

This article was downloaded by: [KU Leuven University Library]

On: 19 March 2015, At: 02:10

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Quantitative Finance

Publication details, including instructions for authors and subscription information:  
<http://www.tandfonline.com/loi/rqf20>

### Stochastic modelling of herd behaviour indices

Florence Guillaume<sup>a</sup> & Daniël Linders<sup>b</sup>

<sup>a</sup> Department of Mathematics and Computer Sciences, University of Antwerp, Middelheimlaan 1, B-2020 Antwerpen, Belgium.

<sup>b</sup> Department of Accountancy, Finance and Insurance, KU Leuven, Naamsestraat 69, B-3000 Leuven, Belgium.

Published online: 12 Mar 2015.



[Click for updates](#)

To cite this article: Florence Guillaume & Daniël Linders (2015): Stochastic modelling of herd behaviour indices, Quantitative Finance, DOI: [10.1080/14697688.2015.1007075](https://doi.org/10.1080/14697688.2015.1007075)

To link to this article: <http://dx.doi.org/10.1080/14697688.2015.1007075>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

# Stochastic modelling of herd behaviour indices

FLORENCE GUILLAUME\*<sup>†</sup> and DANIËL LINDERS<sup>‡</sup>

<sup>†</sup>Department of Mathematics and Computer Sciences, University of Antwerp, Middelheimlaan 1, B-2020 Antwerpen, Belgium

<sup>‡</sup>Department of Accountancy, Finance and Insurance, KU Leuven, Naamsestraat 69, B-3000 Leuven, Belgium

(Received 19 March 2014; accepted 8 January 2015)

This paper proposes different diffusion processes to model herd behaviour indices such as the Herd Behaviour Index (HIX). These models arise by combining popular mean-reverting processes with simple algebraic functions mapping the definition domain of the underlying mean-reverting process to the unit interval. The so obtained Itô processes preserve, to some extent, the mean-reverting trend of the underlying process while satisfying the fundamental properties of the so-called herd behaviour indices. In a numerical study, we calibrate the different model settings to time series data for a period spanning from January 2000 until October 2009 and investigate their ability to predict the future behaviour of herd behaviour indices.

*Keywords:* Herd behaviour modelling; Time-dependent diffusion processes; Mean-reverting processes; Comonotonicity

## 1. Introduction

In order to obtain an equity portfolio with the desired risk–return profile, an investor invests in different stocks, because it is well documented that one of the benefits of owning a stock portfolio is diversification. However, it is well known that this diversification benefit is changing over time. Furthermore, if there is a strong co-movement between the different stocks, the diversification benefit is reduced because the stock prices will tend to move almost in unison and stock picking does not make sense; see e.g. [Linders and Schoutens \(2014\)](#). Therefore, it is important to have an idea of the future level of diversification when composing an equity portfolio. In [Dhaene, Linders, et al. \(2012\)](#), the authors introduce the Herd Behaviour Index (HIX), which is a barometer for the level of diversification between stock prices for the coming 30 days. The HIX is a number between 0 and 1, where 1 indicates that there is no diversification possible.

The HIX is changing overtime and should be modelled as a stochastic process. If today's HIX value is low, this indicates that one can compose a well-diversified equity portfolio. However, using only today's value for the HIX when taking investment decisions may not be sufficient, because one does not have an idea of the future direction of the HIX. As a result, the diversification one initially hopes for may be less than anticipated. Using a stochastic process which is calibrated on

the observed HIX values, one can find the trend and use this information to adjust the investment strategy.

Knowledge about the future trend of the HIX plays a crucial role when considering dispersion trading strategies; see e.g. [Driessen et al. \(2009\)](#) and [Laurence \(2008\)](#). A dispersion strategy is designed to buy co-movement risk. To be more precise, consider an investor who observes that currently there is only a low level of diversification between stock prices. However, he may believe that future diversification will increase. Then, he may consider this information as a sign that the market is oversold and index options are relatively expensive compared to vanilla options. Therefore, he can sell index options and buy an appropriate linear combination of vanilla options. This strategy is not an arbitrage strategy and will only be profitable when the co-movement between the stock prices will turn out to be smaller than anticipated. Therefore, the main difficulty of such dispersion strategies is to determine the right time to enter the trade. In [Laurence \(2008\)](#), the HIX is used to find the most appropriate moment to set up a dispersion trade. Indeed, one enters the dispersion strategy if the value of today's HIX is sufficiently large, compared to historical values.

Having a stochastic model to describe the HIX dynamics may improve the timing of a dispersion strategy. Indeed, such a stochastic model provides a methodology to estimate the future direction of the HIX. Moreover, it allows us to determine the probability that the HIX will be below a certain value at a given maturity. Hence, we can use such a HIX model to have an idea

\*Corresponding author. Email: [florence.guillaume@uantwerpen.be](mailto:florence.guillaume@uantwerpen.be)

of the probability that the dispersion strategy will be profitable. In particular, computing HIX forecasts can be useful to manage the risk of the dispersion strategy by computing, for instance, its profit and loss. By performing such an analysis, we can avoid the situation where we enter a dispersion strategy because the HIX is historically high, but is not expected to decrease in the near future, due, for instance, to a low mean-reverting rate.

In this paper, we consider the observed values of the HIX for the period ranging from January 2000 until October 2009. From the historical trend, we conclude that a suitable model for the HIX should have (to some extent) a mean-reverting behaviour. Popular mean-reverting processes are the Cox–Ingersoll–Ross (CIR) (see, for instance, Cox *et al.* 1985) and the Vasicek process. These two mean-reverting processes lead to tractable models since their transition density is given analytically. However, these processes can take values outside the interval  $[0, 1]$ . Therefore, we use a function to map the definition domain of the underlying mean-reverting process to the unit interval. Applying such a mapping function to a mean-reverting process results in an Itô process and preserves, to some extent, the mean-reverting behaviour of the underlying process. The proposed diffusion models are tractable since their transition density is known analytically so that they can be calibrated to HIX time series in a straightforward way. Remark that the proposed models can be used to model other stochastic processes lying between 0 and 1. Besides, Teng *et al.* (2015) adopted a similar methodology to build stochastic correlation models.

As a numerical study, we assess the ability of different mean-reverting processes and mapping functions to model the HIX for a period spanning from January 2000 until October 2009. The results indicate that the choice of the underlying mean-reverting process is more important than the choice of the mapping function. Indeed, compared with the mapping function, the mean-reverting process impacts the likelihood more severely. Moreover, the results suggest that the CIR process is preferred over the Vasicek process to drive the HIX process, at least for the time period under investigation. In order to investigate the impact of the 2007–2008 credit crisis, we split the HIX time series in a pre-crisis and a post-crisis series and calibrate the proposed models to these two time series separately. We find that during the credit crisis, the long-run mean of the HIX is higher and the mean-reversion rate is smaller. This observation may indicate that a market in distress is likely to have a *comonotonicity cluster*. Indeed, such a market situation is typically characterized by high values for the HIX, indicating that the market is close to be comonotonic, whereas the speed at which the market relaxes is relatively small. In such a situation, one has to be careful when entering a dispersion trade, because such a trading strategy is only profitable if the HIX decreases sufficiently during the investment horizon.

The rest of the paper is organized as follows. In section 2, we introduce the Herd Behaviour Index or HIX and in section 3 we propose a panoply of diffusion models for the HIX. A detailed statistical study of the HIX time series for the period spanning from January 2000 until October 2009 together with the behaviour of the calibration results are given in section 4. In section 5, we consider the predictive power of the proposed stochastic diffusion models. Section 6 concludes.

## 2. The herd behaviour index

### 2.1. The financial market

We assume a financial market<sup>†</sup> where  $n$  different (dividend or non-dividend paying) stocks, labelled from 1 to  $n$ , are traded. Assume that the current time is  $t = 0$ . The price at time  $t$ ,  $0 \leq t \leq T$ , of stock  $i$  is denoted by  $S_t^{(i)}$ . The market index is composed of a linear combination of the  $n$  underlying stocks. Denoting the price of the index at time  $t$  by  $S_t$ ,  $0 \leq t \leq T$ , we have that

$$S_t = w_1 S_t^{(1)} + w_2 S_t^{(2)} + \dots + w_n S_t^{(n)}, \quad (2.1)$$

where  $w_i$ ,  $i = 1, 2, \dots, n$ , are positive weights that are fixed up front.

It is assumed that the financial market is arbitrage-free and that there exists a pricing measure  $\mathbb{Q}$ , equivalent to the physical probability measure  $\mathbb{P}$ . For simplicity in notation and terminology, we consider a constant deterministic interest rate  $r$ . We assume that there is an options exchange where options on the individual stocks and the index are traded. The price of a vanilla call and put option with strike  $K$  and maturity  $T$  is denoted by  $C_i[K, T]$  and  $P_i[K, T]$ , respectively. These prices can be expressed as risk-neutral expectations of the discounted pay-offs:

$$\begin{aligned} C_i[K, T] &= e^{-rT} \mathbb{E} \left[ (S_T^{(i)} - K)_+ \right], \\ P_i[K, T] &= e^{-rT} \mathbb{E} \left[ (K - S_T^{(i)})_+ \right]. \end{aligned}$$

Similar notations are used for index call and put options, resulting in the following expressions:

$$C[K, T] = e^{-rT} \mathbb{E} [(S_T - K)_+], \quad (2.2)$$

$$P[K, T] = e^{-rT} \mathbb{E} [(K - S_T)_+]. \quad (2.3)$$

If no confusion is possible, we will omit the index  $T$  and write, for example,  $C[K]$ ,  $S$  and  $S^{(i)}$  instead of  $C[K, T]$ ,  $S_T$  and  $S_T^{(i)}$ , respectively.

