HAL open science

Data and methods for A threshold model for local volatility: evidence of leverage and mean reversion effects on historical data

Antoine Lejay, Paolo Pigato

To cite this version:

Antoine Lejay, Paolo Pigato. Data and methods for A threshold model for local volatility: evidence of leverage and mean reversion effects on historical data. [Research Report] RT-0494, Inria Nancy - Grand Est; Weierstrass Institute. 2017. hal-01668975v1

HAL Id: hal-01668975

https://inria.hal.science/hal-01668975v1

Submitted on 20 Dec 2017 (v1), last revised 13 Feb 2018 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Data and methods for *A threshold model for local volatility: evidence of leverage and mean reversion effects on historical data*

Antoine Lejay , Paolo Pigato

RESEARCH REPORT
N° 0494
December 20, 2017
Project-Teams TOSCA

ISSN 0249-6399 ISRN INRIA/RR--0494--FR+ENG

Data and methods for *A threshold model for local volatility: evidence of leverage and mean reversion effects on historical data*

Antoine Lejay* , Paolo Pigato†

Project-Teams TOSCA

Research Report n° 0494 — December 20, 2017 — 21 pages

Abstract: This technical report presents the methodology and the numerical results for 21 stock prices under the assumption they follow a Drifted Geometric Oscillating Brownian motion model. Such a model takes leverage and mean-reversion effects into account. Ce rapport completes the article *A threshold model for local volatility: evidence of leverage and mean reversion effects on historical data.*

Key-words: Financial Mathematics; geometric Oscillating Brownian motion; Realized volatility estimator; Maximum likelihood; mean-reversion; leverage effect

* Université de Lorraine, IECL, UMR 7502, Vandœuvre-lès-Nancy, F-54600, France
CNRS, IECL, UMR 7502, Vandœuvre-lès-Nancy, F-54600, France
Inria, Villers-lès-Nancy, F-54600, France
E-mail: `Antoine.Lejay@univ-lorraine.fr`

† Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, Berlin, 10117, Germany
E-mail: `paolo.pigato@wias-berlin.de`

RESEARCH CENTRE
NANCY – GRAND EST

615 rue du Jardin Botanique
CS20101
54603 Villers-lès-Nancy Cedex

Données et méthodes pour *A threshold model for local volatility: evidence of leverage and mean reversion effects on historical data*

Résumé : Ce rapport technique présente la méthodologie et les résultats numériques pour les prix de 21 actifs boursiers sous l'hypothèse qu'ils se comportent comme un mouvement brownien oscillant avec dérive. Un tel modèle prend en compte les effets de levier et de retour à la moyenne. Ce rapport complète l'article *A threshold model for local volatility: evidence of leverage and mean reversion effects on historical data.*

Mots-clés : mathématiques financières ; Mouvement brownien géométrique oscillant ; estimateur de la volatilité réalisée ; maximum de vraisemblance ; retour à la moyenne ; effet de levier

Contents

1. Introduction **4**

2. Theoretical results **4**
2.1. Estimation of the model . . . 4
2.1.1. The model . . . 4
2.1.2. The parameters . . . 5
2.1.3. The estimation procedure . . . 5
2.1.4. The detrended time series . . . 6
2.1.5. Approximation of the log-likelihood . . . 6
2.1.6. Approximation of the skewness parameter . . . 6
2.1.7. Testing the equality of the two volatilities . . . 7
2.2. The data . . . 8
2.2.1. The stocks . . . 8
2.2.2. Software . . . 8

3. Numerical results **9**
3.1. The estimated values . . . 9
3.2. Normalized Occupation times and normalized number of crossings . . . 10
3.3. The detrended time series . . . 11
3.4. Prices of the stocks . . . 12
3.5. The log-likelihood . . . 13
3.6. Estimation of the volatilities in function of the threshold . . . 14
3.7. Estimation of the drift in function of the threshold . . . 15
3.8. Non parametric estimation of the volatility . . . 16
3.9. Non parametric estimation of the drift . . . 17
3.10. Estimation of the skewness parameter . . . 18
3.11. Testing the equality of the two volatilities . . . 19

A. The Girsanov weight **20**

1. Introduction

In [7], we present a simple model — the Geometric Oscillating Brownian motion (GOBM) — which extends the Geometric Brownian motion (also called the Black & Scholes model) by allowing the volatility and the appreciation rate to be piecewise constant. This model can be seen as a continuous time extension of a Self-Exciting Threshold AutoRegressive (SETAR) time series [13, 14]. It is also an alternative to the one proposed by L. Esquível and P. Mota [3] which is based on a regime switching with a thin layer which prevents immediate switching.

In [3], their model have been tested against the stock prices of 21 assets on the period 2005-2009. Their method is derived from the one used in time series [14].

In this technical report, we present the numerical results which we obtain on the same 21 stocks prices as in [3] based on the estimators of the volatility and appreciation rates which studied in [8] and [9].

These numerical results are summarized and commented in [7]. Therefore, we only give a technical presentation of the estimators and the raw results.

Outline. In Section 2, we present the different estimators we use. In Section 3, we give the numerical results.

Acknowledgement. P. Pigato gratefully acknowledges financial support from ERC via Grant CoG-683164.

2. Theoretical results

2.1. Estimation of the model

Notations: $x^+ = \max\{x, 0\}$, $x^- = \max\{-x, 0\}$.

2.1.1. The model

The price S is solution to the SDE

$$S_t = x + \int_0^t \sigma(S_s) S_s \, \mathrm{d}B_s + \int_0^t \mu(S_s) S_s \, \mathrm{d}s, \tag{1}$$

where, for threshold $m \in \mathbb{R}$,

$$\sigma(x) = \begin{cases} \sigma_+ & \text{if } x \geq m, \\ \sigma_- & \text{if } x < m \end{cases} \text{ and } \mu(x) = \begin{cases} \mu_+ & \text{if } x \geq m, \\ \mu_- & \text{if } x < m. \end{cases} \tag{2}$$

The *log-price* $X = \log(S)$ satisfies the SDE

$$X_t = x + \int_0^t \sigma(X_s) \, \mathrm{d}B_s + \int_0^t b(X_s) \, \mathrm{d}s, \tag{3}$$

with

$$r = \log(m) \text{ and } b(x) = \begin{cases} b_+ = \mu_+ - \sigma_+^2/2 & \text{if } x \geq r, \\ b_- = \mu_- - \sigma_-^2/2 & \text{if } x < r. \end{cases} \tag{4}$$

2.1.2. The parameters

There are 5 parameters to identify on the log-price:

- $\sigma_- > 0$, the volatility below the negative side,
- $\sigma_+ > 0$, the volatility above the threshold,
- $b_- \in \mathbb{R}$, the drift below the threshold,
- $b_+ \in \mathbb{R}$, the drift above the threshold,
- $r \in \mathbb{R}$, the threshold.

2.1.3. The estimation procedure

Given the daily observations $\{X_i\}_{i=0}^n$ of the log-price (the time step between two observations is $\Delta t = 1$), the estimation is performed in several steps:

(1) The range of the log-price X is split in values $\{r_i\}_{i=1,\dots,M}$ that are used as possible threshold. We use $M = 60$.

(2) For each r_i, we set $Y = X - r_i$ so that Y is an OBM with a threshold at 0. We estimate the occupation times Λ_+ and Λ_- by

$$\Lambda_+(i) = \Delta t \sum_{i=0}^{n-1} \mathbf{1}_{Y_i \geq 0} \text{ and } \Lambda_-(i) = \Delta t \sum_{i=0}^{n-1} \mathbf{1}_{Y_i \leq 0}. \tag{5}$$

We estimate the volatilities σ_+ and σ_- by

$$\sigma_+(i) = \left(\frac{\sum_{j=0}^{n-1} (Y_{j+1} - Y_j) \times (Y_{j+1}^+ - Y_j^+)}{\Lambda_+(i)} \right)^{1/2}$$
$$\text{and } \sigma_-(i) = \left(\frac{\sum_{j=0}^{n-1} (Y_{j+1} - Y_j) \times (Y_{j+1}^- - Y_j^-)}{\Lambda_-(i)} \right)^{1/2}.$$

We estimate the local time L of Y at 0 at time n by

$$L(i) = \frac{3\sqrt{\pi}}{2\sqrt{2}} \frac{(\sigma_+(i) + \sigma_-(i))}{\sigma_+(i)\sigma_-(i)} \sum_{j=0}^{n-1} (Y_{j+1}^- - Y_j^-) \times (Y_{j+1}^+ - Y_j^+). \tag{6}$$

We also estimate the drifts $b_+(i)$ and $b_-(i)$ by

$$b_+(i) = \frac{-\frac{1}{2}L(i) - Y_n^+ - Y_0^+}{\Lambda_+(i)} \text{ and } b_-(i) = \frac{\frac{1}{2}L(i) - Y_n^- - Y_0^-}{\Lambda_-(i)}.$$

(3) For each threshold r_i, we compute

$$\ell(i) = \widehat{\text{Log-Lik}}(X, b_+(i), b_-(i), \sigma_+(i), \sigma_-(i), r_i)$$

given by (10) below. We select the value $\widehat{\imath}$ of i for which $\ell(i)$ is maximal. This follows the principle of the Akaike Information Principle [1] (the number of parameters is the same whatever r_i).

