DEPARTMENT OF ECONOMICS
WORKING PAPERS

economics.ceu.hu

Counting Process Generated by Boundary-crossing Events Theory and Statistical Applications

by

Péter Farkas[1]

2013/5

[1] Department of Economics, Central European University, Farkas_Peter@ceu-budapest.edu

Abstract

In this paper, we introduce and analyze a new class of stochastic process, named Boundary Crossing Counting Process. We show how to assess its finite-sample probability distribution using first exit time distributions. We discuss and calibrate three different methods for estimating this probability distribution and show how to use this distribution for non-parametric statistical specification testing or model validation. We apply this tool to the problem of unit root testing and to the problem of modeling financial data. First, we find that the BCC-test is less powerful than specific parametric unit-root tests, yet more powerful than another non-parametric test based on probability integral transform. Next, we learn that the BCC-test is relatively powerful in differentiating between commonly used GARCH type of models when the model under the null hypothesis and the model under the alternative hypothesis differ in the distribution of the error term. Finally, we show that the BCC-test often rejects the null hypothesis that the data generating process for major American and Australian stock indexes is well represented by the GARCH(1,1) model. The BCC-test cannot reject the GJR model for any specification, which is additional supportive evidence for the presence and importance of asymmetry in financial markets.

Key words: Stochastic processes, First-exit times to a double boundary, Time series econometrics, Financial econometrics, Nonparametric testing, Specification testing, Model validation, BCC-test, Statistical Power, Sensitivity, Type-2 errors

Acknowledgements

I would like to express my appreciation for the support of Central European University and especially for the personal support of László Mátyás, without whom this paper could not have been written.

Introduction

This paper introduces and analyzes a new class of stochastic processes, which we called Boundary-Crossing Counting Processes or BCC-processes. The name originates from the fact that a BCC-process essentially counts the number of boundary-crossing events. This stochastic process is characterized by an underlying series of random variables X_t^* with memory. Generally speaking, very little needs to be assumed about X_t^*: it can be non-stationary, discrete or continuous, it can be defined parametrically as a stochastic process, or non-parametrically by some data. We begin by defining boundary-crossing events for a general random variable Z_t :

- We first define upper boundary crossing event as the first event when $Z_t \notin (a_t, b_t), Z_t > b_t$ is true.
- We next define lower boundary crossing event as the first event when $Z_t \notin (a_t, b_t), Z_t < a_t$ is true.
- We finally define upper or lower boundary crossing event as the first event when $Z_t \notin (a_t, b_t)$ is true.

We introduce a new random variable $X_t(X_t^*, a_t, b_t)$ derived from the original random variable X_t^* by restarting it at some initial value X_0 upon each boundary-crossing event. Note, that the initial value X_0 may be a fixed value or a random variable, as long as the initial condition $a_0 < X_0 < b_0$ holds. The random variable X_t is therefore enclosed by an a_t lower boundary and a b_t upper boundary with initial condition $a_0 < X_0 < b_0$.

Using $X_t(X_t^*, a_t, b_t)$, we introduce a new class of discrete stochastic processes (Y_t), which counts the number of boundary crossing events. More specifically, let Y_t^U be used for counting upper, while let Y_t^L be used for counting lower boundary crossing events and finally, let Y_t^{UL} be used to count upper and lower crossing events both. Formally, upper-boundary crossing Y_t^U is defined as:

$$Y_t^U(i+1) = Y_t^U(i) + 1, \qquad if\ X_{t-\varepsilon} \in (a_t, b_t), X_t \notin (a_t, b_t), X_{t+\varepsilon} = X_0 \tag{1}$$

where ε indicates an arbitrarily small number. Y_t^L and Y_t^{UL} are defined likewise. Moreover, let us use Y_t to refer to these three processes at once, without any specific upper index. Figure 1 below illustrates one possible BCC-process: The circles indicate upper-boundary crossing events, while the triangles indicate lower-boundary crossing ones. At each boundary crossing event, Y_t^{UL} is incremented by one.