### 2.2. Index options and perfect herd behaviour

In this subsection, we consider the *comonotonic market*. In this particular market situation, the stock price vector  $(S^{(1)}, S^{(2)}, \dots, S^{(n)})$  can be described as follows:

$$\left( S^{(1)}, S^{(2)}, \dots, S^{(n)} \right) \stackrel{d}{=} \left( F_1^{-1}(U), F_2^{-1}(U), \dots, F_n^{-1}(U) \right), \quad (2.4)$$

where  $F_i$  denotes the cdf of the random variable  $S^{(i)}$  and  $U$  a Uniform  $(0, 1)$  random variable. The comonotonic stock market index is denoted by  $S^c$  and defined as follows

$$S^c \stackrel{d}{=} F_1^{-1}(U) + F_2^{-1}(U) + \dots + F_n^{-1}(U). \quad (2.5)$$

Taking into account that the inverse cdf  $F_i^{-1}$  is a non-decreasing function, comonotonicity of the random vector  $(S^{(1)}, S^{(2)}, \dots, S^{(n)})$  implies that the price levels at time  $T$  are moving ‘perfectly together’. Indeed, the random source  $U$  is driving all the  $n$  stock prices and each stock is a non-decreasing function of  $U$ . Note that for each  $i = 1, 2, \dots, n$ , we have that

<sup>†</sup>We use the common approach to describe the financial market via a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ .

$F_i^{-1}(U) \stackrel{d}{=} S^{(i)}$ , which shows that the comonotonic index  $S^c$  and the real index  $S$  are composed of the same  $n$  stocks, but the stock prices in the index  $S^c$  are assumed to be comonotonic. For an extensive overview of the theory of comonotonicity, we refer the reader to [Dhaene et al. \(2002a\)](#). Financial and actuarial applications are described in [Dhaene et al. \(2002b\)](#). An updated overview of applications of comonotonicity can be found in [Deelstra et al. \(2011\)](#).

Consider now an index option written on  $S^c$ . In general, it will not be possible to invest in such a comonotonic index option, because the stock market index  $S$  will not behave in a comonotonic way and, as a result,  $S$  will be different from  $S^c$ . However, we can determine the prices  $C^c[K]$  and  $P^c[K]$  which would be observed for an index call and put option, respectively, provided the market is comonotonic. Indeed, one can prove that the comonotonic index option prices can be expressed as linear combinations of vanilla option prices:

$$C^c[K] = \sum_{i=1}^n w_i C_i[K_i^*], \quad (2.6)$$

$$P^c[K] = \sum_{i=1}^n w_i P_i[K_i^*], \quad (2.7)$$

where the strike prices  $K_i^*$  are given by

$$K_i^* = F_{X_i}^{-1(\alpha_K)}(F_{S^c}(K)), \quad i = 1, 2, \dots, n \quad (2.8)$$

and where  $\alpha_K$  is any element in  $[0, 1]$  such that (see, for instance, [Dhaene et al. 2002a](#))

$$F_{S^c}^{-1(\alpha_K)}(F_{S^c}(K)) = K. \quad (2.9)$$

Vanilla options are traded and their prices can be observed in the market. As first shown by [Breedon and Litzenberger \(1978\)](#), the risk-neutral distribution function  $F_i$  of  $X_i$  and the call option curve  $C_i$  are related by the following equation

$$F_i(x) = 1 + e^{rT} C_i'[x+], \quad (2.10)$$

where  $C_i'[x+]$  is the right derivative of  $C_i$  at  $x$ . The comonotonic cdf  $F_{S^c}$  can be determined as follows:

$$F_{S^c}(x) = \max \left\{ p \in (0, 1) \mid \sum_{i=1}^n w_i F_{X_i}^{-1}(p) \leq x \right\}. \quad (2.11)$$

Combining expressions (2.6), (2.7) (2.10) and (2.11), we conclude that if vanilla option prices can be observed for any strike, the comonotonic prices  $C^c[K]$  and  $P^c[K]$  can be determined in a model-free way using the vanilla option curves. However, one can also approximate the comonotonic index option prices  $C^c[K]$  and  $P^c[K]$  in the so-called finite market case, i.e. in the situation where only finitely many option prices can be observed; see e.g. [Linders et al. \(2012\)](#).

The real index option price is constrained from above by the comonotonic index option price:

$$C[K] \leq C^c[K] \quad \text{and} \quad P[K] \leq P^c[K], \quad \text{for any } K \geq 0. \quad (2.12)$$

In case the stock prices move in a comonotonic way, they will go simultaneously up, or simultaneously down, resulting in a maximal price for an index call and an index put, respectively. A proof of statement (2.12) can be found in [Kaas et al. \(2000\)](#).

### 2.3. The herd behaviour index

A measure for the implied degree of co-movement was introduced in [Dhaene, Linders, et al. \(2012\)](#). This new proxy combines the information conveyed by the whole index option curve and is determined by comparing the variance  $\text{Var}[S]$  of the real sum  $S$  with the comonotonic variance  $\text{Var}[S^c]$ .

*Definition 2.1* (Degree of herd behaviour). Consider the random vector  $\underline{S}$  representing the stock prices  $S^{(i)}$ ,  $i = 1, 2, \dots, n$  at time  $T$ . The degree of herd behaviour of the random vector  $\underline{S}$  is measured by the function  $\pi[\underline{S}]$  as follows

$$\pi[\underline{S}] = \frac{\text{Var}[S]}{\text{Var}[S^c]},$$

provided these variances are finite.

The number  $\pi[\underline{S}]$  takes values between 0 and 1. Moreover, one can prove that a value of 1 characterizes the situation where the components are comonotonic; see e.g. [Cheung et al. \(2013\)](#).

The variance  $\text{Var}[S]$  can always be expressed as a weighted sum of out-of-the-money index option prices; see e.g. [Carr and Madan \(2001\)](#) or [Cheung et al. \(2013\)](#). Similarly,  $\text{Var}[S^c]$  can be expressed as a mixture of comonotonic out-of-the-money index options. Combining these observations we find

$$\pi[\underline{S}] = \frac{\int_{-\infty}^{+\infty} Q[K] dK}{\int_{-\infty}^{+\infty} Q^c[K] dK}, \quad (2.13)$$

where  $Q[K] \equiv Q[K, T]$  denotes the price of an out-of-the-money index option with strike  $K$  and maturity  $T$ . The degree of herd behaviour is a measure for the standardized distance between the actual sum  $S$  and the comonotonic sum  $S^c$  through the index and vanilla option curves.

In practice, option prices can only be observed for a finite number of strike prices such that one can only compute an approximation of (2.13). Let us denote the first traded put option strike price below  $\mathbb{E}[S]$  by  $K_0$ . The traded index put option strikes below  $\mathbb{E}[S]$  are denoted by  $K_{-i}$ ,  $i = 0, 1, \dots, l$  with  $K_{-l} < K_{-l+1} < \dots < K_{-1} < K_0 \leq \mathbb{E}[S]$ , whereas the traded index call option strikes above  $\mathbb{E}[S]$  are denoted by  $K_i$ ,  $i = 1, \dots, h$  with  $\mathbb{E}[S] < K_1 < \dots < K_{h-1} < K_h$ . We then adopt the VIX methodology proposed by the CBOE (Chicago Board Options Exchange) to approximate the degree of herd behaviour in case of a discrete range of listed strikes, measured as the Herd Behaviour Index, or just HIX[T].

*Definition 2.2* Consider the random vector  $\underline{S}$  representing the stock prices  $S^{(i)}$ ,  $i = 1, 2, \dots, n$  at time  $T$ . The  $T$ -year implied Herd Behaviour Index is defined as

$$\text{HIX}[T] = \frac{2e^{rT} \sum_{i=-l}^h \Delta K_i Q[K_i] - (\mathbb{E}[S] - K_0)^2}{2e^{rT} \sum_{i=-l}^h \Delta K_i Q^c[K_i] - (\mathbb{E}[S] - K_0)^2}. \quad (2.14)$$

### 3. Diffusion models for herd behaviour indices

This section features different plausible stochastic models for herd behaviour indices such as the HIX defined in section 2. These models are obtained by combining some mean-reverting process such as the Vasicek or the CIR process with some function mapping the definition domain of the mean-reverting



process to the unit interval. The such obtained Itô processes preserve, to some extent, the mean-reverting trend of the underlying process while satisfying the fundamental properties of the HIX index which has  $[0, 1]$  as definition domain. The idea of combining a mapping function with a Brownian motion first appeared in [van Emmerich \(2006\)](#) to model stochastic correlation but was not further investigated due to its lack of intuitive interpretation. However, this approach gets the advantage of an analytical expression for the conditional transition probability, as opposed to the methodology consisting in modelling directly the desired stochastic process by the following SDE

$$dY_t = \kappa(\eta - Y_t)dt + \zeta\sqrt{1 - Y_t^2}dW_t, \quad Y_0 \in (0, 1),$$

as proposed in [van Emmerich \(2006\)](#). Indeed, van Emmerich could only derive some expression for the stationary transition density whereas the quantity of interest to calibrate financial time series is typically the conditional transition density. On the other hand, the mapping approach considered here benefits from the analytical expression of the transition probability under popular mean-reverting processes, such as the Vasicek and the CIR model. Indeed, considering monotonic mapping functions allows for a direct computation of the transition density in terms of the one of the underlying process by means of the Jacobian formula. Moreover, the numerical study will highlight the fact that, for a given underlying Itô process, the choice of the mapping function does not really matter. This mapping approach has been considered in [Teng et al. \(2015\)](#) to model stochastic correlation. However, Teng et al. focused on the stationary density rather than on the transition density. Since the first aim of modelling herd behaviour indices such as the HIX is to predict their future evolution for the coming days, the relevant parameters are the ones inferred from maximizing the conditional transition density rather than the ones inferred from the stationary density. Moreover, [Teng et al. \(2015\)](#) only worked out one particular correlation model built upon an hyperbolic tangent mapping function and the Vasicek process as underlying whereas we consider and investigate different model sub-classes.