(4) The estimated values of $(r, \sigma_+, \sigma_-, b_+, b_-)$ are then

$$(\widehat{r}, \widehat{\sigma}_+, \widehat{\sigma}_-, \widehat{b}_+, \widehat{b}_-) = (r_{\widehat{\imath}}, \sigma_+(\widehat{\imath}), \sigma_-(\widehat{\imath}), b_+(\widehat{\imath}), b_-(\widehat{\imath})).$$

2.1.4. The detrended time series

Our observations $\{\xi_i\}_{i=0}^n$ with time step $\Delta t = 1$ forms a time series. Given $\beta = (\beta_+, \beta_-) \in \mathbb{R}^2$, we define the *detrended time series* by

$$\xi_i^{\lhd\beta} = \xi_i - \beta_+ \mathbf{1}_{\xi_i \geq 0} \Delta t - \beta_- \mathbf{1}_{\xi_i \leq 0} \Delta t. \tag{7}$$

2.1.5. Approximation of the log-likelihood

With

$$\Theta(x) = \mathbf{1}_{x\geq 0}\sigma_+ + \mathbf{1}_{x<0}\sigma_- \text{ and } \Upsilon(x) = x/\Theta(x),$$

the solution to (3) is transformed into $Y_t = \Upsilon(X_t - r)$ the solution to

$$Y_t = Y_0 + B_t + \int_0^t \frac{b(Y_s)}{\sigma(Y_s)}\,\mathrm{d}s + \theta L_t^0(Y) \text{ with } \theta = \frac{\sigma_- - \sigma_+}{\sigma_- + \sigma_+}, \tag{8}$$

where $L_t^0(Y)$ is the symmetric local at 0 of Y. The SDE (8) is the one of a drifted Skew Brownian motion. When $b_+ = b_- = 0$, then Y is a Skew Brownian motion of parameter $\theta \in (-1, 1)$. Its density is given by

$$p_{\mathrm{SBM}}(t, x, y) = \frac{1}{\sqrt{2\pi t}}\left(\exp\left(\frac{-(x-y)^2}{2t}\right) + \mathrm{sgn}(y)\theta \exp\left(\frac{-(|x|+|y|)^2}{2t}\right)\right). \tag{9}$$

Eq. (9) gives a simple formula for the transition density of X without drift (see also [5]):

$$p(t, x, y; \sigma_+, \sigma_-) = p_{\mathrm{SBM}}\left(t, \frac{x}{\Theta(x)}, \frac{y}{\Theta(y)}\right)\frac{1}{\Theta(y)}.$$

For our data, the ratio b/σ is small (see 3.3). Since the time step between the observations is $\Delta t = 1$, this means that the volatility is the main factor explaining variations of the log-prices. For an estimated drift $\beta = (\beta_+, \beta_-)$, an estimated volatility (s_+, s_-) and a choice ρ of the threshold, the log-likelihood is approximated[1] by

$$\widehat{\text{Log-Lik}}(X, \beta_+, \beta_-, s_+, s_-, \rho) = \sum_{i=1}^n \log p(\Delta t, (X-\rho)_{i-1}^{\lhd(\beta_+,\beta_-)}, (X-\rho)_i^{\lhd(\beta_+,\beta_-)}; s_+, s_-), \tag{10}$$

where $(X-\rho)^{\lhd(\beta_+,\beta_-)}$ is the detrended time series defined by (7) for the observations $\{X_i\}_{i=0}^n$.

2.1.6. Approximation of the skewness parameter

As asserted above, when $b = 0$, then $Y = \Upsilon(X)$ is a Skew Brownian motion (SBM) of parameter $\theta = (\sigma_- - \sigma_+)/(\sigma_- + \sigma_+)$.

The observations $\{X_i\}_{i=0}^n$ are transformed into (with $\Delta t = 1$),

$$Y_i = \frac{X_i - r - \widehat{b}_+ \mathbf{1}_{X_i \geq 0}\Delta t - \widehat{b}_- \mathbf{1}_{X_i<0}\Delta t}{\widehat{\sigma}_+ \mathbf{1}_{X_i\geq 0} + \widehat{\sigma}_- \mathbf{1}_{X_i<0}}.$$

Hence, if X follows an OBM, then the dynamic of the Y should be close to the ones of a SBM of parameter θ.

[1] The approximation comes from the fact that the detrended time series is not exactly the observations of a Skew Brownian motion. Besides, only an approximation of the drift is used.

Given the discrete observations of a SBM of parameter θ, and owing the explicit expression (9) of its density transition function, the parameter θ is easily estimated by θ_{MLE} through a Maximum Likelihood Estimation (MLE). It is shown in [6] that θ_{MLE} is a consistent estimator of θ. When $\theta = 0$ (the Brownian motion), then the rate of convergence is $n^{1/4}$. It is conjectured that this rate of convergence still holds for $\theta \neq 0$.

In Sect. 3.10, we estimate θ_{MLE} which we compare to the predicted value $\widehat{\theta} = (\widehat{\sigma}_- - \widehat{\sigma}_+)/(\widehat{\sigma}_- + \widehat{\sigma}_+)$. We have a good agreement for several of the stocks. A bad agreement could be explain by the following facts:

- The quality of the estimation depends on the local time (which we relate to the number of crossings) which is sometimes very low, leading to a bad estimation.
- The stock log-price does not follows an OBM. An extra local time could be added for example to take a *directional predictability* into account (See [2] for example). We note however that θ_{MLE} is sometimes negative, which seems to contradict a directional predictability in which the (log-)prices tend to rise.

2.1.7. Testing the equality of the two volatilities

In this section, it is assumed that $b_- = b_+ = 0$, that is, the drift terms vanish.

From the data, we consider the hypotheses

(H_0) (null hypothesis) $\sigma_- = \sigma_+$;
(H_1) (alternative hypothesis) $\sigma_- \neq \sigma_+$.

For the sake of simplicity, let us set

$$S_\pm = \sigma_\pm^2,\ S_\pm^n = (\hat{\sigma}_\pm^n)^2,\ \mathbf{S} = \begin{bmatrix} S_- \\ S_+ \end{bmatrix} \text{ and } \mathbf{S}^n = \begin{bmatrix} S_-^n \\ S_+^n \end{bmatrix}.$$

We also define

$$O_\pm = \frac{1}{T}\int_0^T \mathbf{1}_{\pm X_s \geq 0}\,\mathrm{d}s \text{ and } O_\pm^n = \frac{1}{n}\sum_{j=0}^{n-1} \mathbf{1}_{\pm X_j \geq 0}.$$

the normalized positive (resp. negative) occupation time.

The following convergence results was established

$$\sqrt{n}(\mathbf{S}^n - \mathbf{S}) \xrightarrow[n\to\infty]{\text{stably}} \sqrt{2}\begin{bmatrix} S_- \frac{G_-}{\sqrt{O_-}} \\ S_+ \frac{G_+}{\sqrt{O_+}} \end{bmatrix} = M\mathbf{G} \text{ with } M = \sqrt{2}\begin{bmatrix} S_-/\sqrt{O_-} & 0 \\ 0 & S_+/\sqrt{O_+} \end{bmatrix}, \tag{11}$$

where $\mathbf{G} = \begin{bmatrix} G_- & G_+ \end{bmatrix}' \sim \mathcal{N}(0, \mathrm{Id})$ is a Gaussian vector independent from the process X.

