to include many of the forms used in practice (for example : power functions $\sigma(x) = \sigma x^\alpha$, where σ and α are the unknown parameters that we estimate). For this example of $\sigma(x) = x^\alpha$ we have the process S is a strict local martingale that is always strictly positive if $\alpha > 1$, and is a martingale if $\frac{1}{2} \leq \alpha < 1$ which however can assume the value 0. If $\alpha = 1$ we are in the case of geometric Brownian motion. We then also use our non-parametric estimators. If these estimators are comparable, we have a conclusive test for divergence of the volatility integral. If they are not comparable, then the test is inconclusive.

5.1.1 The Comparison Theorem

This section states and proves the comparison theorem.

Theorem 8 (Comparison Theorem) *Assume that $dS_t = \sigma(t, S_t)dW_t$ and that there exist two functions Σ and $\overline{\sigma}$ such that : for all t and x, $\overline{\sigma}(x) \leq \sigma(t,x) \leq \Sigma(x)$ and such that σ, Σ and $\overline{\sigma}$ are continuous, locally Hölder continuous with exponent $\frac{1}{2}$, then:*

(i) if for all $c > 0$, $\int_c^\infty \frac{x}{\Sigma^2(x)}dx = \infty$ then S is a martingale.

(ii) if there exists $c > 0$ such that $\int_c^\infty \frac{x}{\overline{\sigma}^2(x)}dx < \infty$ then S is a strict local martingale.

In order to prove this theorem, we need the following lemma (see [3]):

Lemma 9 *Let g be a concave function, α_i, $i = 1, 2$ be two continuous functions, locally Hölder continuous with exponent $\frac{1}{2}$ such that for all (x,t), $\alpha_1(x,t) \leq \alpha_2(x,t)$. Let $T > 0$ be fixed. We consider $dX_t^{\alpha_{1,2}} = \alpha_{1,2}(t, X_t^{\alpha_{1,2}})dW_t$ and $u_{1,2}(x,t) = E_{(x,t)}(g(X_T^{\alpha_{1,2}}))$. Then for all $x \in \mathbb{R}^+$ and $t \in [0,T]$, $u_1(x,t) \geq u_2(x,t)$.*

Proof. [**of Theorem 8**] **(i)** Since g(x)=x is concave, we can apply the previous lemma and get that for all (x,t), $u(x,t) \geq u^\Sigma(x,t)$. If $\int_c^\infty \frac{x}{\Sigma^2(x)}dx = \infty$, then by Theorem 7, $u^\Sigma(x,t) = x$, for all (x,t). Thus for all (x,t), $u(x,t) \geq x$. But, we know that S is positive local martingale and hence a super martingale by Jensen's lemma, thus for all (x,t), $u(x,t) \leq x$. This proves that : $E(S_T|S_t = x) = x$ and S is a martingale.

(ii) Let $T > 0$ be fixed, and $\overline{u}(x,t) = E(S_T^{\overline{\sigma}}|S_t^{\overline{\sigma}} = x)$. Let $c > 0$ such that $\int_c^\infty \frac{x}{\overline{\sigma}^2(x)}ds < \infty$. We know that $S^{\overline{\sigma}}$ is strict local martingale by Theorem 7 and $\overline{u}(x,t) \leq x$ and it exists t such that $\overline{u}(x,t) < x$. Again since $g(x) = x$ is concave, $x \geq \overline{u}(x,t) \geq u(x,t)$, for all (t,x), and there exists t such that $u(x,t) < \overline{u}(x,t) < x$. Hence, S is a strict local martingale. ■

5.1.2 Illustrative Examples

To illustrate this procedure, we use market price data from the alleged internet dotcom bubble (and beyond), from 1999 to 2005. As explained above, we can use the previous theorem as follows: first we choose a parametric form for the diffusion coefficient and we estimate the parameters as explained in [7] by Genon-Catalot and Jacod after choosing a contrast function to minimize. That is, we choose a parametric form $\sigma(\nu, x)$ where ν is the multidimensional parameter that need to be estimated and a contrast $f(G, x)$. Their estimator is defined as $\hat{\nu}_n = \arg\min \frac{1}{n}\sum_{i=1}^n f(\sigma^2(\nu, S_{t_{i-1}}), S_i^n)$ where $S_n^i = \sqrt{n}(S_{t_i} - S_{t_{i-1}})$.

Usual choices for the contrast function f in our one dimensional setting are $f_1(G,x) = \ln(G)+\frac{x^2}{G}$ or $f_2(G,x) = (x^2-G)^2$. We do not provide further details in this paper and refer the interested reader to [7] for a detailed description of the estimation procedure.

Then we estimate the volatility function using our non-parametric estimators. If the two volatility function estimates are similar, then by applying the criteria of Theorem 7 to the parametric estimator, we can test for the existence of a price bubble using the comparison theorem as in Theorem 8.

A Conclusive Test: When applied to the stock *Lastminute.com*, the methodology is conclusive.

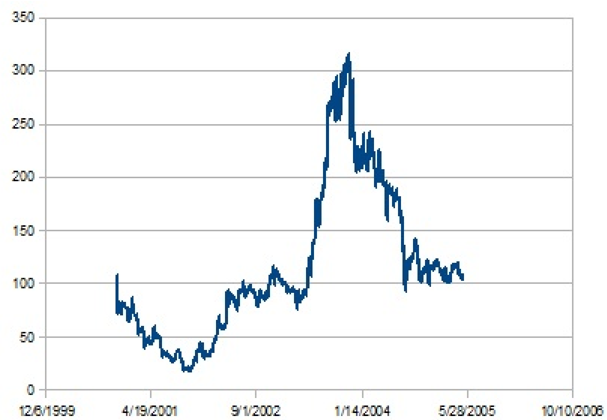

Figure 1: Lastminute.com Stock Prices during the alleged Dotcom Bubble.

Intuitively, given the stock price time series as given in Figure 1, one suspects the existence of a price bubble. Our test confirms this belief.

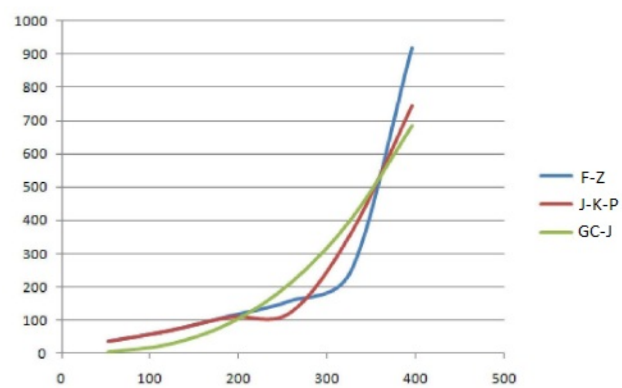

Figure 2: Lastminute.com. Estimates of $\sigma(x)$.