### 3.1. Mapping functions and conditional transition densities

Let us consider an Itô process given by

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t, \quad X_0 = x_0, \quad (3.1)$$

where  $W = \{W_t, t \geq 0\}$  is a Brownian motion, and some function  $f(x)$  twice continuously differentiable on  $\mathbb{R}$ . Applying Itô's Lemma gives the differential form of  $Y_t = f(X_t)$ :

$$dY_t = \tilde{a}(t, X_t)dt + \tilde{b}(t, X_t)dW_t,$$

where

$$\begin{aligned} \tilde{a}(t, X_t) &= a(t, X_t)\frac{\partial f}{\partial x}(X_t) + \frac{1}{2}b^2(t, X_t)\frac{\partial^2 f}{\partial x^2}(X_t) \\ \text{and } \tilde{b}(t, X_t) &= b(t, X_t)\frac{\partial f}{\partial x}(X_t). \end{aligned} \quad (3.2)$$

By construction, herd behaviour indices only take value in the unit interval  $[0, 1]$  such that the choice of  $f$  is restricted to functions mapping the definition domain of  $X$  to the unit interval. In particular, we will consider the four mapping functions

represented on figure 1, namely

$$\begin{aligned} f_1(x) &= \frac{1}{1 + \exp(-x)}, & f_2(x) &= \frac{\tanh(x) + 1}{2}, \\ f_3(x) &= 1 - \exp(-x) & \text{and } f_4(x) &= \tanh(x). \end{aligned} \quad (3.3)$$

The mapping functions  $f_1$  and  $f_2$  can be combined with any Itô process defined on  $(-\infty, \infty)$  whereas the mapping functions  $f_3$  and  $f_4$  can be associated with any Itô process defined on  $[0, \infty)$ .

Figure 1 clearly indicates that the four functions are monotonic. Hence, the transition probability<sup>†</sup> of  $Y_t = f(X_t)$  can be directly obtained from the transition probability of  $X_t$  by noting that, for any monotonic function  $f$ ,

$$p_Y(y|y_0; \theta, \Delta t) = \left| \frac{d}{dy} f^{-1}(y) \right| p_X(f^{-1}(y)|f^{-1}(y_0); \theta, \Delta t),$$

where  $\theta$  is the set of model parameters. The inverse of the mapping functions is

$$\begin{aligned} f_1^{-1}(y) &= \log\left(\frac{y}{1-y}\right), & f_2^{-1}(y) &= \frac{1}{2}\log\left(\frac{y}{1-y}\right), \\ f_3^{-1}(y) &= \log\left(\frac{1}{1-y}\right) & \text{and } f_4^{-1}(y) &= \frac{1}{2}\log\left(\frac{1+y}{1-y}\right), \\ & & & y \in (0, 1). \end{aligned}$$

If the process  $Y$  is defined using the mapping function  $f_i$ , i.e. if  $Y_t = f_i(X_t)$  for any  $t \geq 0$ , then we denote the transition density of  $Y$  by  $p_Y^{(i)}$ . We then find that for the different mapping functions, the corresponding transition densities are given by:

$$\begin{aligned} p_Y^{(1)}(y|y_0; \theta, \Delta t) &= \frac{1}{y(1-y)} p_X\left(\log\left(\frac{y}{1-y}\right) \middle| \log\left(\frac{y_0}{1-y_0}\right); \theta, \Delta t\right); \end{aligned} \quad (3.4)$$

$$\begin{aligned} p_Y^{(2)}(y|y_0; \theta, \Delta t) &= \frac{1}{2y(1-y)} p_X\left(\frac{1}{2}\log\left(\frac{y}{1-y}\right) \middle| \frac{1}{2}\log\left(\frac{y_0}{1-y_0}\right); \theta, \Delta t\right); \end{aligned} \quad (3.5)$$

and

$$\begin{aligned} p_Y^{(3)}(y|y_0; \theta, \Delta t) &= \frac{1}{1-y} p_X\left(\log\left(\frac{1}{1-y}\right) \middle| \log\left(\frac{1}{1-y_0}\right); \theta, \Delta t\right); \end{aligned} \quad (3.6)$$

$$\begin{aligned} p_Y^{(4)}(y|y_0; \theta, \Delta t) &= \frac{1}{1-y^2} p_X\left(\frac{1}{2}\log\left(\frac{1+y}{1-y}\right) \middle| \frac{1}{2}\log\left(\frac{1+y_0}{1-y_0}\right); \theta, \Delta t\right). \end{aligned} \quad (3.7)$$

The process  $X = \{X_t, t \geq 0\}$  is chosen among the set of mean-reverting processes for which the transition probability is known analytically. We will thus assume that the underlying process  $X$  follows either a Vasicek or a CIR process. Let us first consider the situation where the process  $X$  follows a Vasicek process. In this case

$$dX_t = \kappa(\eta - X_t)dt + \zeta dW_t, \quad X_t \in (-\infty, \infty), \quad (3.8)$$

where  $W = \{W_t, t \geq 0\}$  is a standard Brownian motion. The Vasicek process is of the generic form (3.1) with  $a(t, X_t) =$

<sup>†</sup>The transition density is defined as the conditional density that  $Y$  is at level  $y$  at time  $t + \Delta t$  given that it is at level  $y_0$  at time  $t$ .

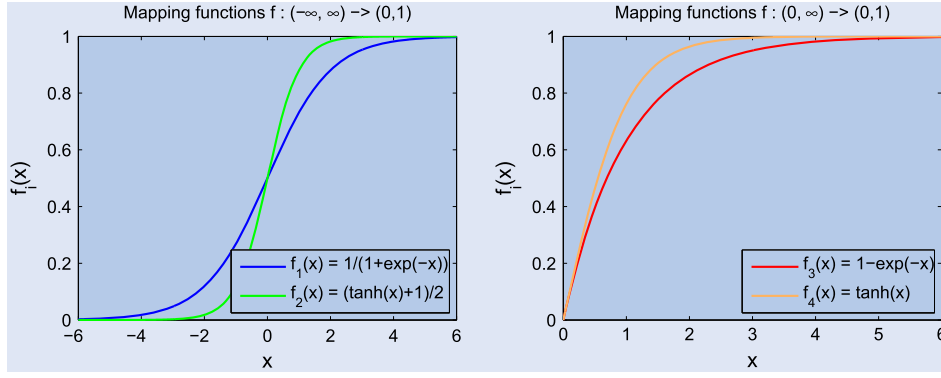


Figure 1. Mapping functions.

$\kappa(\eta - X_t)$  and  $b(t, X_t) = \zeta$ . Furthermore, the transition density is given in closed form; see for instance [Aït-Sahalia \(1999\)](#):

$$p_X(x|x_0; \theta, \Delta t) = \left(\frac{\pi\xi^2}{\kappa}\right)^{-1/2} \exp\left(-\frac{(x - \eta - (x_0 - \eta)\exp(-\kappa\Delta t))^2 \frac{\kappa}{\xi^2}}{2}\right),$$

where  $\xi^2 = \zeta^2(1 - \exp(-2\kappa\Delta t))$ .

In case the process  $X$  is assumed to be a CIR process, we have that

$$dX_t = \kappa(\eta - X_t)dt + \zeta\sqrt{X_t}dW_t, \quad X_t \in (0, \infty). \quad (3.9)$$

The CIR process can be written as (3.1) with  $a(t, X_t) = \kappa(\eta - X_t)$  and  $b(t, X_t) = \zeta\sqrt{X_t}$ . Furthermore, the transition density is given by (see, for instance, [Aït-Sahalia 1999](#))

$$p_X(x|x_0; \theta, \Delta t) = c \exp(-u - v) \left(\frac{v}{u}\right)^{q/2} I_q(2\sqrt{uv}),$$

where  $q \equiv \frac{2\kappa\eta}{\zeta^2} - 1 \geq 0$ ,  $c \equiv \frac{2\kappa}{\zeta^2(1 - \exp(-\kappa\Delta t))}$ ,  $u = cx_0 \exp(-\kappa\Delta t)$  and  $v = cx$  and where  $I_q$  is the modified Bessel function of the first kind of order  $q$ . We do not consider the more general CEV process which nests the Vasicek and CIR processes as special cases since the transition probability is then not known analytically, and has, thus, to be approximated numerically.

The parameter set  $\theta = \{\kappa, \eta, \zeta\}$  can be efficiently estimated by the maximum likelihood estimation technique. Given the Markovian nature of the Vasicek process (3.8) and the CIR process (3.9), the maximum likelihood (ML) estimate  $\hat{\theta}$  is obtained by maximizing

$$\log L_N(\theta) = \frac{1}{M-1} \sum_{i=1}^{M-1} \log(p_Y(y_{t_{i+1}}|y_{t_i}; \theta, \Delta t_{i+1})),$$

where  $M$  denotes the number of observations in the time series and  $y_{t_1}, y_{t_2}, \dots, y_{t_M}$  denote the observations of the process  $Y$ . The time grid satisfies  $t_1 < t_2 < \dots < t_M$  and  $\Delta t_{i+1} = t_{i+1} - t_i$ . For sake of comparison, we will also consider the Brownian motion (with drift)

$$dX_t = \eta dt + \zeta dW_t, \quad X_t \in (-\infty, \infty) \quad (3.10)$$

as underlying process. The transition density is then given by

$$p_X(x|x_0; \theta, \Delta t) = \frac{1}{\sqrt{2\pi\zeta^2\Delta t}} \exp\left(-\frac{(x - x_0 - \eta\Delta t)^2}{2\zeta^2\Delta t}\right).$$

Combining the mapping functions (3.3) with the three Itô processes leads to six model specifications as shown in table 1, where the coefficients  $\tilde{a}(t, X_t)$  and  $\tilde{b}(t, X_t)$  are obtained by plugging (3.3) into (3.2).

### 3.2. Properties of the diffusion models

Let assume that the process  $X$  driving the herd behaviour index process  $Y$  is a mean-reverting process and can be expressed as follows:

$$dX_t = \kappa(\eta - X_t)dt + \zeta v(X_t)dW_t, \quad X_t \in (c, +\infty). \quad (3.11)$$

Note that if  $c = -\infty$  and  $v(x) = 1$  for all  $x \in \mathbb{R}$ , we find back the Vasicek model, while for the CIR model, we need to set  $c = 0$  and  $v(x) = \sqrt{x}$  for  $x > 0$ . From equations (3.1) and (3.2), we find that  $Y_t = f(X_t)$  can be expressed as follows:

$$dY_t = \left(\kappa(\eta - X_t) \frac{\partial f}{\partial x}(X_t) + \frac{1}{2}\zeta^2 v^2(X_t) \frac{\partial^2 f}{\partial x^2}(X_t)\right) dt + \zeta v(X_t) \frac{\partial f}{\partial x}(X_t) dW_t \quad (3.12)$$

$$= \left(\kappa a_1(X_t)(\eta - X_t) + \zeta^2 a_2(X_t)\right) dt + \zeta b_1(X_t) dW_t, \quad (3.13)$$

where  $a_1(X_t) = \frac{\partial f}{\partial x}(X_t)$ ,  $a_2(X_t) = \frac{1}{2}v^2(X_t) \frac{\partial^2 f}{\partial x^2}(X_t)$  and  $b_1(X_t) = v(X_t) \frac{\partial f}{\partial x}(X_t)$ . Figure 2 shows the mean-reverting distortion function  $a_1$ , the drift function  $a_2$  and the diffusion function  $b_1$  for the models built upon a Vasicek or a CIR process.†

Equation (3.13) indicates that the function  $a_1$  distorts the mean-reverting rate  $\kappa$  of the underlying process. For the functions  $f_3$  and  $f_4$ , the distortion function  $a_1$  decreases with  $X_t$ , i.e. with  $Y_t$  since the mapping functions are monotonically increasing, whereas it first increases (until a HIX level of approximately 0.5) and then decreases with  $X_t$ , or equivalently  $Y_t$  when the mapping function  $f_1$  or  $f_2$  is used.