The matrix M is approximated by

$$M^n = \sqrt{2}\begin{bmatrix} S_-^n/\sqrt{O_-^n} & 0 \\ 0 & S_+^n/\sqrt{O_+^n} \end{bmatrix}.$$

Thanks to the isotropy of the Gaussian vector $\mathbf{G}$, we define for a level of confidence α the quantity q_α by $\mathbb{P}(|\mathbf{G}| \leq q_\alpha) = 1 - \alpha$. The random variable $|\mathbf{G}|^2$ follows a χ^2 distribution with two degrees of freedom so that q_α is easily computed by the `R` command `qchisq`.

Our *acceptance region* of level α is then the ellipsis

$$\mathcal{R}_\alpha = \left\{ \mathbf{S}^n + \frac{q_\alpha}{\sqrt{n}} M^n \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} \,\middle|\, \theta \in [0, 2\pi) \right\}.$$

Our **rule of decision** is to reject the Null Hypothesis (H_0) if the diagonal line $s : [0, +\infty) \mapsto (s, s)$ does not cross $\mathcal{R}_\alpha$.

2.2. The data

2.2.1. The stocks

The data are the same as the one in [11]. The daily quotations of the stocks are downloaded from the Web site Yahoo Finance [15]. The dates range from 01/02/2005 to 30/11/2009. Clearly, the time series for KO differs from the one in [11].

GOOG	Google	HP	Hewlett-Packard	AAPL	Apple
ADBE	Adobe	CA	CA	C	CitiGroup
KO	Coca-cola	CSCO	Cisco	IBM	IBM
JPM	JP Morgan	MCD	McDonalds	SBUX	Starbucks
PM	Philip Morris	PG	P & G	PFE	Pfizer
PCG	PG&E	NYT	New-York Times	MSFT	Microsoft
MSI	Motorola	MON	Monsanto	AMZN	Amazon

The number of samples is 1217 for each stock, excepted for PM where there are only 432 samples.

2.2.2. Software

All the estimations have been performed using the `R` software [12].

3. Numerical results

3.1. The estimated values

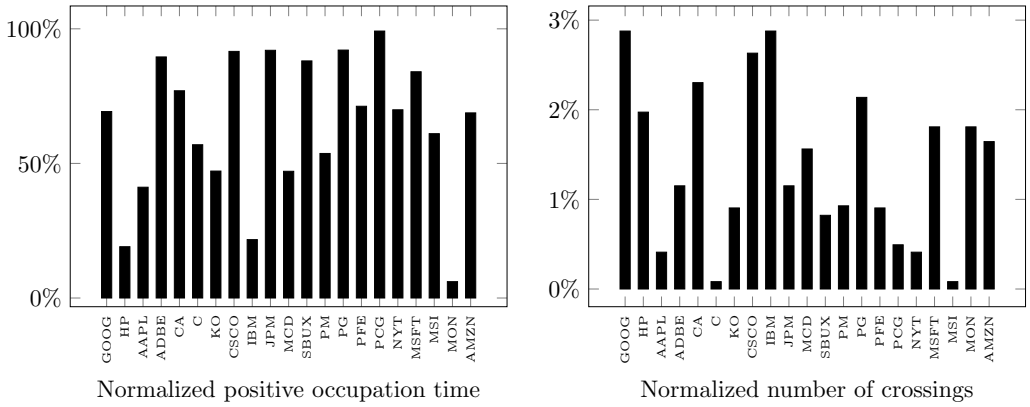
Estimated parameters for the GOBM model: $\sigma_\pm$ is the daily volatility $\mu_\pm = b_\pm + \sigma_\pm^2/2$ is the daily mean return rate, $b_\pm$ is the drift of the log-price and *signs* are the respective signs of b_-, b_+ (a $+-$ indicates a mean-reversion effect).

Oscillating Brownian motion (OBM)								
Index	m [\$]	σ_- [%]	σ_+ [%]	μ_- [‰]	μ_+ [‰]	b_- [‰]	b_+ [‰]	signs
GOOG	378.1	2.81	2.07	3.41	0.19	3.02	−0.02	+−
HP	57.9	4.18	2.53	1.78	−3.58	0.91	−3.90	+−
AAPL	117.0	3.78	2.56	2.19	0.11	1.48	−0.22	+−
ADBE	25.9	4.37	3.00	5.03	−0.48	4.07	−0.93	+−
CA	21.6	3.20	1.61	2.00	−0.56	1.49	−0.69	+−
C	40.4	7.47	1.09	−1.24	−0.48	−4.03	−0.54	−−
KO	47.6	1.49	1.13	0.54	0.14	0.43	0.08	++
CSCO	17.1	3.65	1.92	10.01	−0.44	9.35	−0.63	+−
IBM	115.4	1.64	1.27	0.71	−0.87	0.57	−0.95	+−
JPM	32.2	8.33	2.63	12.66	−0.34	9.19	−0.68	+−
MCD	51.6	1.28	1.77	1.33	−0.06	1.25	−0.22	+−
SBUX	13.3	4.52	2.92	2.00	−0.55	0.98	−0.98	+−
PM	45.3	2.66	1.76	0.76	−0.31	0.41	−0.47	+−
PG	52.2	1.81	1.27	2.48	0.03	2.31	−0.05	+−
PFE	18.9	2.51	1.30	0.63	−0.36	0.32	−0.44	+−
PCG	33.9	7.09	1.45	24.20	0.08	21.69	−0.02	+−
NYT	15.6	4.98	1.64	0.08	−1.16	−1.16	−1.29	−−
MSFT	23.0	3.28	1.64	6.17	−0.83	5.64	−0.96	+−
MSI	14.3	4.18	1.64	−0.35	−0.02	−1.22	−0.15	−−
MON	119.2	3.41	2.73	1.32	−5.72	0.74	−6.09	+−
MZN	39.4	2.42	3.44	3.28	0.63	2.98	0.04	++

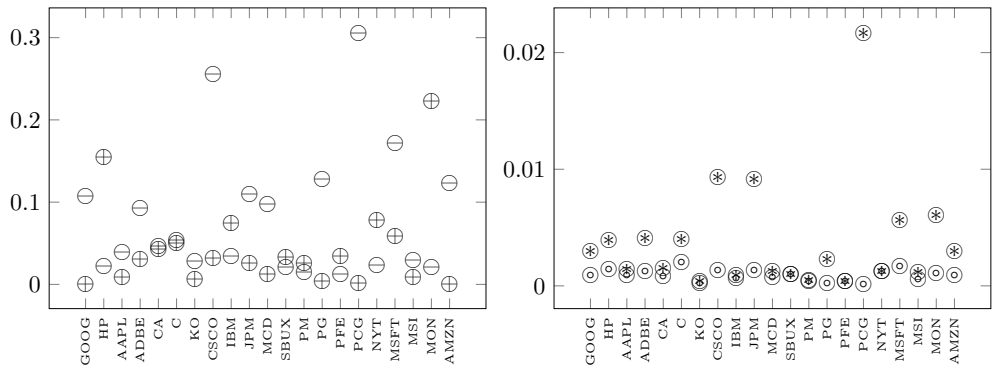
3.2. Normalized Occupation times and normalized number of crossings

When there is no drift, the positive and negative occupation times are known to follow a variant of the Arc-Sine distribution [5, 10]. With a drift, this distribution changes. However, it is expected that the normalized positive occupation time $\Lambda(T)/T$ up to time T is "close" to 0 or 1.

The normalized number of crossings is related to the local time. Actually, for the Skew Brownian motion, a consistent local time estimator may be built from this quantity [6].