This can be seen from Figure 2 above that displays the estimators of Florens-Zmirou (F-Z), the smoothed kernel (J-K-P), and the parametric estimator of Genon-Catalot and Jacod (GC-J), using the power parametric form $\sigma(\sigma_0, \alpha, x) = \sigma_0 x^\alpha$ (here $\nu = (\sigma_0, \alpha)$

is a two-dimensional parameter) and the loglikelihood like contrast $f_1(G, x)$. Using this estimation technique, we find an estimate $\hat{\sigma}(x) = \sigma(\hat{\sigma}_0, \hat{\alpha}, x)$ whose tail behavior leads to the convergence of the integral $\int_\epsilon^\infty \frac{x}{\hat{\sigma}(x)^2} dx$. Also our estimator (J-K-P) lies above the estimated function (GC-J) hence Theorem 8 guarantess that the price process is a strict local martingale and we have bubble pricing.

An Inconclusive Test: A weakness of this procedure is that the comparison test using parametric estimators might be inconclusive, even when intuitively one suspects a bubble. An example of this phenomenon is that of *Etoys*.

Figure 3: eToys Stock Prices during the alleged Dotcom Bubble.

This stock price graph suggests the existence of a price bubble. Our methodology is illustrated in Figure 4.

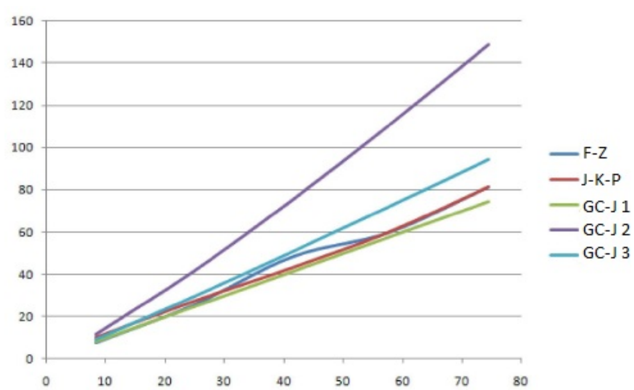

Figure 4: eToys. Estimates of $\sigma(x)$.

The three curves included in this figure and labelled GC-J1, GC-J2 and GC-J3 represent the estimators of Genon-Catalot and Jacod, with different parametric forms and contrasts to minimize. GC-J1 and GC-J3 are obtained from the parametric form $\sigma(\sigma_0, \alpha, x) = \sigma_0 x^\alpha$ and the contrasts f_1 and f_2 respectively. GC-J2 is obtained from $\sigma(\sigma_0, \alpha, \beta, x) = \sigma_0 x^\alpha \ln^\beta |x|$ and the contrast f_1. In theory, we have a bubble if $\alpha > 1$ or if $\alpha = 1$ and

$\beta > 1$, however the estimated parameters lie in these boundaries values and we see that the curves are so close to linear that we cannot conclude either convergence nor divergence of the integral in Theorem 4. Using this methodology, our test is inconclusive as to whether or not there was a bubble in the stock price of *eToys* during the 1999-2001 period.

5.2 Method 2: *RKHS* theory

This section presents our second method for extrapolation. This method is different, in that it is based on *RKHS* theory. Previously in the paper we have considered parameterized families of functions, so that once the parameter is chosen the tail behavior is determined. We can observe the volatility coefficient σ only on a bounded interval, of course, so it is a leap of faith to assume that (a) it is of the form of the parameterized family of functions considered, and (b) its behavior continues unchanged into the tail. Nevertheless, this is more or less the standard technique in situations such as this.

Our second method is a bit more subtle. Our procedure here consists of two steps:

- We first interpolate an estimate of σ within the bounded interval where we have observations, and in this way we lose the irregularities of non parametric estimators;
- We next extrapolate our function σ by choosing a *RKHS* from a family of Hilbert spaces in such a way as to remain as close as possible (on the bounded interval of observations) to the interpolated function provided in the previous step.

This represents a new methodology which allows us to choose a *good* extrapolation method. We do this via the choice of a certain extrapolating RKHS, which – once chosen – again determines the tail behavior of our volatility σ. If we let $(H_m)_{m\in\mathbb{N}}$ denote our family of RKHS, then any given choice of m, call it m_0, allows us to interpolate *perfectly* the original estimated points, and thus provides a valid *RKHS* H_m with which we extrapolate σ. But this represents a choice of m_0 and not an estimation. So if we stop at this point the method would be as arbitrary as parametric estimation. That is, choosing m_0 is analogous to choosing the parameterized family of functions which fits σ best. The difference is that we do not arbitrarily choose m_0. Instead we choose the index m *given the data available.* In this sense we are using the data twice. To do this we evaluate different RKHSs is order to find the most appropriate one *given the arrangement of the finite number of grid points* from our observations.

The *RKHS* method (see [12]) is intimately related to the reconstruction of functions from scattered data in certain linear functional spaces. The reproducing kernel $Q(x, x')$ that is associated with an *RKHS* $H(\mathcal{D})$ in the spatial domain $\mathcal{D}$, over the coordinate x, is unique and positive and thus constitutes a natural basis for generic interpolation problems.

5.2.1 Reproducing Kernel Hilbert Spaces

Let $H(\mathcal{D})$ be a Hilbert space of continuous real valued functions $f(x)$ defined on a spatial domain $\mathcal{D}$. A reproducing kernel Q possesses many useful properties for data interpolation and function approximation problems.

Proposition 1 *There exists a kernel function $Q(x, x^{'})$, the reproducing kernel, in $H(\mathcal{D})$ such that the following properties hold:*

(i) ***Reproducing property.*** *For all x and y,*

$$\begin{aligned} f(x) &= \langle f(x^{'}), Q(x, x^{'})\rangle^{'} \\ Q(x, y) &= \langle Q(x, x^{'}), Q(y, x^{'})\rangle^{'}. \end{aligned}$$

The prime indicates that the inner product $\langle\cdot,\cdot\rangle^{'}$ is performed over $x^{'}$.

(ii) ***Uniqueness.*** *The RKHS $H(\mathcal{D})$ has one and only one reproducing kernel $Q(x, x^{'})$.*

(iii) ***Symmetry and Positivity.*** *The reproducing kernel $Q(x, x^{'})$ is symmetric, i.e. $Q(x^{'}, x) = Q(x, x^{'})$, and positive definite, i.e.:*

$$\sum_{i=1}^{n}\sum_{k=1}^{n} c_i Q(x_i, x_k) c_k \geq 0$$

for any set of real numbers c_i and for any countable set of points $(x_i)_{i\in[1,n]}$.