†Note that for models 4 and 5 (i.e. for the models built upon a Brownian motion), equation (3.13) becomes

$$dY_t = \left(a_1(X_t)\eta + \zeta^2 a_2(X_t)\right) dt + \zeta b_1(X_t) dW_t,$$

where the functions  $a_1$ ,  $a_2$  and  $b_1$  coincide with the ones relative to the Vasicek model built upon the same mapping function.

Table 1. Herd behaviour index model specifications.

	Underlying process	Mapping function	$\tilde{a}(t, X_t)$	$\tilde{b}(t, X_t)$
1	Vasicek	$f_1(x)$	$\kappa \frac{\exp(X_t)}{(\exp(X_t)+1)^2} (\eta - X_t) + \frac{\zeta^2}{2} \frac{\exp(X_t)(1-\exp(X_t))}{(\exp(X_t)+1)^3}$	$\zeta \frac{\exp(X_t)}{(\exp(X_t)+1)^2}$
2	Vasicek	$f_2(x)$	$\frac{\kappa}{2} \operatorname{sech}^2(X_t) (\eta - X_t) - \frac{\zeta^2}{2} \tanh(X_t) \operatorname{sech}^2(X_t)$	$\frac{\zeta}{2} \operatorname{sech}^2(X_t)$
3	CIR	$f_3(x)$	$\kappa \exp(-X_t) (\eta - X_t) - \frac{\zeta^2}{2} \exp(-X_t) X_t$	$\zeta \exp(-X_t) \sqrt{X_t}$
4	CIR	$f_4(x)$	$\kappa \operatorname{sech}^2(X_t) (\eta - X_t) - \zeta^2 \tanh(X_t) \operatorname{sech}^2(X_t) X_t$	$\zeta \operatorname{sech}^2(X_t) \sqrt{X_t}$
5	BM	$f_1(x)$	$\eta \frac{\exp(X_t)}{(\exp(X_t)+1)^2} + \frac{\zeta^2}{2} \frac{\exp(X_t)(1-\exp(X_t))}{(\exp(X_t)+1)^3}$	$\zeta \frac{\exp(X_t)}{(\exp(X_t)+1)^2}$
6	BM	$f_2(x)$	$\frac{\eta}{2} \operatorname{sech}^2(X_t) - \frac{\zeta^2}{2} \tanh(X_t) \operatorname{sech}^2(X_t)$	$\frac{\zeta}{2} \operatorname{sech}^2(X_t)$

The function  $b_1$  affects the volatility of the HIX process  $Y_t$ . The weight function  $b_1$  first increases and then decreases with  $X_t$ , or equivalently  $Y_t$ , whatever the model settings. It reaches a maximum for a HIX level of 0.5 (models 1 and 2) or somewhat lower (0.3935 and 0.4791 under model 3 and 4, respectively), whereas it tends to zero when the HIX approaches its boundaries, whatever the mapping function under investigation. This indicates that the closer the HIX to its boundaries, the less volatile the HIX; in the limit, the process  $Y_t$  becomes deterministic.

The function  $a_2$  is an additional drift function added to the mean-reverting trend of the process. For models 1 and 2, the drift function is an odd function ( $a_2(-x) = -a_2(x)$ ); it first increases until  $Y_t \approx 0.2113$  before decreasing until  $Y_t \approx 0.7887$  and finally, increasing. The drift function is then positive for a HIX level lower than 0.5 and negative for a HIX level higher than 0.5. Under models 3 and 4, the additional drift function first decreases and then increases with  $X_t$ , or equivalently  $Y_t$ . Its minimum value is reached for a HIX index around 0.63 and 0.77 under model 3 and model 4, respectively. In that case, the additional drift is negative whatever the value of the herd behaviour index.

The shape of the different functions clearly indicates that the models built upon the same mean-reverting process share similar characteristics. This can easily be explained by the fact that the mapping functions  $f_1$  and  $f_2$  or  $f_3$  and  $f_4$  cannot deviate too much from each others since they have to map the same portion of the real line to the unit interval and to be monotonic (see also figure 1).

#### 4. The HIX: a mean-reverting index

In this section, we calibrate the different models featured in section 3 to the HIX index for a time period spanning from January 2000 until October 2009, and including therefore, the 2007–2008 credit crunch (see figure 3). In a first time, we consider three time series windows: the first one consists of the whole time series window, the second one is the pre-crisis (sub-)period and the last one the post-crisis (sub-)period.

The transition date between these two sub-periods is taken as 9 October 2007, when the S&P 500 index price reached a maximum value of 1565.15 US dollars. This date (or more precisely the next day) is considered as the first day of the credit crisis period and turns out to be close to the month of August 2007, which is seen by many as the actual start of the credit crunch.

##### 4.1. The HIX time series

Table 2 shows the statistics of the HIX index and the HIX daily changes for the three above-mentioned reference time series. From the results we observe that the average level of the HIX during the post crisis period is higher than during the pre-crisis. Hence, stock prices tend to move more strongly together during distress periods. These findings are in line with previous empirical studies (see, for instance, Skintzi and Refenes 2005, Dhaene, Dony, et al. 2012 and Rubbaniy et al. 2013). Besides, the autocorrelation function (ACF) of the different original time series is significantly positive, at least up to five lags. At short lags, the ACF is higher during the credit crisis period than before the crisis, which suggests that, for a short time horizon, comonotonicity tends to cluster more during distress periods such that the probability of then observing a high HIX value in the near future is relatively high. To further test for comonotonicity clustering, the ACF of the original time series for the first 100 lags is shown on figure 4. The ACF remains significantly positive up to 79 lags for the post-crisis sub-period and up to more than 100 lags for the pre-crisis and the whole period. This highlights a long-run dependence in the herd behaviour: periods characterized by a low (high) degree of comonotonicity in the market are more likely to be followed by periods of low (high) degree of comonotonicity. We also see that, although higher at short lags, the ACF of the HIX becomes lower for the post-crisis period than for the whole period and the pre-crisis period after 13 and 32 lags, respectively. Moreover, the significant negative ACF at the first lag for the differenced series reported in table 2 suggests a mean-reverting behaviour for the HIX index, whatever the time series window. These findings imply that a mean-reverting

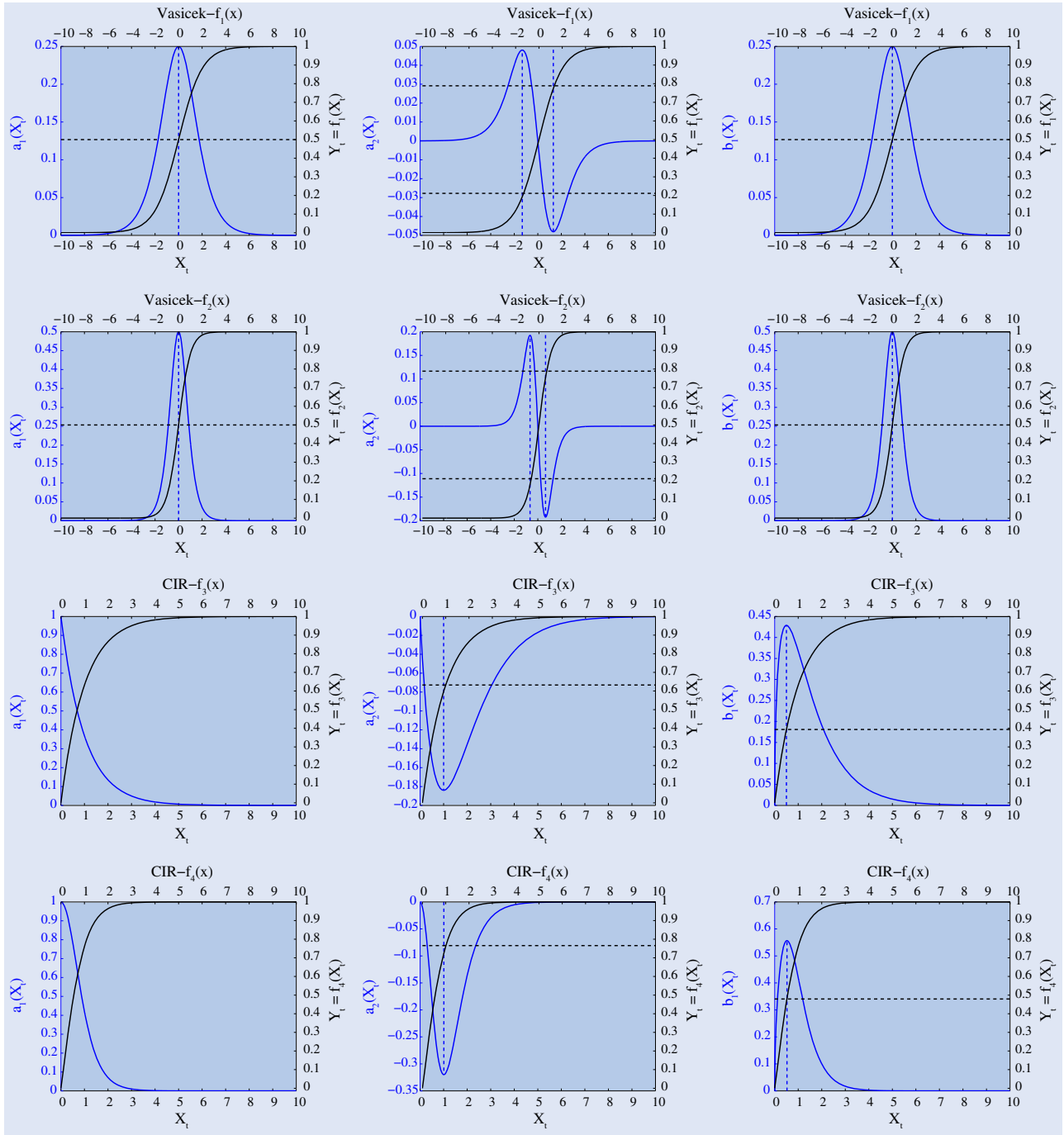


Figure 2. The functions  $a_1$ ,  $a_2$  and  $b_1$  for different mapping functions and underlying processes  $X$ .

process should be more appropriate than a Brownian motion for modelling the underlying process  $X$ . Indeed, in section 3.2, we showed that, although altered by the distortion function  $a_1$ , the mean-reverting trend is transferred from the underlying process  $X$  to the HIX process  $Y$ .