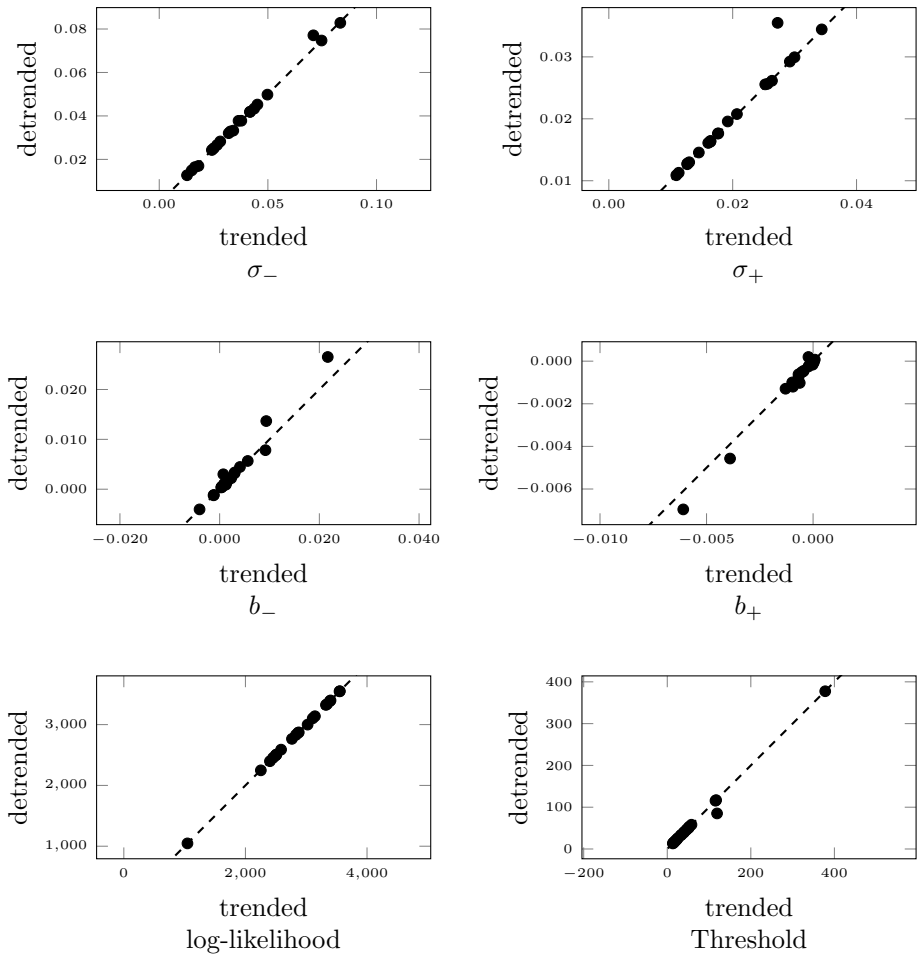
3.3. The detrended time series



(a) $\oplus = \left|\widehat{b}_+/\widehat{\sigma}_+\right|$, $\ominus = \left|\widehat{b}_-/\widehat{\sigma}_-\right|$

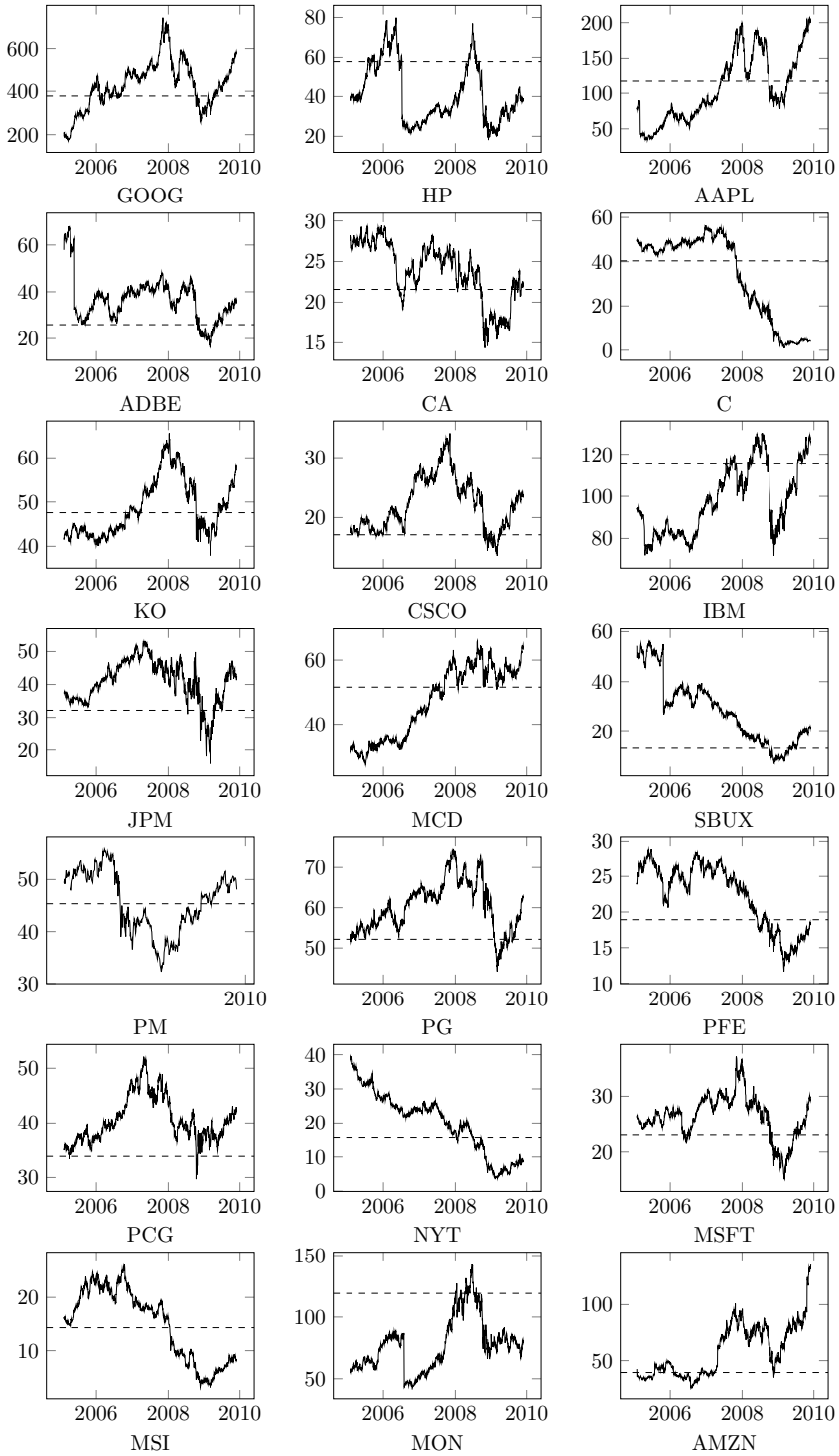
(b) $\circledast = \text{mean}(|X_i - X_i^\beta|)$, $\odot = \max(|X_i - X_i^\beta|)$ with $\beta = (\widehat{b}_+, \widehat{b}_-)$

The different parameters are estimated for the trended and the detrended time series. The estimations are similar. The stock for which there is a difference in (σ_-, σ_+) in MON, for which the log-likelihood is flat on a rather large range (See 3.5).