In this framework, interpolation is seen as an inverse problem. The inverse problem is the following. Given a set of real valued data $(f_i)_{i\in[1,M]}$ at M distinct points $S_M = x_i, i \in [1, M]$ in a domain $\mathcal{D}$, and a *RKHS* $H(\mathcal{D})$, find a suitable function $f(x)$ that interpolates these data points. Using the reproducing property, this interpolation problem is reduced to solving the following linear inverse problem :

$$\forall i \in [1, M], f(x_i) = \langle f(x^{'}), Q(x_i, x^{'})\rangle^{'} \tag{8}$$

where we need to invert this relation and exhibit the function $f(x)$ in $H(\mathcal{D})$. We refer the reader to [12] for a detailed discussion.

We first present the normal solution that allows an exact interpolation, and second the regularized solution that yields quasi interpolative results, accompanied by an error bound analysis. Then in the next section, we will construct a family of *RKHS*'s that enable us to interpolate not $\sigma(x)$ but $\frac{1}{\sigma(x)^2}$. This transformation makes natural the choice of the family of *RKHS*'s. Note that for every choice of an *RKHS*, on can construct an interpolating function using the input data. For this reason, we define a family of Reproducing Kernel Hilbert Spaces that encapsulate different assumptions on the asymptotic forms and smoothness constraints. From this set, we choose that *RKHS* which best fits the input data in the sense explained below.

Normal Solutions: The most straightforward interpolation approach is to find the normal solution that has the minimal squared norm $||f||^2 = \langle f(x^{'}), f(x^{'})\rangle^{'}$ subject to the interpolation condition (8).

That is, given a set of real valued data $\{f_i\}, 1 \leq i \leq K$ specified at K distinct points in a domain $\mathcal{D}$, we wish to find a function f that is the normal solution:

$$f(x) = \sum_{i=1}^{M} c_i Q(x_i, x)$$

where the coefficients c_i satisfy the linear relation :

$$\forall k \in [1, M], \sum_{i=1}^{M} c_i Q(x_i, x_k) = f_k. \tag{9}$$

If the matrix Q_M whose entries are the $Q(x_i, x_k)$ is "well conditioned," then the linear algebraic system above can be efficiently solved numerically. Otherwise, we use regularized solutions.

Regularized Solutions: When the matrix Q_M is "ill conditioned," regularization procedures may be invoked for approximately solving the linear inverse problem. In particular, the Tikhonov regularization procedure produces an approximate solution f_α, which belongs to $H(\mathcal{D})$ and that can be obtained via the minimization of the regularization functional

$$||Qf - F||^2 + \alpha ||f||^2$$

with respect to $f(x)$. Note that here F is the data vector (f_i) and the residual norm $||Qf - F||^2$ is defined as:

$$||Qf - F||^2 = \sum_{i=1}^{M} (\langle f(x'), Q(x_i, x') \rangle' - f_i)^2.$$

The regularization parameter α is chosen to impose a proper balance between the residual constraint $||Qf - F||$ and the magnitude constraint $||f||$. The regularized solution has the form

$$f_\alpha(x) = \sum_{i=1}^{M} c_i^\alpha Q(x_i, x) \tag{10}$$

where the coefficients c_i^α satisfy the linear relation:

$$\forall k \in [1, M], \sum_{i=1}^{M} c_i^\alpha (Q(x_i, x_k) + \alpha \delta_{i,k}) = f_k \tag{11}$$

where $\delta_{i,k}$ is the Kronecker delta function. Note that for $\alpha > 0$, Q_M^α whose entries are $[Q(x_i, x_k) + \alpha \delta_{i,k}]$ is symmetric and positive definite and the problem can now be solved efficiently. Also, the *RKHS* interpolation method leads to an automatic error estimate of the regularized solution (see [12] for more details).

5.2.2 Construction of the Reproducing Kernels

We consider reciprocal power reproducing kernels that asymptotically behave as some reciprocal power of x, over the interval $[0, \infty[$. We are interested in this type of *RKHS* because this is a reasonable assumption for $f(x) = \frac{1}{\sigma^2(x)}$. The CEV model $dS_t = S_t^\alpha dW_t$ where $\alpha > 0$ is an important local volatility model proposed in the literature and satisfies this assumption, with $f_{cev}(x) = \frac{1}{x^{2\alpha}}$. We also assume that the function $f(x)$ possesses the asymptotic property

$$\lim_{x\to\infty} x^k f^{(k)}(x) = 0, \forall k \in [1, n-1].$$

for some $n \geq 1$ that controls the minimal required regularity. This property is often satisfied by the volatility functions used in practice. For instance, $x^k f_{cev}^{(k)}(x) = \frac{\prod_{i=0}^{k-1}(-2\alpha-i)}{x^{2\alpha}}$ converges to 0 as x tends to infinity, for all k. This is also satisfied by many volatility functions that explode faster than any power of x, for example $\sigma(x) = x^\alpha e^{\beta x}$, with $\alpha > 0$ and $\beta > 0$. The condition appears restrictive only when σ and its derivatives explode too slowly or when σ is bounded, however in these cases, it is likely that there is no bubble and no extrapolation using this *RKHS* theory will be required. We would like to emphasize that the asymptotic property satisfied by f is the key point for the whole method to work as this may be seen from Proposition 3 below.

Concerning the degree of smoothness, we usually take in practice n to be 1, 2 or 3. We can define now our Hilbert space

$$H_n = H_n([0, \infty[) = \left\{ f \in C^n([0, \infty[) \mid \lim_{x\to\infty} x^k f^{(k)}(x) = 0, \forall k \in [1, n-1] \right\}.$$

We now need to define an inner product. A smooth reproducing kernel $q^{RP}(x, x')$ can be constructed via the choice:

$$< f, g >_{n,m} = \int_0^\infty \frac{y^n f^{(n)}(y)}{n!} \frac{y^n g^{(n)}(y)}{n!} \frac{dy}{w(y)}$$

where $w(y) = \frac{1}{y^m}$ is the asymptotic weighting function. From now on we consider the *RKHS* $H_{n,m} = (H_n, <,>_{n,m})$. The next proposition can be shown following the steps in [12].

Proposition 2 *The reproducing kernel is given by*

$$q_{n,m}^{RP}(x, y) = n^2 x_>^{-(m+1)} B(m+1, n) F_{2,1}(-n+1, m+1, n+m+1, \frac{x_<}{x_>})$$

where $x_>$ and $x_<$ are respectively the larger and smaller of x and y, $B(a, b)$ is the beta function and $F_{2,1}(a, b, c, z)$ is Gauss's hypergeometric function.