#### 4.2. ML estimation

Table 3 shows the ML estimate of the model parameters for the HIX time series, and this for the three above-mentioned reference time periods. Table 3 clearly indicates that the two map-

ping functions  $f : (-\infty, +\infty) \rightarrow (0, 1)$ , i.e.  $f_1(x)$  and  $f_2(x)$ , lead to the same ML goodness of fit. This can be explained by the fact that the two models (i.e. models 1 and 2 as well as models 5 and 6) are actually equivalent since  $f_2(x) = f_1(2x)$ . The equivalence is obtained by further setting  $\kappa^{(2)} = \kappa^{(1)}$ ,  $\eta^{(2)} = \frac{1}{2}\eta^{(1)}$  and  $\zeta^{(2)} = \frac{1}{2}\zeta^{(1)}$ , where the superscripts  $(1)$  and  $(2)$  refer to model 1 or 5 and to model 2 or 6, respectively. Hence, in the following, we will only consider models 1, 3, 4 and 5. We also observe that, as expected from figure 2, the choice of the underlying process matters more than the shape of the mapping function since the likelihood



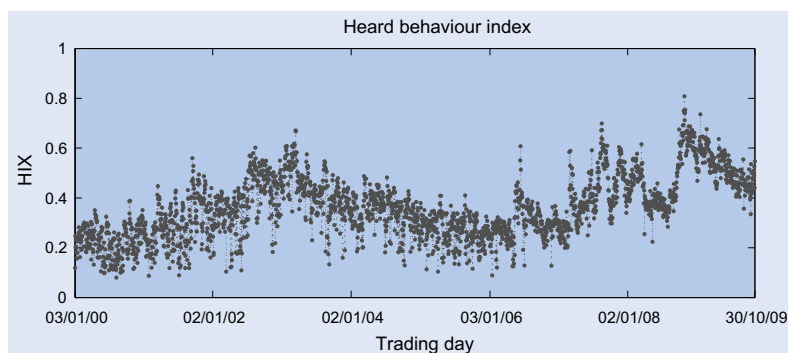


Figure 3. Evolution through time of the HIX index.

Table 2. HIX time series statistics.

Time series	$M$	Mean	Var	Skew	Kurt	ACF(1)	ACF(2)	ACF(3)	ACF(4)	ACF(5)
HIX (2000–2009)	2470	0.361336	0.015293	0.346408	2.725549	0.8544*	0.8295*	0.8102*	0.7838*	0.7677*
$\Delta$ HIX (2000–2009)	2469	0.000173	0.004416	0.162024	5.434606	-0.4133*	-0.0217	0.0255	-0.0352	0.0015
HIX (pre-crisis)	1950	0.326938	0.011324	0.386486	2.982369	0.7741*	0.7374*	0.7100*	0.6692*	0.6474*
$\Delta$ HIX (pre-crisis)	1948	0.000091	0.005097	0.149345	4.977160	-0.4180*	-0.0215	0.0305	-0.0421	0.0004
HIX (post-crisis)	520	0.490327	0.009112	0.267414	2.683532	0.8939*	0.8667*	0.8405*	0.8229*	0.7982*
$\Delta$ HIX (post-crisis)	519	0.000444	0.001873	0.313177	5.383072	-0.3668*	-0.0210	-0.0272	0.0373	0.0099

Note: Values with \* are significant at a level  $\alpha = 0.05$ .

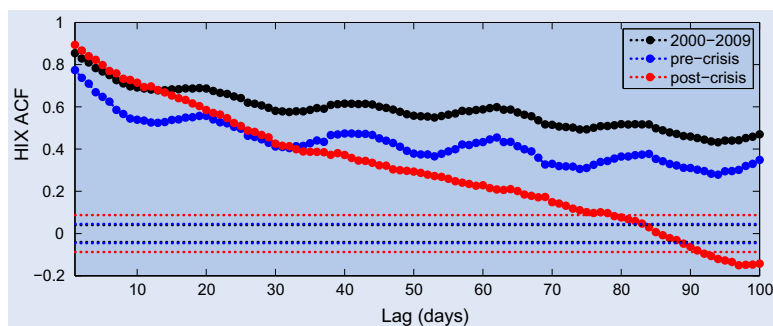


Figure 4. HIX index ACF.

Table 3. ML estimate for the HIX time series.

Model	2000–2009				Pre-crisis				Post-crisis			
	$\hat{\kappa}$	$\hat{\eta}$	$\hat{\zeta}$	$L(\hat{\theta})$	$\hat{\kappa}$	$\hat{\eta}$	$\hat{\zeta}$	$L(\hat{\theta})$	$\hat{\kappa}$	$\hat{\eta}$	$\hat{\zeta}$	$L(\hat{\theta})$
1	49.094	-0.60788	5.6531	3.520	77.014	-0.76330	6.3286	3.421	29.552	-0.02249	2.9964	5.644
2	49.094	-0.30394	2.8265	3.520	77.014	-0.38165	3.1643	3.421	29.552	-0.01125	1.4982	5.644
3	42.155	0.47083	2.7325	<b>3.761</b>	68.374	0.41002	3.0105	<b>3.588</b>	29.419	0.70013	1.7545	<b>5.691</b>
4	43.994	0.38795	2.2510	3.674	70.150	0.34501	2.4990	3.536	29.542	0.54990	1.3221	<b>5.691</b>
5	0	0.2228	5.3898	3.360	0	0.1469	5.8770	3.187	0	0.4657	2.9130	5.482
6	0	0.1114	2.6949	3.360	0	0.0735	2.9385	3.187	0	0.2328	1.4565	5.482

Note: The bold figures correspond to the best performing model for each sample period.

$L(\hat{\theta})$  differs more across the underlying models than across the mapping functions. Moreover, the results seem to indicate that the underlying process  $X$  should be characterized by a mean-reverting trend since the likelihood is higher for models

built upon the CIR, and, to a smaller extent, the Vasicek model than for models built upon the Brownian motion. This confirms the intuition inferred from the mean-reverting trend of the herd behaviour index highlighted by the HIX ACF shape.

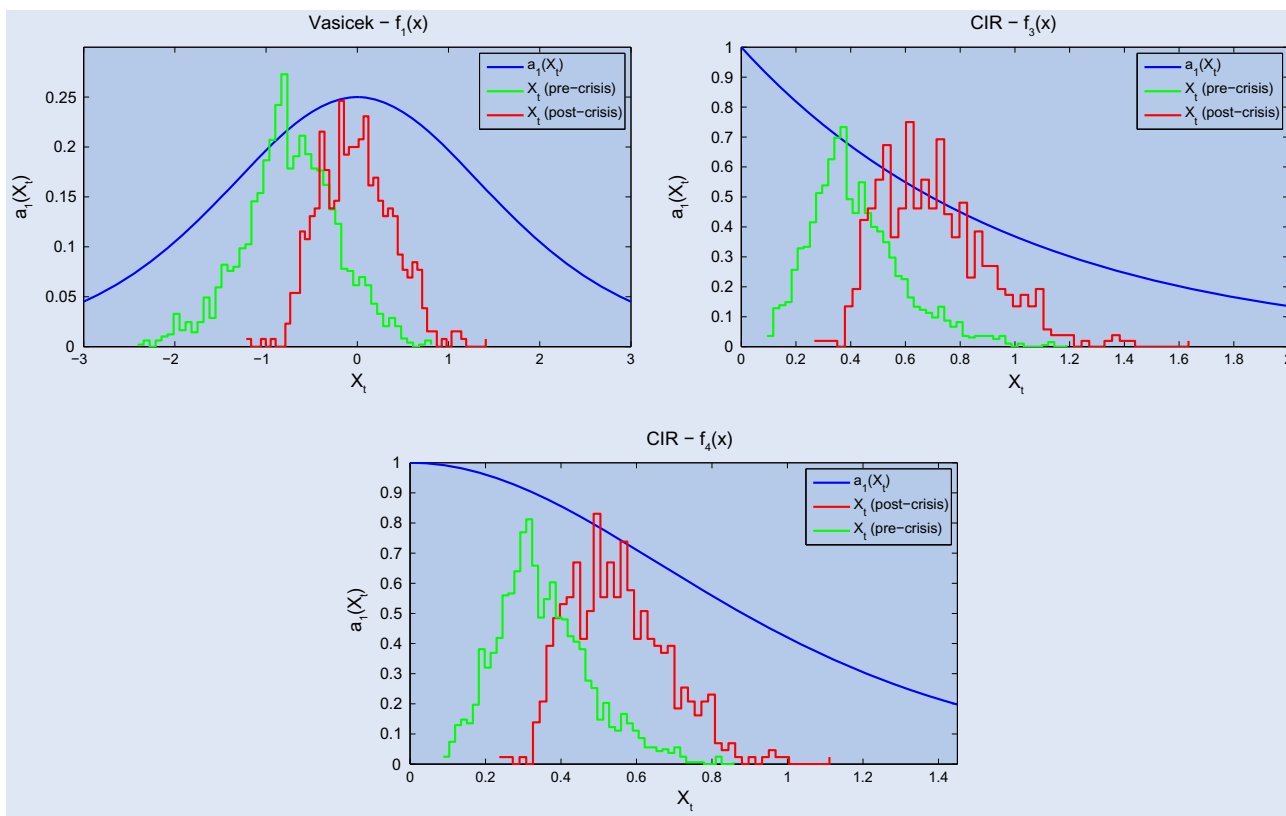


Figure 5. Mean-reverting rate function  $a_1$  for models 1 (upper left), 3 (upper right) and 4 (lower left).