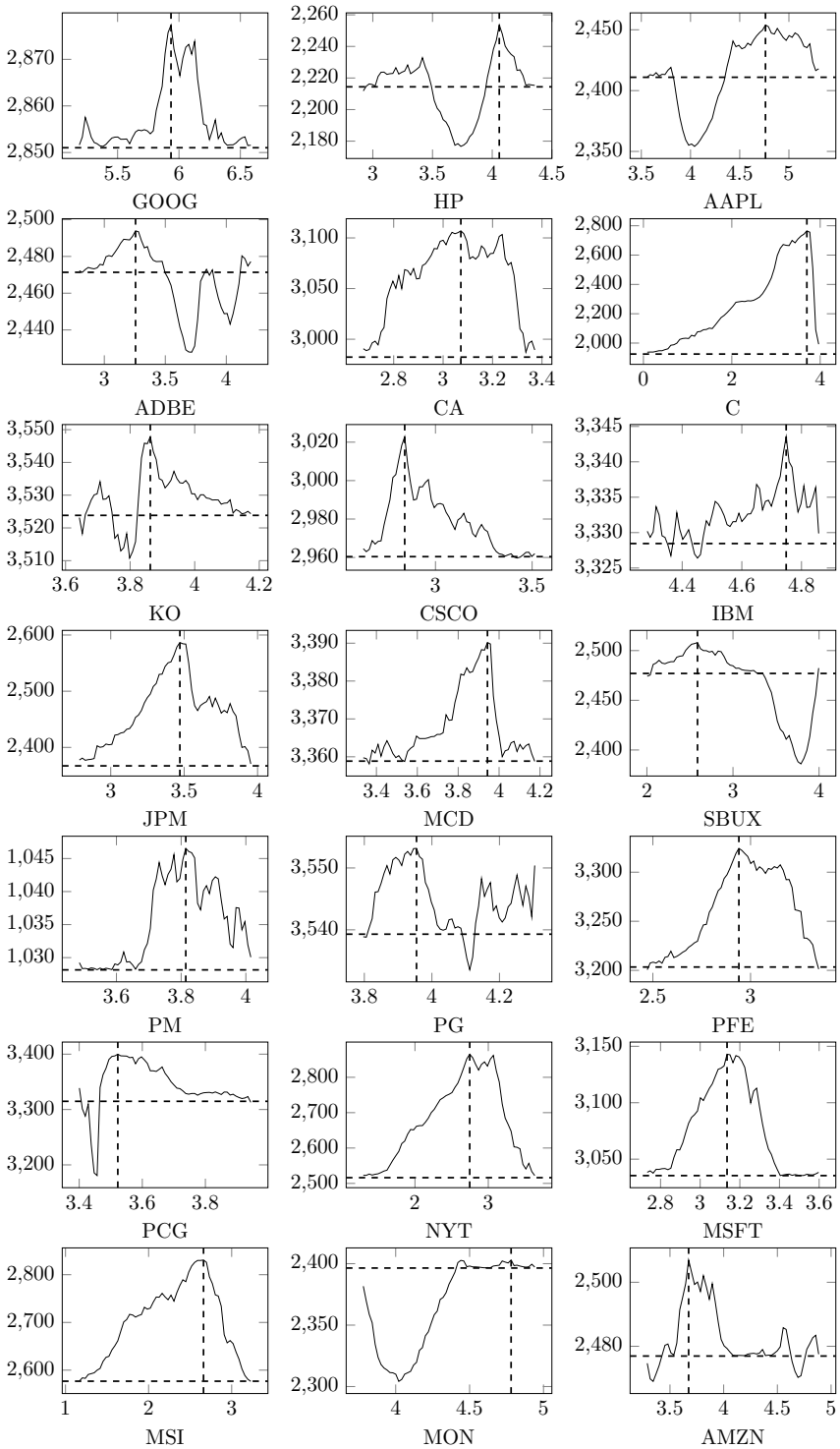
3.4. Prices of the stocks

The prices are expressed in USD. The horizontal dashed line is the threshold m.



3.5. The log-likelihood

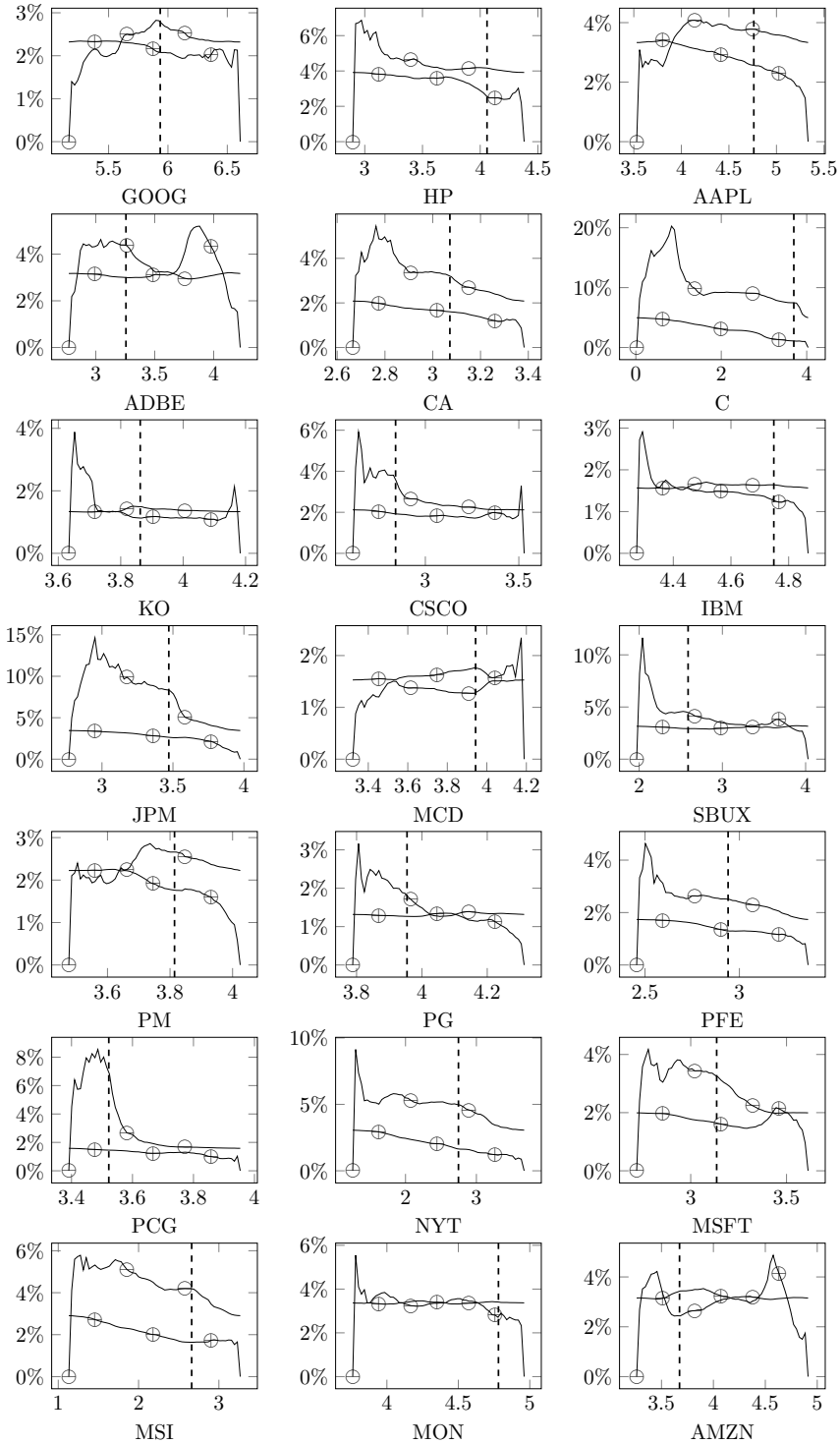
The dashed vertical lines are at $r(\widehat{\imath})$, the threshold which maximizes the log-likelihood. The dashed horizontal lines are at the log-likelihood of the Geometric Brownian motion.



3.6. Estimation of the volatilities in function of the threshold

The estimated volatilities $\sigma_+(i)$ (marked by $\oplus$) and $\sigma_-(i)$ (marked by $\ominus$) are plotted in function of the possible thresholds $r(i)$ for the log-prices.

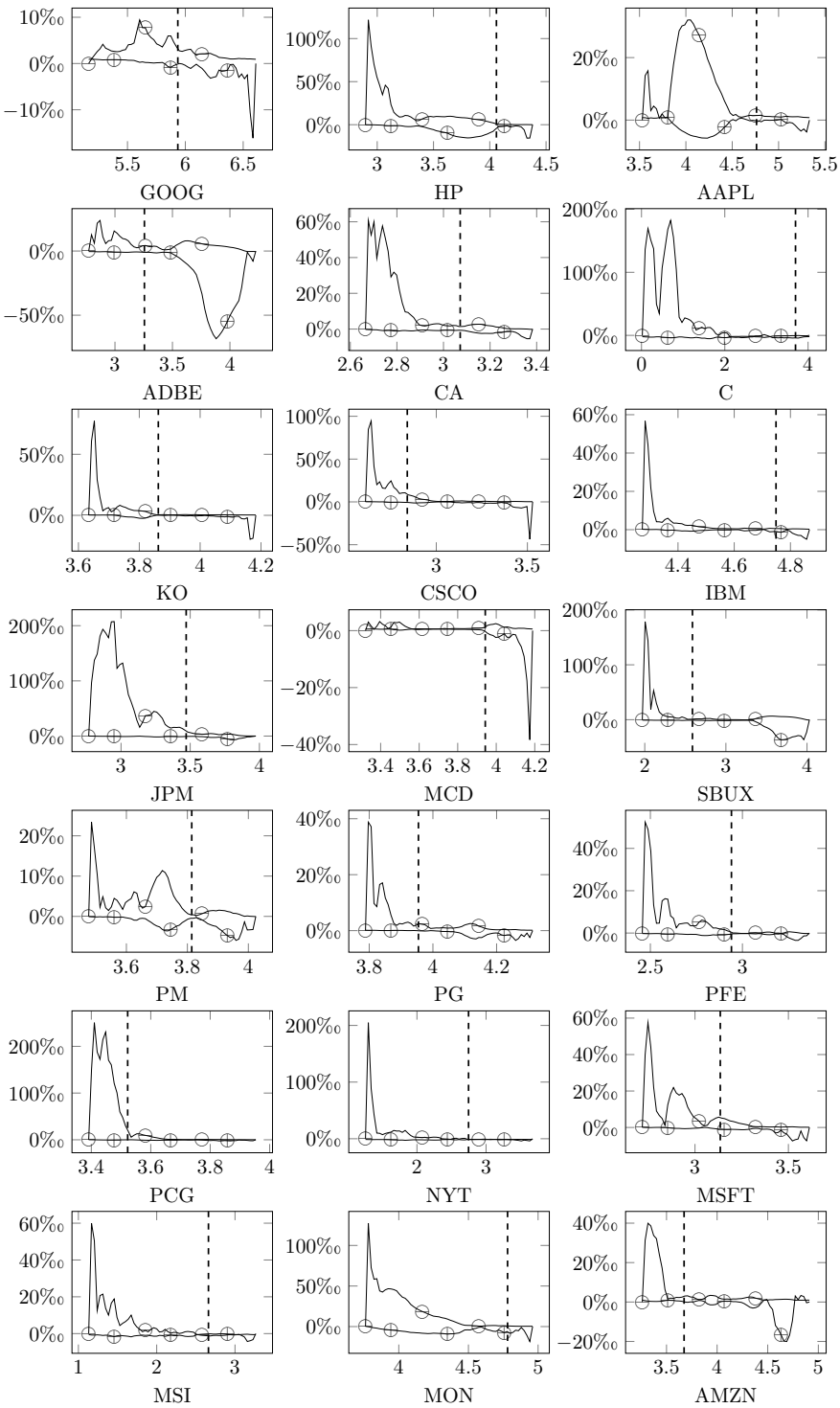
The dashed vertical lines are at $r(\widehat{i})$, the threshold which maximizes the log-likelihood.