Remark 9 *The integers $n-1$ and $m+1$ are respectively the order of smoothness and the asymptotic reciprocal power behavior of the reproducing kernel $q^{RP}(x, y)$. This kernel is a rational polynomial in the variables x and y and has only a finite number of terms, so it is computationally efficient.*

As pointed out above, any choice of n and m creates an *RKHS* $H_{n,m}$ and allows one to construct an interpolating function $f_{n,m}(x)$ with a specific asymptotic behavior. The following result gives the exact asymptotic behavior.

Proposition 3 *For every x, $q^{RP}(x,y)$ is equivalent to $\frac{n^2}{y^{m+1}}B(m+1,n)$ at infinity as a function of y and*

$$\lim_{x\to\infty} x^{m+1} f_\alpha(x) = n^2 B(m+1,n) \sum_{i=1}^{M} c_i^\alpha$$

where f_α is defined as in (10) *and the constants c_i^α are obtained as in* (11)*. Hence, if $\sum_{i=1}^M c_i^\alpha \neq 0$, then $f_\alpha(x)$ is equivalent to $\frac{n^2 B(m+1,n)}{x^{m+1}} \sum_{i=1}^M c_i^\alpha$.*

5.2.3 Choosing the Best m

The choice of m allows one to decide if the integral in Theorem 7 converges or diverges. If $m > 1$, there is a bubble. This section explains how to choose m. Let us first summarize the idea. We choose the *RKHS* by optimizing over the asymptotic weight m that allows us to construct a function that interpolates the input data points and remains as close as possible to the interpolated function on the finite interval $\mathcal{D}$. This optimization provides an $\overline{m}$ which allows us to construct $\sigma_{\overline{m}}(x)$. We employ a four step procedure:

(i) Non-parametric estimation over $\mathcal{D}$: Estimate $\sigma(x)$ using our non-parametric estimator on a fixed grid $x_1, ..., x_M$ of the bounded interval $\mathcal{D} = [\min S, \max S]$ where $\min S$ and $\max S$ are the minimum and the maximum reached by the stock price over the estimation time interval $[0,T]$. In our illustrative examples, we use the kernel $\phi(x) = \frac{1}{c} e^{\frac{1}{4x^2-1}}$ for $|x| < \frac{1}{2}$, where c is the appropriate normalization constant. The number of data available n and the restriction on the sequence $(h_n)_{n\geq 1}$ makes the number of grid points M relatively small in practice. In our numerical experiments, $7 \leq M \leq 25$.

(ii) Interpolate $\sigma(x)$ over $\mathcal{D}$ using *RKHS* theory: Use any interpolation method on the finite interval $\mathcal{D}$ to interpolate the data points $(\sigma(x_i))_{i\in[1,M]}$. Call the interpolated function $\sigma^b(x)$. For completeness, we provide a methodology to achieve this using the *RKHS* theory. However, any alternative interpolation procedure for a finite interval could be used.

Define the Sobolev space: $H^n(\mathcal{D}) = \left\{u \in L^2(\mathcal{D}) \mid \forall k \in [1,n], u^{(k)} \in L^2(\mathcal{D})\right\}$ where $u^{(k)}$ is the weak derivative of u. The norm that is usually chosen is $||u||^2 = \sum_{k=0}^n \int_{\mathcal{D}} (u^{(k)})^2(x)dx$. Due to Sobolev inequalities, an equivalent and more appropriate norm is $||u|| = \int_{\mathcal{D}} u^2(x)dx + \frac{1}{\tau^{2n}} \int_{\mathcal{D}} (u^{(n)})^2(x)dx$. We denote by $K_{n,\tau}^{a,b}$ the kernel function of $H^n(]a,b[)$, where in this case $\mathcal{D} =]a,b[$. This reproducing kernel is provided for $n=1$ and $n=2$ in the following lemma.

Lemma 10

$$K_{1,\tau}^{a,b}(x,y) = \frac{\tau}{\sinh(\tau(b-a))} \cosh(\tau(b-x_>)) \cosh(\tau(x_< - a))$$
$$K_{2,\tau}^{a,b}(x,y) = L_{x_>}(x_<)$$

and $L_x(t)$ is of the form $\sum_{i=1}^{4}\sum_{k=1}^{4} l_{ik} b_i(\tau t) b_k(\tau x)$.

We refer to [28, Equation (22) and Corollary 3 on page 28] for explicit analytic expressions for l_{ik} and b_k, which while simple, are nevertheless tedious to write. In both equalities, $x_>$ and $x_<$ respectively stand for the larger and smaller of x and y. In practice, one should check the quality of this interpolation and carefully study the outputs by choosing different τ's before using the interpolated function $\sigma^b = \frac{1}{\sqrt{f^b}}$ in the algorithm detailed above, where $f^b(x) = \sum_{i=1}^{M} c_i^b K_{n,\tau}^{\mathcal{D}}(x_i, x)$, for all $x \in \mathcal{D}$ and for all $k \in [1, M]$, $\sum_{i=1}^{M} c_i^b K_{n,\tau}^{\mathcal{D}}(x_i, x_k) = f_k = \frac{1}{\sqrt{\sigma^{est}(x_k)}}$.

(iii) Deciding if an extrapolation is required: If the extended form of the estimated $\sigma(x)$ implies that the volatility does not diverge to ∞ as $x \to \infty$ and remains bounded on $\mathbb{R}^+$, no extrapolation is required. In such a case $\int_\epsilon^\infty \frac{x}{\sigma^2(x)}$ is infinite and the process is a true martingale. If one decides, however, that $\sigma(x)$ diverges to ∞ as $x \to \infty$, then the next step is required to obtain a 'natural' candidate for its asymptotic behavior as a reciprocal power.

(iv) Extrapolate $\sigma^b(x)$ to $\mathbb{R}^+$ using *RKHS*: Fix $n = 2$ and define

$$\overline{m} = \arg\min_{m \geq 0} \sqrt{\int_{[a,\infty[\cap\mathcal{D}} |\sigma_m - \sigma^b|^2 ds} \tag{12}$$

where $f_m = \frac{1}{\sigma_m^2}$ is in the *RKHS* $H_{2,m} = (H_{2,m}([0,\infty[), \langle,\rangle_{RP})$. By definition, all σ_m will interpolate the input data points and $\sigma_{\overline{m}}$ has the asymptotic behavior that best matches our function on the estimation interval. a is the threshold determining closeness to the interpolated function. Choosing a too small is misleading since then it would account more (and unnecessarily) for the interpolation errors over the finite interval $\mathcal{D}$ than desirable. We should choose a large a since we are only interested in the asymptotic behavior of the volatility function. In the illustrative examples below, the threshold a in (12) is chosen to be $a = \max S - \frac{1}{3}(\max S - \min S)$.