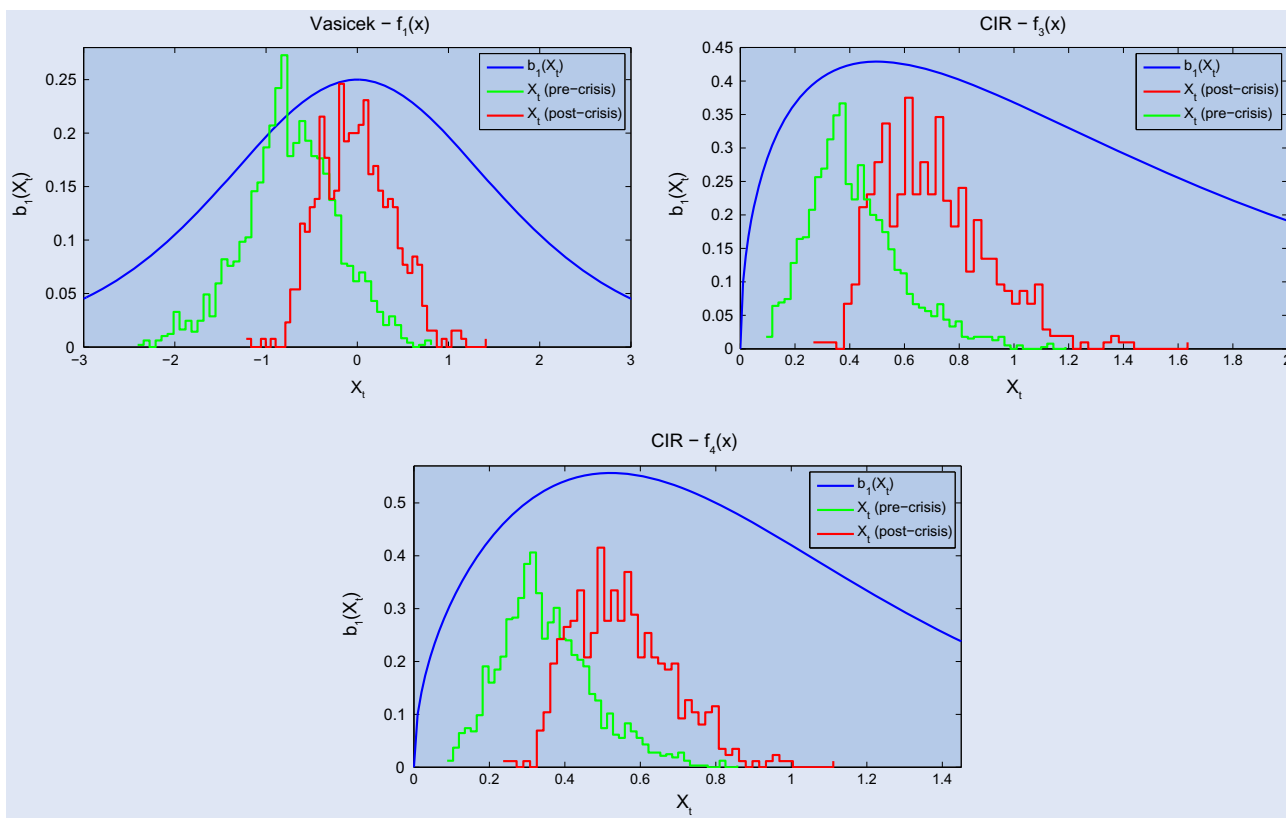


Figure 6. Volatility function  $b_1$  for models 1 and 5 (upper left), 3 (upper right) and 4 (lower left).

Table 4. ML estimate of  $\tilde{\eta} = f_i(\eta)$ .

Period	Model			
	1	2	3	4
2000–2009	0.3525	0.3525	0.3755	0.3696
Pre-crisis	0.3179	0.3179	0.3364	0.3319
Post-crisis	0.4944	0.4944	0.5035	0.5004

Note: The bold figures correspond to the best performing model for each sample period.

The ML estimate  $\hat{\kappa}$  is higher and lower, with respect to the whole sample period, during the pre-crisis and the post-crisis sub-period, respectively. Hence, the underlying process  $X$  and thus, to some extent the HIX process  $Y$  too, reverts faster towards its long-run mean during the pre-crisis period; indicating that the degree of herd behaviour is more persistent during distress periods. Indeed, in section 3.2, we showed that if  $X$  is a mean-reverting process, the process  $Y$  has again a mean-reverting trend, although altered by the distortion function  $a_1$ . Figure 5 shows the distortion function together with the histogram of the empirical process  $X$  before and during the credit crisis period. We clearly observe that, on average, the distortion function is closer to its maximum value during the post-crisis and the pre-crisis period for model 1 and for models 3 and 4, respectively. Hence, the distortion function amplifies, on average, the decrease in the mean-reversion speed of the HIX during the distress period under models built upon the CIR process. On the other hand, the effect of the distortion function tends to be opposite for models built upon the Vasicek process; the impact of the distortion function being then nevertheless quite limited due to its narrow value range. These findings are in line with the fact that, during the post-crisis period, the HIX ACF is higher at short lags and the HIX daily change ACF is smaller in absolute value at the first lag.

The ML estimate  $\hat{\zeta}$  is higher and lower, with respect to the whole time series window, during the pre-crisis and the post-crisis sub-period, respectively. In particular, during the crisis period,  $\hat{\zeta}$  is roughly half of its value before the crisis. Hence, the volatility of the HIX process is expected to decrease by half, *ceteris paribus*. Nevertheless, the sample variance during the pre-crisis sub-period is far from being four times bigger than during the post-crisis sub-period, as indicated in table 2. This might be explained by the fact that the weight function  $b_1$  has the opposite effect on the volatility of the process  $Y$  than the volatility parameter  $\zeta$ . Figure 6 shows the volatility weight function  $b_1$  together with the histogram of the empirical process  $X$  before and during the credit crisis period. For models 1 and 4, it is clear that, on average, the weight function is closer to its maximum value during the post-crisis period than during the pre-crisis period. Hence, the weight function then tends to be higher during the credit crunch, leading to an increase in the total volatility of  $Y$ . A similar conclusion can be drawn for model 3 since at the maximum of the weight function, the left slope is much sharper than the right slope and since the empirical values of the underlying process  $X$  tend to be lower before the crisis.

We also observe that for the mean-reverting processes (i.e. for the models built upon a Vasicek or a CIR process), the long-run mean of the herd behaviour index can be

approximated by  $\tilde{\eta} = f_i(\hat{\eta})$ . Indeed, comparing the sample mean of the three reference time series reported in table 2 with the ML estimate of  $\tilde{\eta}$  reported in table 4, the maximum relative difference amounts to 3.92 per cent only.

## 5. The predictability power of mapping models for the HIX

In order to assess the predictive performance of the HIX diffusion model with different specifications for the underlying process  $X$  and the mapping function  $f$ , we compute one-day- up to one-month-ahead forecasts of the HIX index for one-year daily observations of the HIX. Hence, we first calibrate each model settings to daily observations of the HIX for a one-year moving time series window by maximizing the likelihood function. Then, we simulate HIX forecasts using the ML model parameter estimates  $\hat{\theta}$  and replacing  $y_0$  by the last observation of the one-year period. The Monte Carlo simulation is based on the discretization of the underlying process  $X$ . More particularly, we proceed as follows for the illustrative example of model 1:

- (1) compute  $x_0$ :

$$x_0 = f_1^{-1}(y_0) = \log\left(\frac{y_0}{1 - y_0}\right);$$

- (2) simulate  $N$  one-day-ahead forecasts of the underlying process  $X$

$$X_{t+\Delta t}^{(i)} = X_t + \hat{\kappa}(\hat{\eta} - X_t)\Delta t + \hat{\zeta}\sqrt{\Delta t}\epsilon_i, \quad i = 1, \dots, N,$$

where  $\epsilon_i$  are i.i.d. standard normal random variables;

- (3) infer the one-day-ahead forecasts of the HIX process  $Y$ :

$$Y_{t+\Delta t}^{(i)} = f_1\left(X_{t+\Delta t}^{(i)}\right) = \frac{1}{1 + \exp\left(-X_{t+\Delta t}^{(i)}\right)},$$

$$i = 1, \dots, N;$$

- (4) the one-day-ahead forecast is obtained by averaging the path forecasts over the number of paths:

$$\hat{Y}_{t+\Delta t} = \frac{1}{N} \sum_{i=1}^N Y_{t+\Delta t}^{(i)};$$

- (5) repeat the same methodology to compute the  $k$ -days-ahead forecast for  $k = 1, \dots, m$ , where  $m$  denotes the number of days in the current month (i.e. the month following the one-year observation period); but where the value of  $y_0$  is updated to be equal to the last computed forecast, i.e. the  $(k - 1)$ -days-ahead forecast.

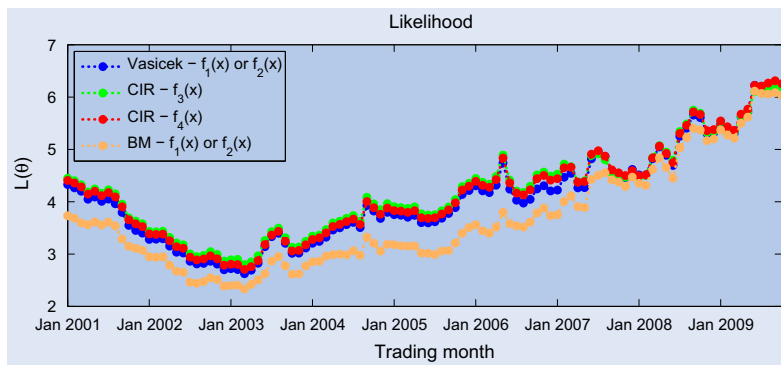


Figure 7. Maximum likelihood.