In theory, there may be many other useful BCC-processes as well, but in this paper we focus on Y_t^U, Y_t^L and Y_t^{UL} exclusively. For example, another BCC-process may be defined as a process that is incremented by one upon upper boundary-crossing and decreased by one upon lower boundary crossing, but in this paper we will not deal with them.

The aim of the paper is to first analyze and then to apply the discrete probability distribution $p_i(T)$, which describes the probability that $Y_T = i$. Similarly to the notations introduced above, $p_i^U(T)$ refers to

upper boundary-crossing probabilities, $p_i^L(T)$ refers to lower boundary crossing probabilities, and $p_i^{UL}(T)$ refers to boundary-crossing events in general.

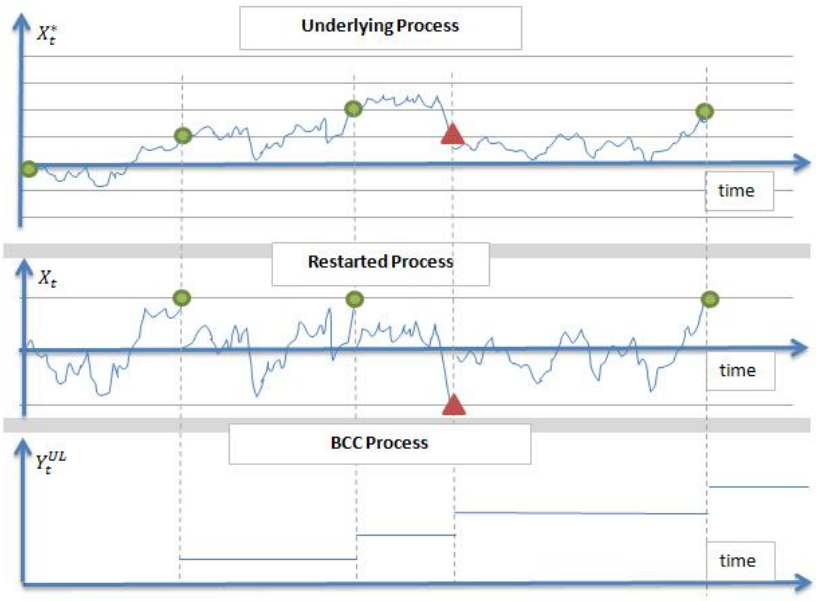

Figure 1: What is a Boundary-Crossing Counting Process or BCC-process?

The study of the above BCC-processes is motivated by its useful applications. These stochastic processes appear to have diverse applications in the field of statistics, finance and management sciences, but in this paper we focus only on their statistical applications. The theory applied in this paper builds heavily on the mathematical results related with boundary-crossing events. There is no clear consensus on how to name the various events, therefore let us briefly review the terminology. 'First passage time' or 'hitting time' is typically used in situations, where there is only one boundary. 'Expected first passage time' describes the expected amount of time needed to reach that boundary. 'First passage time distribution' aims to characterize the full distribution. The case of two boundaries is usually referred to as 'first exit time' or 'double-barrier hitting time' although the notation "first exit time" is also used to describe first passage time, see for example Wilmot (1998), p. 144 . 'Exit times' should not be confused with 'first range time', as range is generally used to describe the difference between the maximum and the minimum value. In this paper, we follow the terminology of Borodin and Salminen (2002), they use the name "first exit time" to describe the case of double boundaries, therefore, we also use the same term.

The first wave of literature on this topic was written by classical authors, like Bachelier (1900) or Kolmogorov (1931) and was motivated by the gambler's ruin problem in finance and by repeated (independent) sampling problem in statistics. The second wave of literature aimed to formalize these early results, along with some corresponding results in physics, and is summarized in Feller's classical book (Feller, 1968 and 1971). Another general treatment is given by Karlin and Taylor (1981, p. 192), who characterize boundary-crossing probabilities and expected first passage-times using certain

functionals and the concept of scale function and speed function under fairly general assumptions. A less general approach focusing on the Geometric Brownian Motion may be found in Karlin and Taylor, (1998). A third wave of literature consists of articles partially motivated by pricing certain financial options (barrier-options). Lin (1998), for example, proposes to use Gerber-Shiu technique along with the Laplace Transform in order to calculate the first exit time (double-barrier hitting time) distributions. Linetsky (2004) proposes to use spectral expansion approach for calculating hitting time distributions.