3.7. Estimation of the drift in function of the threshold

The estimated drift $b_+(i)$ (marked by $\oplus$) and $b_-(i)$ (marked by $\ominus$) are plotted in function of the possible thresholds $r(i)$ for the log-prices.

The dashed vertical lines are at $r(\widehat{i})$, the threshold which maximizes the log-likelihood.

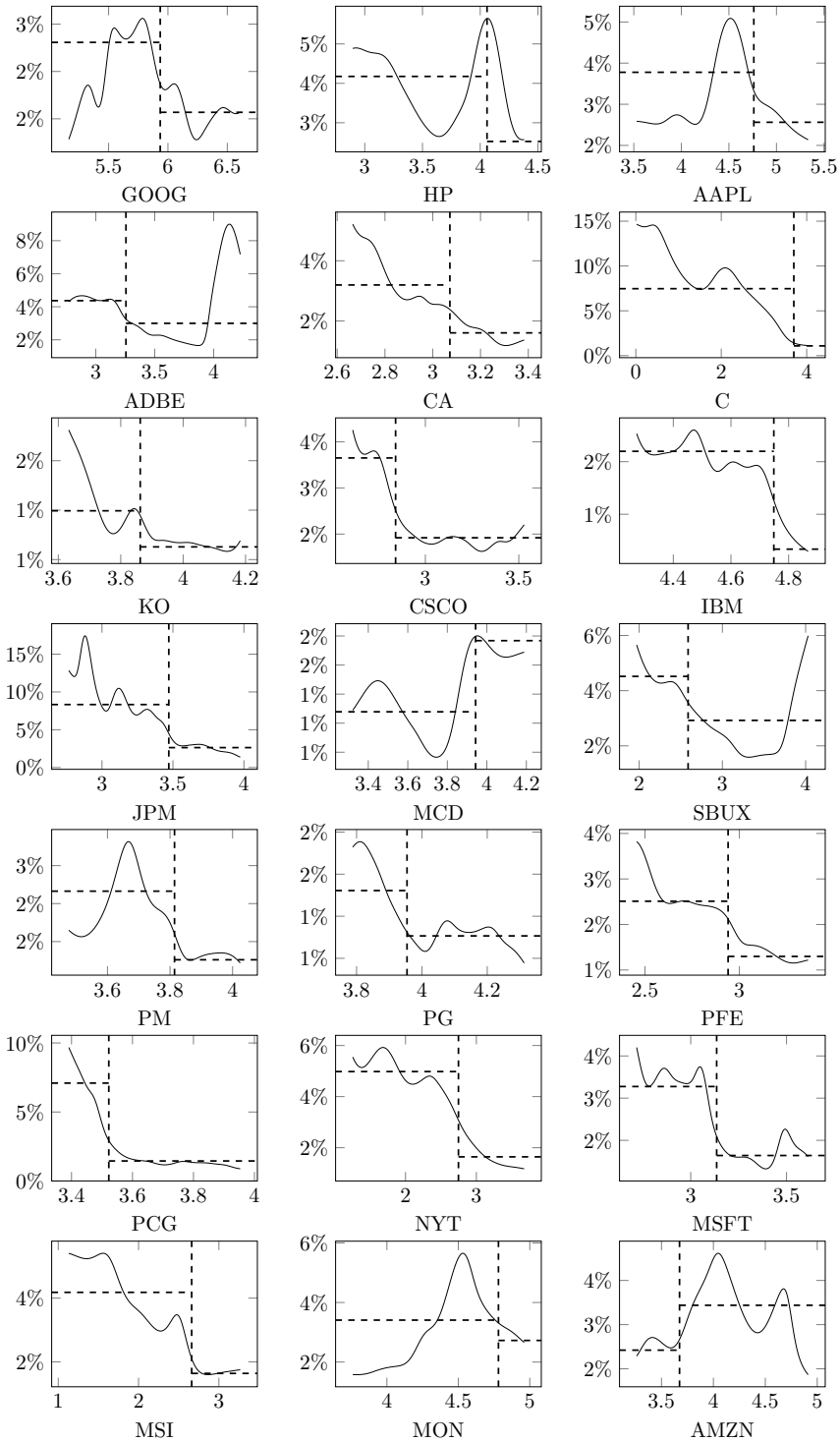


3.8. Non parametric estimation of the volatility

The dashed horizontal lines are at σ_- and σ_+, the estimated volatilities.

The dashed vertical lines are at $r(\widehat{\imath})$, the threshold which maximizes the log-likelihood.

The non parametric estimation is performed using `ksdiff` of the `R` package `sde` [4], which implement de Nadaraya-Watson estimator.