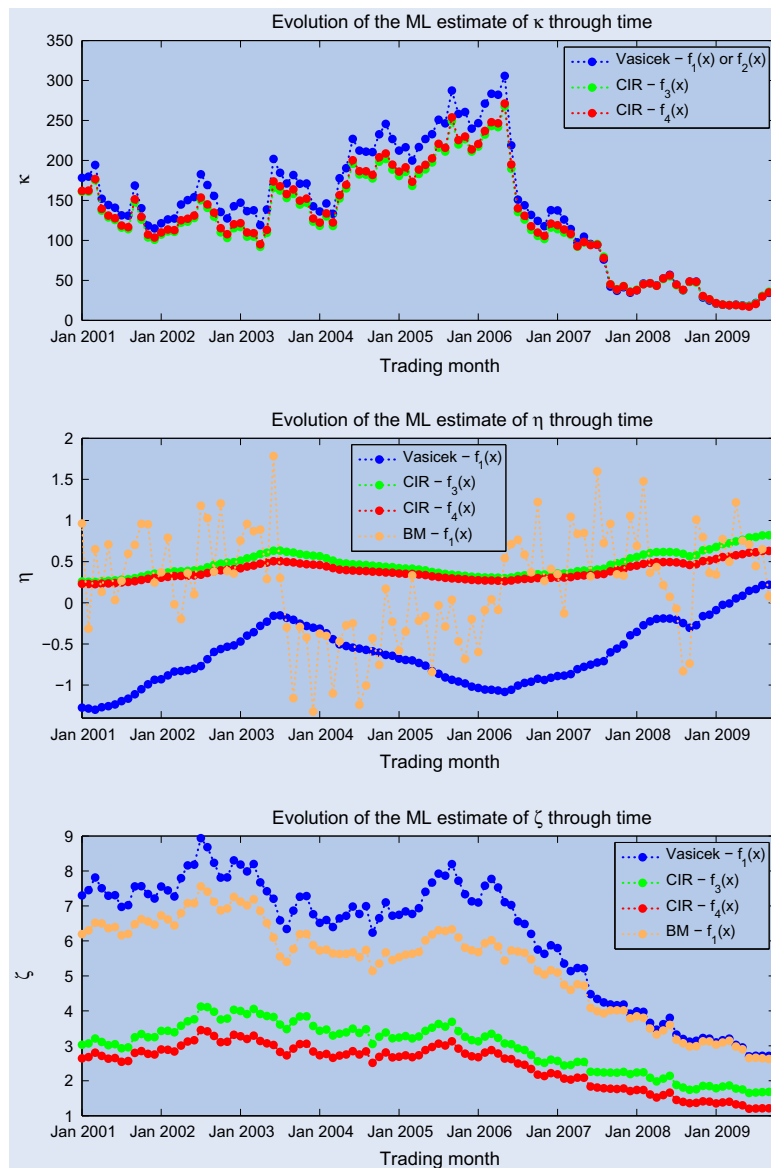


Figure 8. ML estimate of the model parameters.

For the numerical study, we consider  $N = 100\,000$  sample paths. Figures 7 and 8 show the likelihood function and the ML estimate of the model parameters, respectively, for

the 106 one-year overlapping time series windows included in the period under investigation, namely January 2000–December 2000, February 2000–January 2001, ..., October



Table 5. Average likelihood for different lengths of the time series window.

Length (in months)	Model			
	1	3	4	5
3	4.637	4.731	4.693	3.985
4	4.491	4.592	4.553	3.905
5	4.395	4.499	4.461	3.856
6	4.318	4.424	4.387	3.819
7	4.262	4.369	4.332	3.787
8	4.210	4.318	4.282	3.756
9	4.169	4.278	4.242	3.729
10	4.136	4.248	4.212	3.709
11	4.105	4.219	4.182	3.688
12	4.075	4.191	4.154	3.669
13	4.052	4.170	4.132	3.654
14	4.030	4.151	4.112	3.641
15	4.009	4.132	4.092	3.628
16	3.987	4.113	4.072	3.613
17	3.964	4.093	4.050	3.597
18	3.944	4.074	4.031	3.584
19	3.924	4.055	4.012	3.570
20	3.906	4.039	3.995	3.559
21	3.887	4.022	3.978	3.548
22	3.870	4.007	3.961	3.536
23	3.852	3.992	3.945	3.524
24	3.836	3.977	3.930	3.512

Table 6. Average monthly relative variation of the model parameters.

	Model			
	1	3	4	5
Mean( $ \Delta\kappa/\kappa $ )	0.1051	0.1022	0.1037	
Mean( $ \Delta\eta/\eta $ )	0.1728	0.0280	0.0243	1.7273
Mean( $ \Delta\zeta/\zeta $ )	0.0367	0.0324	0.0337	0.0311

2008–September 2009. Hence, the calibration and forecasting procedures are repeated every month. From figure 7, it is clear that models built upon an underlying mean-reverting process lead to a better fit of the historical HIX than models built upon a Brownian motion. Indeed, the likelihood  $L(\hat{\theta})$  is, on average, 12.32, 16.05 and 14.69 per cent higher under model 1, 3 and 4 than under model 5, respectively. Besides, the best fit is usually obtained for model 3, i.e. for the CIR process and the mapping function  $f_3(x) = 1 - \exp(-x)$ , for which the likelihood amounts to 4.1913, on average. For sake of robustness, we tried other time series windows, with a width ranging from 3 months to 2 years. Whatever the time series window, the model leading to the overall best fit is model 3, followed by model 4 and then model 1. Indeed, the average likelihood is always higher for model 3, as indicated in table 5. Note that, as expected, the wider the time series window, the lower the likelihood, on average. In the following, we will focus on a time series window of one year,<sup>†</sup> the conclusions drawn remaining to hold whatever the length of the time series. From figure 7, we also observe that, although the performance (measured as the likelihood function) of the proposed diffusion models is higher during the crisis period, the difference between models built

upon a mean-reverting process or a Brownian motion is then significantly less marked. Indeed, taking October 2007 as the first month of the credit crunch, the relative difference of the likelihood with respect to model 5 amounts to 14.95, 19.70 and 17.82 per cent and to 3.36, 3.57 and 3.99 per cent for models 1, 3 and 4 during the pre-crisis and the post-crisis period, respectively. Hence, the most suited model among models 1, 3 and 4 depends on the one-year sample period used to calibrate the model, the three models leading nevertheless to roughly the same likelihood.

A way to further discriminate between these three models is to look at the monthly variation of the calibrated parameters. Figure 8 indicates that the stability of the parameter  $\kappa$  is approximately the same for the three models, whereas the parameters  $\eta$  and  $\zeta$  are much more stable for the models built upon the CIR process.<sup>‡</sup> This pattern is confirmed by the average relative change of the model parameters reported in table 6. Indeed, among those three models, model 3 leads, on average, to the most stable parameters  $\kappa$  and  $\zeta$  whereas model 4 leads, on average, to the most stable long-run parameter  $\eta$ . Overall, the results clearly highlight the supremacy of the models built upon the underlying CIR process in terms of parameter stability.

<sup>†</sup>A time series of one year seems to be a reasonable choice given the fact that the predictions are computed for up to 1 month ahead.

<sup>‡</sup>Note that choosing a wider time series window leads to more stable calibrated parameters, which has to be counterbalanced with the lower likelihood.

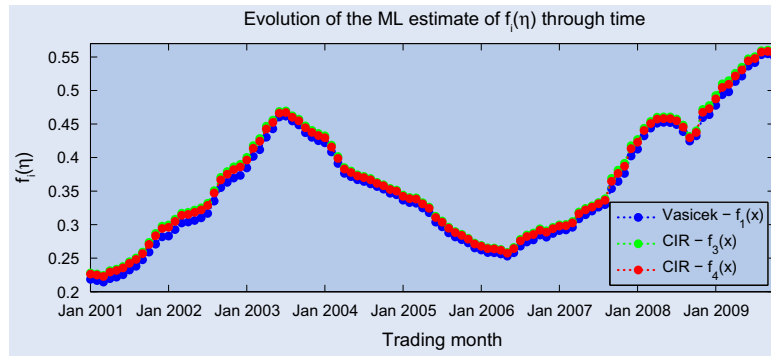


Figure 9. ML estimate of  $\tilde{\eta} = f_i(\eta)$ .

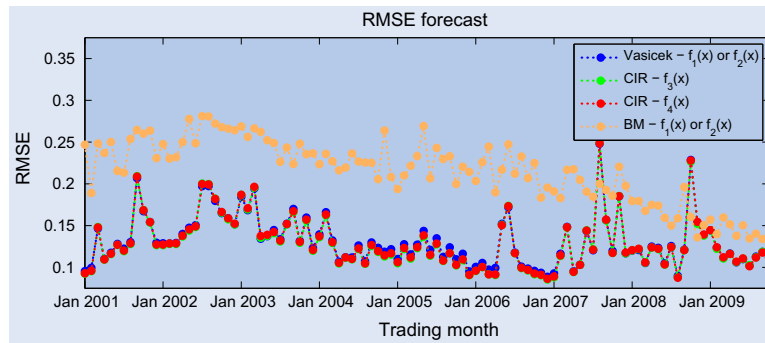


Figure 10. RMSE for the HIX forecast.

Figure 8 and table 6 also indicate the huge instability of the drift parameter  $\eta$  of the Brownian motion through time.

Observing the ML estimates  $\hat{\kappa}$ ,  $\hat{\eta}$  and  $\hat{\zeta}$  through time can be considered as a decomposition of the HIX in three components, each explaining a particular driver of the HIX process. For example, the evolution of the ML estimate  $\hat{\kappa}$  seems to indicate that there is a sudden change in the market herd behaviour around June 2006<sup>†</sup> and August 2007, characterized by a steep decrease in the mean-reverting parameter  $\kappa$ . This might suggest that recording the evolution of the model parameter  $\kappa$  could allow to detect some warning sign for market distress. Indeed, if the HIX is at a relatively high level, such a decrease in the mean-reverting rate might indicate that there is a comonotonicity cluster, i.e. an increased probability that the HIX remains for a while at a relatively high level. Another more obvious indicator of a change in the degree of market comonotonicity can be detected from the “long-run mean” of the HIX, i.e. from  $\tilde{\eta}$  (see figure 9). There are two clear changes in the trend of market herd behaviour during the period under consideration: the first one in July 2003 and the second one in June 2006. The transition between the long-term downturn and long-term upturn which occurred in June 2006 could have been, if monitored, a clear warning sign of the 2007–2008 credit crunch. Moreover, the calibrated parameters could give us further insight in the right time to enter a dispersion strategy than simply looking at the current value of the HIX. In particular, an investor should favour entering such position when both the current value of

the HIX and the calibrated mean-reversion rate are historically high rather than when a historically high value of the HIX is associated with a low value of  $\hat{\kappa}$ .