5.2.4 Illustrative Examples

We illustrate our testing methodology for price bubbles using the stocks that are often alleged ([29] and [24]) as experiencing internet dotcom bubbles. We consider those stocks for which we have tick data. The data was obtained from WRDS [30]. We apply this methodology to four stocks: Lastminute.com, eToys, Infospace, and Geocities. The methodology performs well. The weakness of the method is the possibility of inconclusive tests as illustrated by *eToys*. For *Lastminute.com* and *Infospace* our methodology supports the existence of a price bubble. For *Infospace*, we reproduce the methodology step-by-step. Finally, the study of *Geocities* provides a stock commonly believed to have exhibited a bubble (see for instance [29] and [24]), but for which our method says it did not. We now provide our analyzes.

Lastminute.com: Our methodology confirms the existence of a bubble. The stock prices are given in Figure 5.

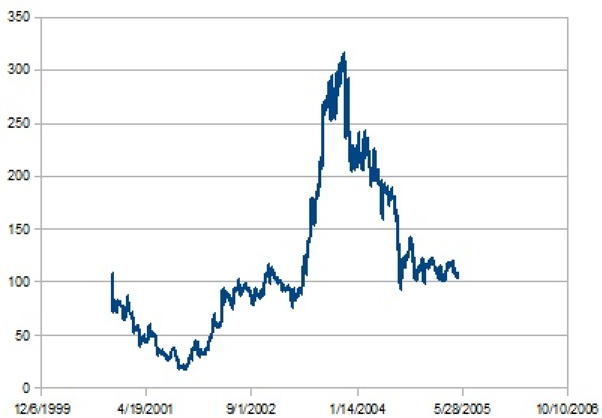

Figure 5: Lastminute.com Stock Prices during the alleged Dotcom Bubble.

The optimization performs as expected with the asymptotic behavior given by $\overline{m} = 8.26$, which means that $\sigma(x)$ is equivalent at infinity to a function proportional to x^{α} with $\alpha = 4.63$. We plot in Figure 6 the different extrapolations obtained using different reproducing kernel Hilbert spaces $H_{2,m}$ and their respective reproducing kernels $q_{2,m}^{RP}$.

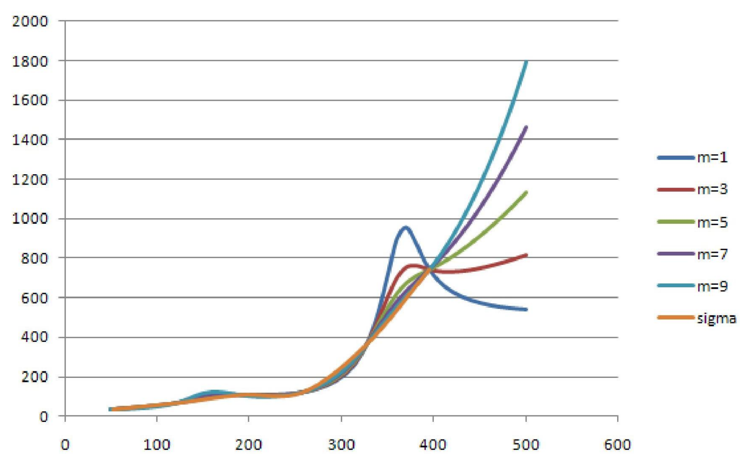

Figure 6: Lastminute.com. RKHS estimates of $\sigma(x)$.

Figure 6 shows that m is between 7 and 9 as obtained by the optimization procedure. The orange curve labelled (sigma) is the interpolation on the finite interval $\mathcal{D}$ obtained from the non-parametric estimation procedure where the interpolation is achieved using the *RKHS* theory as described in step **(ii)** with the choice of the repoducing kernel Hilbert space $H^1(\mathcal{D})$ and the reproducing kernel $K_{1,6}^{\min S, \max S}$. Then m is optimized as in step **(iv)** so that the interpolating function $\sigma_{\overline{m}}(x)$ is as close as possible to the orange curve in the last third of the domain $\mathcal{D}$, i.e. the threshold a in (12) is chosen to be $a = \max S - \frac{1}{3}(\max S - \min S)$.

eToys: While the graph of the stock price of eToys as given in Figure 7 makes the existence of a bubble plausible, *the test nevertheless is inconclusive*. Different choices of m giving different asymptotic behaviors are all close to linear (see Figure 8).

Figure 7: Etoys.com Stock Prices during the alleged Dotcom Bubble.

Because they are so close to being linear, we cannot tell with any level of assurance that the integral in question diverges, or converges. We simply cannot decide which is the case. If it were to diverge we would have a martingale (and hence no bubble), and were it to converge we would have a strict local martingale (and hence bubble pricing).

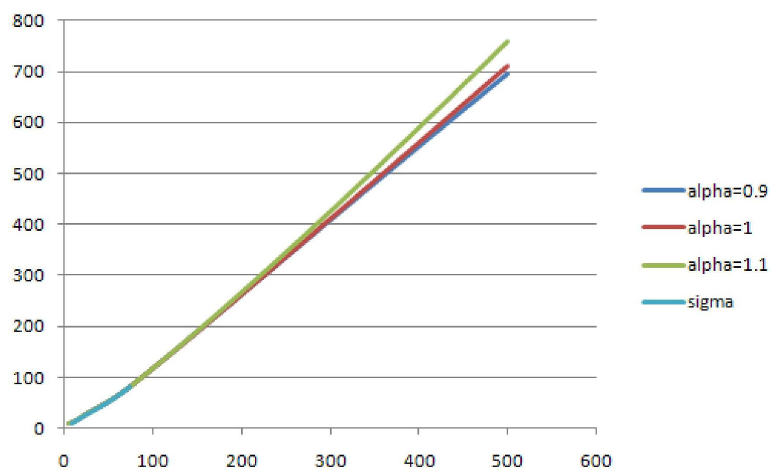

Figure 8: eToys. RKHS estimates of $\sigma(x)$.

The estimated $\overline{m}$ is close to one. In Figure 8, the powers α are given by $\frac{1}{2}(m+1)$ where m is the weight of the reciprocal power used to define the Hilbert space and its inner product. We plot the extrapolated functions obtained using different Hilbert spaces

$H_{2,m}$ together with their reproducing kernels $q_{2,m}^{RP}$. Figure 9 shows that the extrapolated functions obtained using these different $RKHS$ $H_{2,m}$ produce the same quality of fit on the domain $\mathcal{D}$.