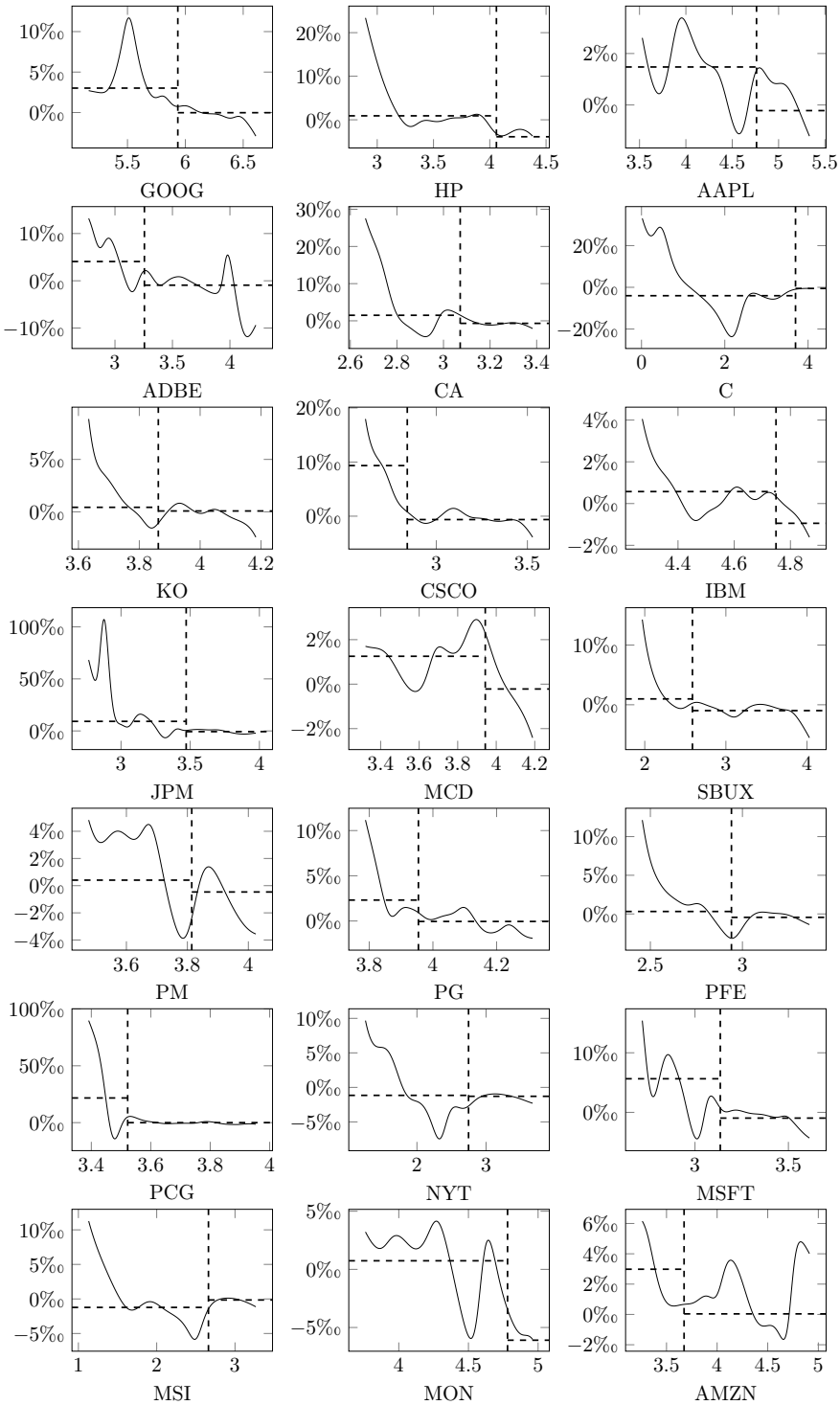


3.9. Non parametric estimation of the drift

The dashed horizontal lines are at b_- and b_+, the estimated volatilities.

The dashed vertical lines are at $r(\widehat{\imath})$, the threshold which maximizes the log-likelihood.

The non parametric estimation is performed using `ksdrift` of the `R` package `sde` [4], which implement de Nadaraya-Watson estimator.

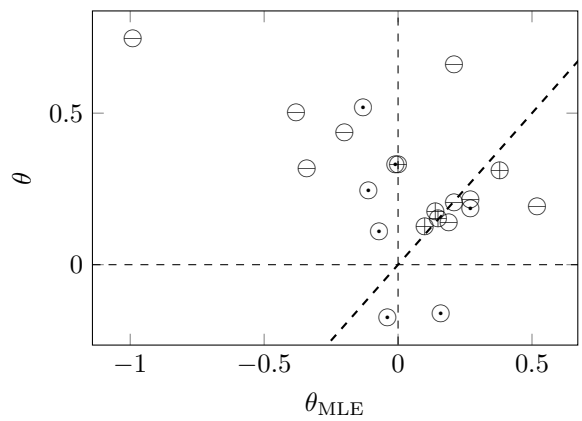


3.10. Estimation of the skewness parameter

The maximum likelihood estimator θ_{MLE} for the parameter $\theta \in (-1, 1)$ of the SBM is then estimated and compared to $\theta = (\widehat{\sigma}_- - \widehat{\sigma}_+)/(\widehat{\sigma}_- + \widehat{\sigma}_+)$.

The data are classified according to the range of the normalized number n_{c} of crossings $\ominus$: $[0, 1\%]$ $\odot$: $[1\%, 2\%]$ $\oplus$: $[2\%, 3\%]$. This classification gives an indication on how big is the local time [6], hence on the quality of the estimator.

Index	θ_{MLE}	θ	n_{c} [%]	Class
GOOG	0.15	0.15	2.88	$\oplus$
HP	−0.11	0.25	1.97	$\odot$
AAPL	0.52	0.19	0.41	$\ominus$
ADBE	0.27	0.19	1.15	$\odot$
CA	0.00	0.33	2.30	$\oplus$
C	−0.99	0.75	0.08	$\ominus$
KO	0.19	0.14	0.90	$\ominus$
CSCO	0.38	0.31	2.63	$\oplus$
IBM	0.10	0.13	2.88	$\oplus$
JPM	−0.13	0.52	1.15	$\odot$
MCD	0.16	−0.16	1.56	$\odot$
SBUX	0.27	0.21	0.82	$\ominus$
PM	0.21	0.20	0.93	$\ominus$
PG	0.14	0.18	2.14	$\oplus$
PFE	−0.34	0.32	0.90	$\ominus$
PCG	0.21	0.66	0.49	$\ominus$
NYT	−0.38	0.50	0.41	$\ominus$
MSFT	−0.01	0.33	1.81	$\odot$
MSI	−0.20	0.44	0.08	$\ominus$
MON	−0.07	0.11	1.81	$\odot$
AMZN	−0.04	−0.17	1.64	$\odot$