The supremacy of mean-reverting processes can also be inferred from the root-mean-square error (RMSE) of the forecasts:

$$FRMSE = \left( \frac{1}{m} \sum_{i=1}^m \frac{1}{N} \sum_{j=1}^N \left( Y_{t+i\Delta t} - Y_{t+i\Delta t}^{(j)} \right)^2 \right)^{1/2}, \quad (5.1)$$

where  $\{Y_{t+i\Delta t}, i = 1, \dots, m\}$  are the out-of-sample observations of the 13 month (i.e. the month following the one-year observation period). Figure 10 shows the forecast RMSE (5.1) for the period 2001–2009, for monthly recalibration whereas figure 11 shows the HIX forecasts for the different model settings. The results indicate that models built upon mean-reverting processes lead to a significantly lower RMSE forecast than models built upon Brownian motions. Indeed, the forecast RMSE is on average 37.36, 38.44 and 38.09 per cent lower under models 1, 3 and 4 than under model 5. Moreover, the difference is more marked before the crisis since the relative difference of the FRMSE with respect to model 5 amounts to 42.22, 43.51 and 43.12 per cent and to 20.76, 21.10 and 20.89 per cent, for models 1, 3 and 4 during the pre-crisis and the post-crisis period, respectively. These findings are in line with the conclusion drawn from the ML.

<sup>†</sup>June 2006 corresponds to the starting of the decline in the value of subprime mortgages caused by the steep drop in house prices.

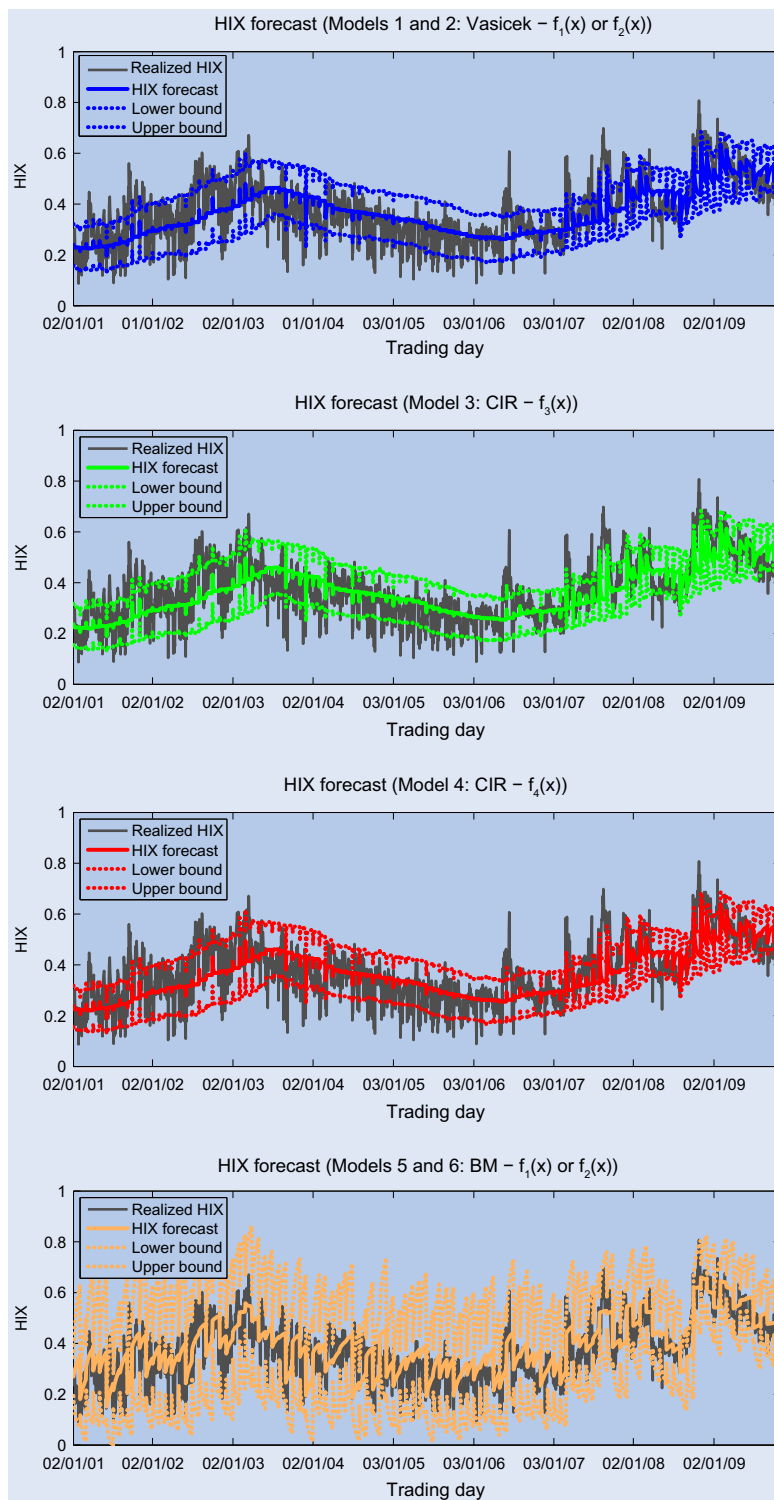


Figure 11. HIX forecast.

## 6. Conclusion

This paper proposes a time-dependent diffusion process for modelling herd behaviour indices which is built upon some popular mean-reverting process and a monotonic function which maps the definition domain of the underlying process to the unit interval. For the numerical study, we work out

different model settings, built upon either a Brownian motion, a Vasicek or a CIR process. The results have highlighted the fact that models built upon a mean-reverting process outperform models built upon a Wiener process, in terms of their ability to both fit historical HIX values and predict future values of the HIX. Moreover, the results have indicated that models built upon a CIR process lead to more stable calibrated parameters

through time. By way of conclusion, diffusion models obtained by mapping monotonically a CIR process to the unit interval are promising to predict future values of herd behaviour indices such as the HIX and hence to deliver crucial information on how to adjust the relative positions in individual stocks and their pooled index.

## References

- Aït-Sahalia, Y., Transition densities for interest rate and other diffusions. *J. Finance*, 1999, **54**, 1361–1395.
- Breedon, D. and Litzenberger, R., Prices of state contingent claims implicit in option prices. *J. Bus.*, 1978, **51**(6), 621–651.
- Carr, P. and Madan, D. Towards a theory of volatility trading. In *Option Pricing, Interest Rates and Risk Management*, edited by E. Jouini, J. Cvitanic and M. Musiela, pp. 458–476, 2001 (Cambridge University Press: Cambridge).
- Cheung, K.C., Dhaene, J., Kukush, A. and Linders, D., Ordered random vectors and equality in distribution. *Scand. Actuar. J.*, 2013. Available online at: <http://dx.doi.org/10.1080/03461238.2013.807470>.
- Cox, J.C., Ingersoll, J.E. and Ross, S.A., A theory of the term structure of interest rates. *Econometrica*, 1985, **53**(2), 385–407.
- Deelstra, G., Dhaene, J. and Vanmaele, M., An overview of comonotonicity and its applications in finance and insurance. In *Advanced Mathematical Methods for Finance*, edited by B. Oksendal and G. Nunno, pp. 155–179, 2011 (Springer: Berlin).
- Dhaene, J., Denuit, M., Goovaerts, M., Kaas, R. and Vyncke, D., The concept of comonotonicity in actuarial science and finance: Theory. *Insur. Math. Econ.*, 2002a, **31**(1), 3–33.
- Dhaene, J., Denuit, M., Goovaerts, M., Kaas, R. and Vyncke, D., The concept of comonotonicity in actuarial science and finance: Applications. *Insur. Math. Econ.*, 2002b, **31**(2), 133–161.
- Dhaene, J., Dony, J., Forays, M., Linders, D. and Schoutens, W., FIX—the fear index: Measuring market fear. In *Topics in Numerical Methods for Finance*, edited by M. Cummins, F. Murphy, and J. Miller, 2012 (Springer Proceedings in Mathematics & Statistics: New York, NY).
- Dhaene, J., Linders, D., Schoutens, W. and Vyncke, D., The herd behavior index: A new measure for the implied degree of co-movement in stock markets. *Insur. Math. Econ.*, 2012, **50**(3), 357–370.
- Driessen, J., Maenhout, P.J. and Vilkov, G., The price of correlation risk: Evidence from equity options. *J. Finance*, 2009, **64**(3), 1377–1406.
- van Emmerich, C., Modelling correlation as a stochastic process. Technical report, Bergische Universität, Wuppertal, June 2006.
- Kaas, R., Dhaene, J. and Goovaerts, M.J., Upper and lower bounds for sums of random variables. *Insur. Math. Econ.*, 2000, **27**(2), 151–168.
- Laurence, P., A new tool for correlation risk management: The market implied comonotonicity gap. Global Derivatives, Paris, Invited Talk, May 2008.
- Linders, D., Dhaene, J., Hounnon, H. and Vanmaele, M., Index options: A model-free approach. Research report AFI-1265 FEB, Leuven: KU Leuven - Faculty of Business and Economics, 2012.
- Linders, D. and Schoutens, W., A framework for robust measurement of implied correlation. Research report AFI-1386 FEB, Leuven: KU Leuven - Faculty of Business and Economics, 2014.
- Rubbaniy, G., Asmerom, R. and Rizvi, S.K.A., Do fear indices help predict stock returns? 2013. Available online at SSRN: <http://ssrn.com/abstract=2121274>.
- Skintzi, V.D. and Refenes, A.N., Implied correlation index: A new measure of diversification. *J. Futures Markets*, 2005, **25**, 171–197.
- Teng, L., van Emmerich, C., Ehrhardt, M. and Günther, M., A versatile approach for stochastic correlation using hyperbolic functions. *Int. J. Comput. Math.*, 2015. Available online at: <http://dx.doi.org/10.1080/00207160.2014.1002779>.