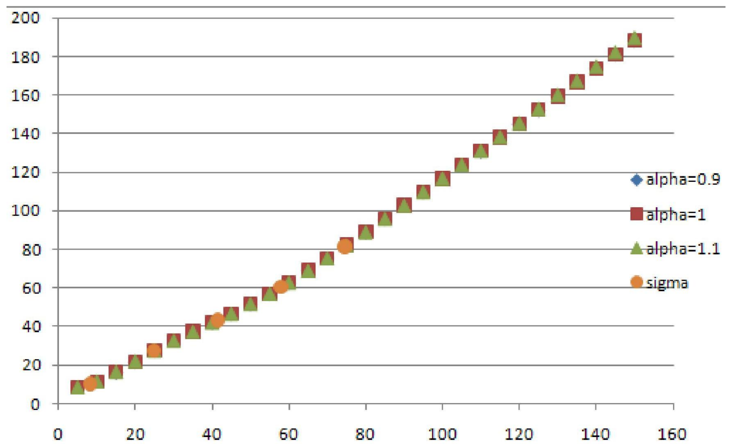

Figure 9: eToys. RKHS estimates of $\sigma(x)$, Quality of Fit.

Infospace: Our methodology shows that Infospace exhibited a price bubble. We detail the methodology step by step in this example. The graph of the stock prices in Figure 10 suggests the existence of a bubble.

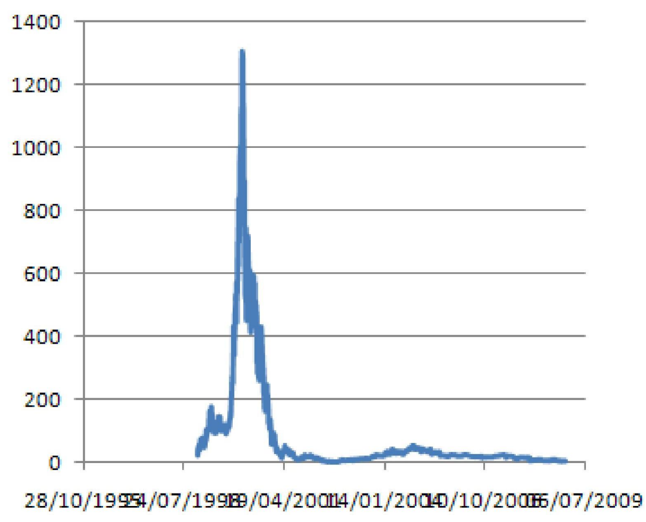

Figure 10: Infospace Stock Prices during the alleged Dotcom Bubble.

(i) We compute the Florens-Zmirou's estimator and our smooth kernel local time based estimator, using a sequence $h_n = \frac{1}{n^{\frac{1}{3}}}$. The result is not smooth enough as seen in Figure 11.

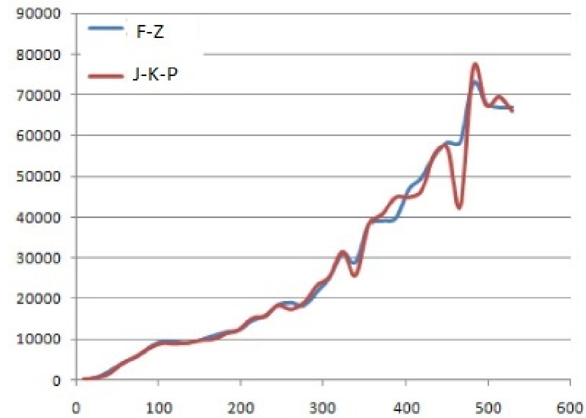

Figure 11: Infospace. Non-parametric Estimation using $h_n = \frac{1}{n^{\frac{1}{3}}}$.

(ii) We use the sequence $h_n = \frac{1}{n^{\frac{1}{4}}}$ to compute our estimators (the number of points where the estimation is performed is smaller, M=11). Theoretically, we no longer have the convergence of the Florens-Zmirou's estimator. However, as seen in Figure 12, this estimator is robust with respect to the constraint on the sequence h_n. F-Z, LowerBound and UpperBound are Florens-Zmirou's estimator together with the 95% confidence bounds her estimation procedure provides. J-K-P is our estimator.

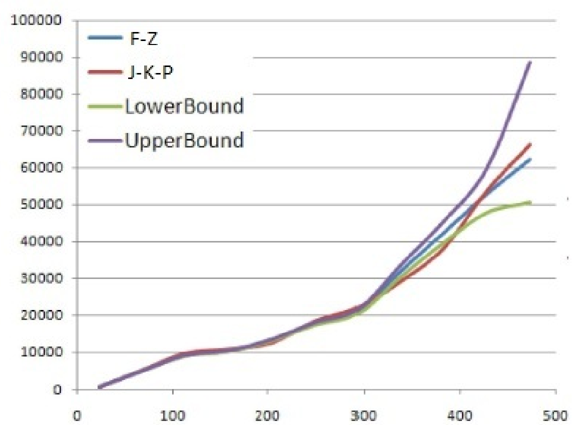

Figure 12: Infospace. Non-parametric Estimation using $h_n = \frac{1}{n^{\frac{1}{4}}}$.

(iii) We obtained in (ii) estimations on a fixed grid containing $M = 11$ points, and we now construct a function $\sigma^b(x)$ on the finite domain (see Figure 13) which perfectly interpolates those points. Here the *RKHS* used is $H^1(\mathcal{D})$ where $\mathcal{D} = [\min S, \max S]$ together with the reproducing kernels $K^{\mathcal{D}}_{1,\tau}$, where τ takes the values 1, 3, 6 and 9. The functions obtained using these different reproducing kernels provide the same quality of fit within $\mathcal{D}$ and we can use any of the four outputs as the interpolated function, σ^b, over the finite interval $\mathcal{D}$.

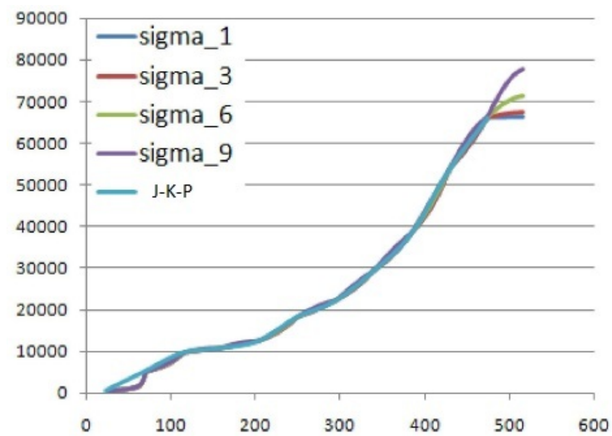

Figure 13: Infospace. Interpolation $\sigma^b(x)$ on the compact domain.