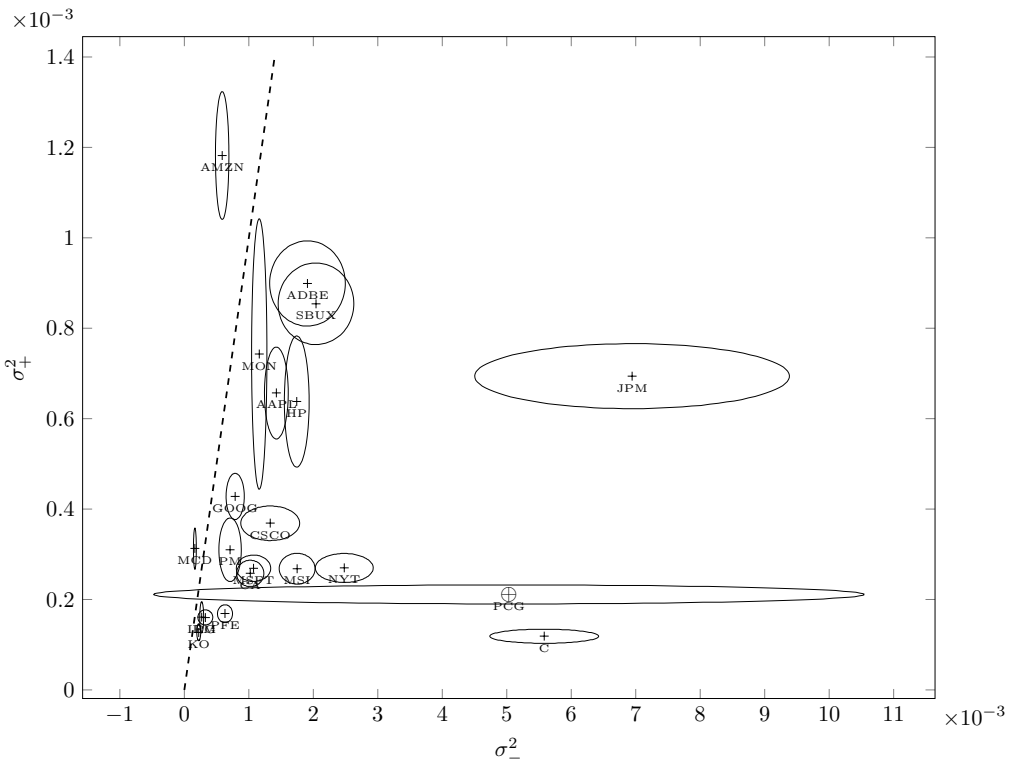


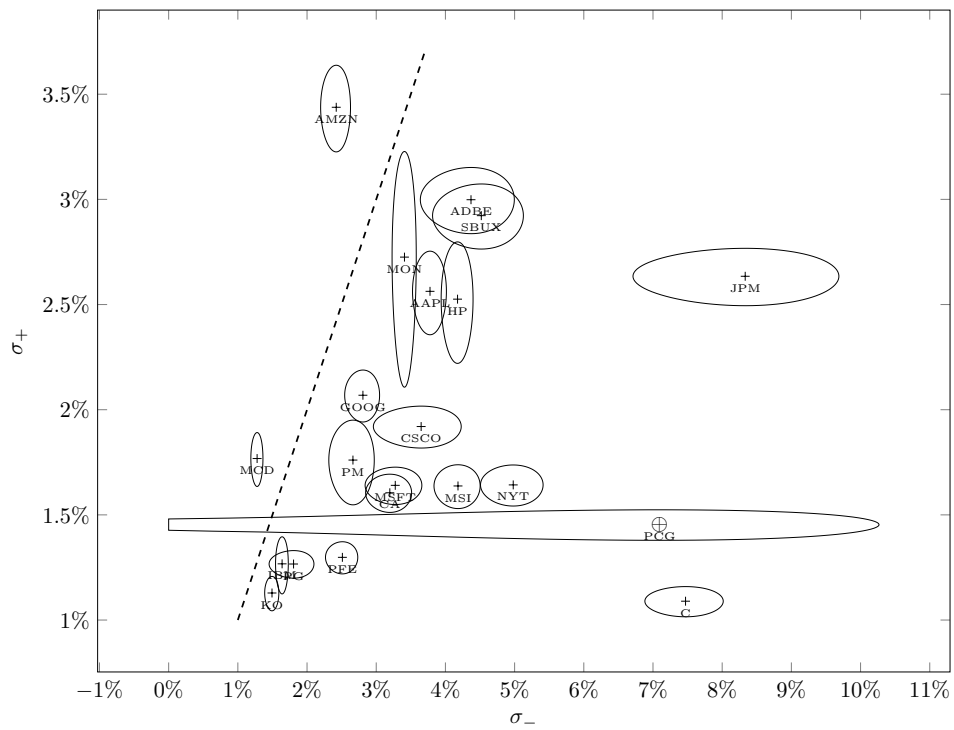
3.11. Testing the equality of the two volatilities

The 95 %-confidence regions for testing the null hypothesis $\sigma_- = \sigma_+$ are plotted for each of the 21 stocks. They are plotted with respect to (σ_-^2, σ_+^2), hence the scale, and to (σ_-, σ_+).

The dashed line is the diagonal line $x \mapsto x$.

Only one confidence ellipsis is crossed by this line, the one for PCG, but this is due to a very small negative occupation time.





A. The Girsanov weight

In this appendix, we show that the Girsanov weight may be estimated as well, thanks to the particular structure of our drift and volatility.

The drift of the centered log-price $Y = X - r$ could be removed using a Girsanov transform. More precisely, for a function Ξ on the space $\mathcal{C}([0,T];\mathbb{R})$ of continuous functions, $\mathbb{E}[G\Xi(Y)] = \mathbb{E}[\Xi(\check{Y})]$, where $\check{Y}$ is the solution to

$$\check{Y}_t = x + \int_0^t \sigma(\check{Y}_s)\,\mathrm{d}B'_s$$
$$\text{and } G = \exp\left(-\int_0^T \Gamma(Y_s)\,\mathrm{d}B_s - \frac{1}{2}\int_0^T \Gamma^2(Y_s)\,\mathrm{d}s\right) \text{ with } \Gamma(x) = \frac{b(x)}{\sigma(x)},$$

for a Brownian motion B'.

With the Itô-Tanaka formula,

$$Y_t^+ = Y_0^+ + \sigma_+ \int_0^t \mathbf{1}_{Y_s \geq 0}\,\mathrm{d}B_s + b_+\Lambda(t) + \frac{1}{2}L_t^0(Y), \tag{12}$$
$$Y_t^- = Y_0^- - \sigma_- \int_0^t \mathbf{1}_{Y_s \leq 0}\,\mathrm{d}B_s - b_-(t - \Lambda(t)) + \frac{1}{2}L_t^0(Y) \tag{13}$$

where $\Lambda(t) = \int_0^t \mathbf{1}_{Y_s \geq 0}\,\mathrm{d}s$ and $\{L_t^0(Y)\}_{t\geq 0}$ is the symmetric local time of Y at 0. Multiplying (12) by b_+/σ_+^2 and (13) by b_-/σ_-^2 leads to

$$\frac{b_+}{\sigma_+^2}Y_t^+ = \frac{b_+}{\sigma_+^2}Y_0^+ + \int_0^t \mathbf{1}_{Y_s \geq 0}\Gamma(Y_s)\,\mathrm{d}B_s + \frac{b_+^2}{\sigma_+^2}\Lambda(t) + \frac{b_+}{2\sigma_+^2}L_t^0(Y),$$

and

$$\frac{b_-}{\sigma_-^2}Y_t^- = \frac{b_-}{\sigma_-^2}Y_0^- - \int_0^t \mathbf{1}_{Y_s \geq 0}\Gamma(Y_s)\,\mathrm{d}B_s - \frac{b_-^2}{\sigma_-^2}(t - \Lambda(t)) + \frac{b_-}{2\sigma_-^2}L_t^0(Y).$$

Hence,

$$\int_0^t \Gamma(Y_s)\,\mathrm{d}B_s = \frac{b_+}{\sigma_+^2}Y_t^+ - \frac{b_+}{\sigma_+^2}Y_0^+ - \frac{b_-}{\sigma_-^2}Y_t^- + \frac{b_-}{\sigma_-^2}Y_0^- - \frac{b_+^2}{\sigma_+^2}\Lambda(t) - \frac{b_-^2}{\sigma_-^2}(t - \Lambda(t)) - \left(\frac{b_+}{2\sigma_+^2} + \frac{b_-}{2\sigma_-^2}\right)L_t^0(Y).$$

On the other hand,

$$\frac{1}{2}\int_0^t \Gamma(Y_s)^2\,\mathrm{d}s = \frac{b_+^2}{2\sigma_+^2}\Lambda(t) + \frac{b_-^2}{2\sigma_-^2}(t - \Lambda(t)).$$

It is then possible to estimate the Girsanov weight from the estimations of $\sigma_\pm$, $b_\pm$ and $\Lambda(t)$ and $L_t^0(Y)$. The quantities $\Lambda(t)$ and $L_t^0(Y)$ can be approximated as in (5) and (6).

References

[1] H. Akaike. "Information theory and an extension of the maximum likelihood principle". In: *Second International Symposium on Information Theory (Tsahkadsor, 1971)*. Akadémiai Kiadó, Budapest, 1973, pp. 267–281.

[2] L. H. R. Alvarez and P. Salminen. "Timing in the Presence of Directional Predictability: Optimal Stopping of Skew Brownian Motion". In: (2016). Preprint arxiv:1608.04537.

[3] M. L. Esquível and P. P. Mota. "On some auto-induced regime switching double-threshold glued diffusions". In: *J. Stat. Theory Pract.* 8.4 (2014), pp. 760–771. DOI: 10.1080/15598608.2013.854184.

[4] S. M. Iacus. *Simulation and inference for stochastic differential equations, with R examples.* Springer Series in Statistics. Springer, New York, 2008. DOI: 10.1007/978-0-387-75839-8.

[5] J. Keilson and J. A. Wellner. "Oscillating Brownian motion". In: *J. Appl. Probability* 15.2 (1978), pp. 300–310.

[6] A. Lejay, E. Mordecki, and S. Torres. *Two consistent estimators for the Skew Brownian motion.* Preprint. 2017.

[7] A. Lejay and P. Pigato. "A threshold model for local volatility: evidence of leverage and mean reversion effects on historical data". In: (2017). Preprint.

[8] A. Lejay and P. Pigato. "Statistical estimation of the Oscillating Brownian Motion". In: *Bernoulli* (2017). To appear.

[9] A. Lejay and P. Pigato. *Estimation of the drift parameters of a Drifted Oscillating Brownian motion.* In preparation. 2017.

[10] A. Lejay and P. Pigato. "Statistical estimation of the Oscillating Brownian Motion". In: *Bernoulli* (2017).

[11] P. P. Mota and M. L. Esquível. "On a continuous time stock price model with regime switching, delay, and threshold". In: *Quant. Finance* 14.8 (2014), pp. 1479–1488. ISSN: 1469-7688. DOI: 10.1080/14697688.2013.879990.

[12] R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing. Vienna, Austria, 2016. URL: https://www.R-project.org/.

[13] H. Tong. "Threshold models in time series analysis — 30 years on". In: *Statistics and its Interface* 4 (2011).

[14] H. Tong. *Threshold models in nonlinear time series analysis.* Vol. 21. Lecture Notes in Statistics. Springer-Verlag, New York, 1983. DOI: 10.1007/978-1-4684-7888-4.

[15] Yahoo, ed. *Yahoo Finance.* Data downloaded on 2017-02-03. June 2017. URL: https://finance.yahoo.com.

RESEARCH CENTRE
NANCY – GRAND EST

615 rue du Jardin Botanique

CS20101

54603 Villers-lès-Nancy Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399