(iv) Finally we optimize over m and find the *RKHS* $H_{2,m}$ that allows the best interpolation of the $M = 11$ estimated points and such that the extrapolated function $\overline{\sigma}(x)$ remains as close as possible to $\sigma^b(x)$ on the third right side of $\mathcal{D}$. Of course, the reproducing kernels used in order to construct the functions σ_m and minimize the target error as in (12) are $q^{RP}_{2,m}$. We obtain $\overline{m} = 6.17$ (i.e. $\alpha = \frac{\overline{m}+1}{2} = 3.58$) and we can conclude that there is a bubble.

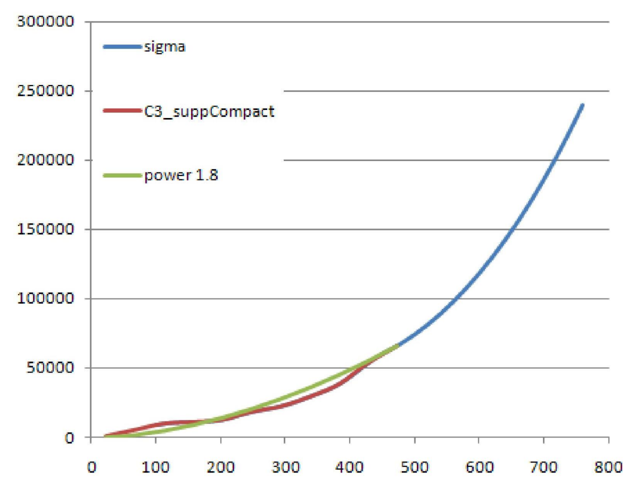

Figure 14: Infospace. Final estimator and RKHS Extrapolation.

Remark 10 *One might expect $\alpha \approx 1.8$ as suggested by the green curve in Figure 14. But this is different from what the RKHS extrapolation has selected. Why? In Figure 14, we plot the RKHS extrapolation obtained when $\alpha = 1.8$. We have proved that*

$$\lim_{x\to\infty} \frac{x^{m+1}}{\overline{\sigma}^2(x)} = 4B(m+1,2)\sum_{i=1}^{M} c_i.$$

The numerical computations give: $\overline{\sigma}(x) \approx \frac{x^{3.58}}{127009}$ *when using optimization over* m *and* $\overline{\sigma}(x) \approx \frac{x^{1.8}}{5.66}$ *when fixing* $\alpha = 1.8$*. Independent of the power chosen, the* c_i*'s and hence the constant of proportionality are automatically adjusted to interpolate the input points. But, as can be seen in Figure 15, the power 3.58 is more consistent in terms of extending 'naturally' the behavior of* $\sigma^b(x)$ *to* $\mathbb{R}^+$*.*

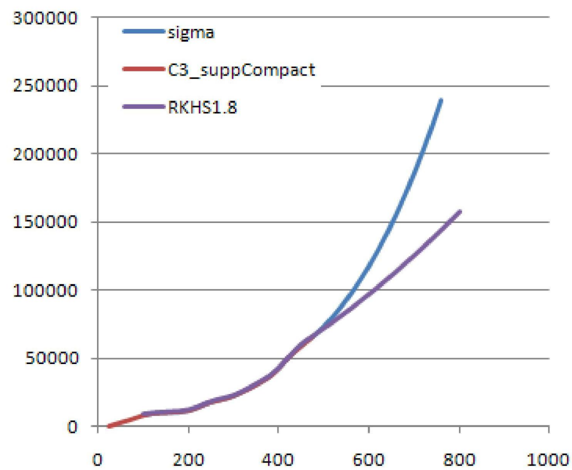

Figure 15: Infospace

Geocities: Our methodology shows that this stock did not have a price bubble. The stock prices are graphed in Figure 16.

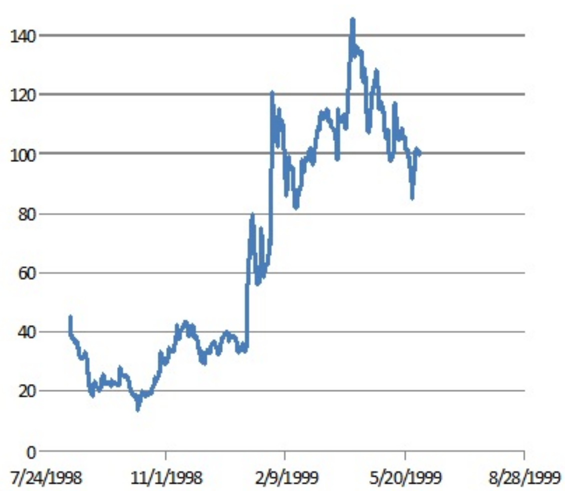

Figure 16: Geocities Stock Prices during the alleged Dotcom Bubble.

This is an example where we can stop at step (iii) : we do not need to use *RKHS* theory to extrapolate our estimator in order to determine its asymptotic behavior. As seen from

Figure 17, the volatility is a nice bounded function, and any natural extension of this behavior implies the divergence of the integral $\int_\epsilon^\infty \frac{x}{\sigma^2(x)}dx$. Hence the price process is a true martingale.

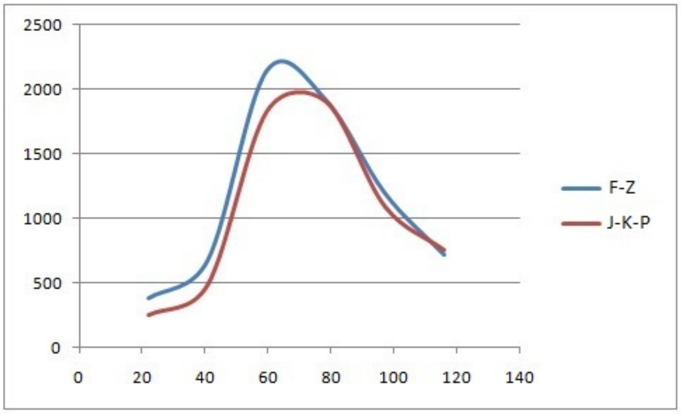

Figure 17: Geocities. Estimates of σ.

6 Conclusion

Given the price process of a risky asset that follows a stochastic differential equation under the risk neutral measure of the form

$$dX_t = \sigma(X_t)dW_t$$

where W is a standard one dimensional Brownian motion, we provide methods for estimating the volatility coefficient $\sigma(x)$ at the values where it is observed. If the behavior of $\sigma(x)$ is reasonable, we extend this estimator to all of $\mathbb{R}_+$ via the technology of Reproducing Kernel Hilbert Spaces. Having done this, we are then able to decide on the convergence or the divergence of the integral

$$\int_\epsilon^\infty \frac{x}{\sigma(x)^2}dx,$$

for any $\epsilon > 0$, which in turn determines whether or not the risky price process is experiencing, or has experienced, a bubble. Unfortunately, the test does not always work, since it depends on the behavior of $\sigma(x)$.

We illustrate our methodology using data from the alleged internet dotcom bubble of 1998-2001. Not surprisingly, we find that all three eventualities occur: in one case we are able to confirm the presence of a bubble; in a second case we confirm the lack of a bubble, and in a third case we find that the test is inconclusive. It is our hope that our methodology opens some new avenues for the testing of stock price bubbles in real time.

Acknowledgements

We wish to thank Jean Jacod for giving us a lot of help and advice in the revision of this paper, and in particular for alerting us to reference [15] and indicating how we could shorten and improve some of the proofs in the paper.

References

[1] P. Carr, A. Cherny and M. Urusov, On the Martingale Property of Time Homogeneous Diffusions, Preprint, http://www.math.tu-berlin.de/÷urusov/papers/cev.pdf, 2007.

[2] A. M. G. Cox and D. G. Hobson, 2005, Local martingales, bubbles and option prices, *Finance and Stochastics*, 9 (4), 477 - 492.

[3] E. Ekström and J. Tysk, Bubbles, Convexity and the Black-Scholes Equation, *Annals of Applied Probability*, **19**, 2009, 1369–1384

[4] H.J. Engelbert and W. Schmidt, Strong Markov continuous local martingales and solutions of one-dimensional stochastic differential equations, Parts I, II, and III, *Math. Nachr.*, **143**, 1989, 167-184; **144**, 1989, 241-281; **151**, 1991, 149-197.

[5] F. Delbaen and W. Schachermayer. The fundamental theorem of asset pricing for unbounded stochastic processes, *Math. Ann.*, 312:215–250, 1998.

[6] D. Florens-Zmirou, On Estimating the Diffusion Coefficient from Discrete Observations, *Journal of Applied Probability*, **30**, 1993, 790-804.

[7] V. Genon-Catalot and J. Jacod, *On the Estimation of the Diffusion Coefficient for Multi Dimensional Diffusion Processes*, Annales de l'I.H.P, section B, tome 29, 1993, p.119-151

[8] Jacob Goldstein, interview with William Dudley, the President of the New York Federal Reserve; in *Planet Money*, April 9, 2010

[9] S. Heston, M. Loewenstein, and G. A. Willard, 2007, Options and bubbles, *Review of Financial Studies*, preprint.

[10] P. Hitczenko, Best Constants in Martingale Version of Rosenthal's Inequality, *Annals of Probability*, **18**, 1990, 1656-1668.

[11] M. Hoffmann, L_p Estimation of the Diffusion Coefficient, *Bernoulli*, **5**, 1999, 447-481.

[12] T. Hollebeek, T.S. Ho and H. Rabitz, Constructing Multidimensional Molecular Potential Energy Surfaces From AB Initio Data, *Annu. Rev. Phys. Chem.*,1999, **50**, 537-570.

[13] J. Jacod, Grossissement Initial, Hypothèse H' et Théorème de Girsanov, *Grossissements de filtrations: exemples et application*, Lecture Notes in Math, **1118**, 1985, 15-35.

[14] J. Jacod, Rates of Convergence to the Local Time of a Diffusion, *Annales de L'Institut Henri Poincaré, Section B*, **34**, 1998, 505-544.

[15] J. Jacod, Non-parametric Kernel Estimation of the Coefficient of a Diffusion, *Scandinavian Journal of Statistics*, **27**, 2000, 83-96.

[16] R. Jarrow, Y. Kchia and P. Protter, On Detecting an Asset Bubble, Preprint 2010

[17] R. Jarrow, P. Protter and K. Shimbo, Asset Price Bubbles in Complete Markets, *Advances in Mathematical Finance*, Springer-Verlag, M.C. Fu et al, editors, 2007, 97–122.

[18] R. Jarrow, P. Protter and K. Shimbo, Asset Price Bubbles in Incomplete Markets, *Mathematical Finance*, **20**, 2010, 145-185.

[19] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, 1991, Springer-Verlag.

[20] S. Kotani, On a Condition that One Dimensional Diffusion Processes are Martingales, *In Memoriam Paul-André Meyer*, Séminaire de Probabilitiés XXXIX, Lecture Notes in Math **1874**, 2006, 149-156.

[21] D. Kramkov, Discussion on *How to Detect an Asset Bubble, by R. Jarrow, Y. Kchia, and P. Protter*, Remarks given after the presentation of P. Protter at the meeting *Contemporary Issues and New Directions in Quantitative Finance,* 9^{th} -11^{th} *July 2010*; Oxford, England, July 10, 2010.

[22] M. Loewenstein and G. A. Willard, 2000, Rational equilibrium asset-pricing bubbles in continuous trading models, *Journal of Economic Theory*, 91 (1), 17- 58.

[23] M. Loewenstein and G. A. Willard, 2000, Local martingales, arbitrage and viability: free snacks and cheap thrills, *Economic Theory*, 16, 135 - 161.

[24] J. Markham, *A Financial History of Modern U.S. Corporate Scandals: From Enron to Reform*, M.E. Sharpe, 2005.

[25] A. Mijatovic and M.Urusov, On the Martingale Property of Certain Local Martingales, Arxiv preprint arXiv:0905.3701, 2009 - arxiv.org

[26] P. Protter, *Stochastic Integration and Differential Equations, Second Edition, Version 2.1*, 1995, Springer-Verlag, Heidelberg.

[27] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion, Second Edition*, 1994, Springer-Verlag.

[28] C. Thomas-Agnan, Computing a Family of Reproducing Kernels for Statistical Applications, *Numerical Algorithms*, **13**, 1996, 21-32.

[29] Wikipedia page *http://en.wikipedia.org/wiki/Dot-com_bubble*

[30] WRDS, Lastminute.com *bid prices, from 14 March 2000 to 30 July 2004*
WRDS, eToys, *bid prices, from 20 May 1999 to 26 February 2001*
WRDS, Infospace, *bid prices, from 15 December 1998 to 19 September 2002*

WRDS, Geocities, *bid prices, from 11 August 1998 to 28 May 1999*