The results for common continuous stochastic processes have been summarized in a handbook format by Borodin and Salminen (2002). Here, authors also discuss the theory used in deriving these results, although actual proofs are generally not included. An even more general approach, connecting the theory of boundary-crossing events with the theory of integral equations, has been proposed by Valov (2009). In general, the theory of exit times and hitting times are much more developed for continuous processes than for discrete ones, that is why applying simulations when dealing with discrete data is fairly common. Valenti, Spagnolo and Bonanno (2006), for example, compared empirical and simulated hitting time distributions: More precisely they have compared the hitting time distributions derived from the Brownian motion, from the GARCH model and from the Heston model, with the hitting time distribution obtained from actual data on security prices. If one decides to carry out simulations of this kind without lack of theoretical/analytical benchmark, then the insights of Primožič (2011) concerning Monte Carlo simulation shall be taken into account.

The main innovations of our paper are two-fold: on the one hand, section 2 extends existing results described above by introducing and analyzing these new Boundary Crossing Counting stochastic processes. To our knowledge BCC-processes have not been dealt with or characterized in this manner. On the other hand, we show how to use BCC processes for non-parametric statistical specification testing, which we call Boundary-Crossing Counting test or BCC-test. This name is appropriate because BCC-test compares the number of boundary crossing events under some null hypothesis with the number of boundary crossing events in the actual sample observation (statistical testing), or with the number of crossing events in the alternative specifications.

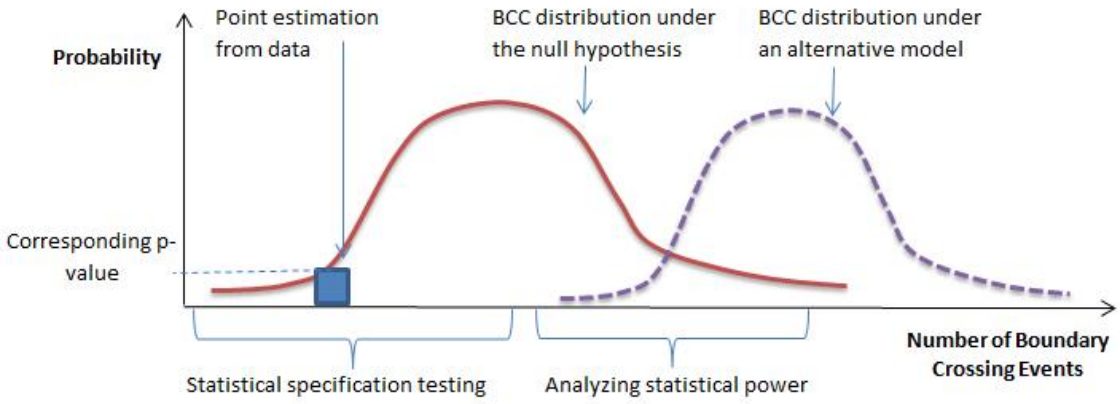

Figure 2: How to use the Boundary-Crossing Counting process for statistical specification-testing?

By using the BCC-test, one can, for instance, evaluate the probability that an observed realization of a potentially unknown stochastic process is generated by some pre-defined stochastic process. A financial analyst may, for example, calculate the probability that some security prices (a stock index, the price of an individual stock or the price of a commodity) is well represented by the Gaussian Random Walk process or by some GARCH(p,q) process. The logic of the test illustrated above is as follows. Under the null hypothesis, it is possible to estimate the BCC-process $Y_t(\omega)$, which can then be evaluated at sample lengths T. The $Y_t(\omega)$ distribution (plotted above) can then be compared with the observed number of boundary crossing, and the corresponding p* value van be calculated. It is important to highlight that the H_0 hypothesis may not need to be a parametric one. The BCC-test can also be useful for evaluating whether two sets of dependent observations are generated by the same, potentially unknown, data generating process. For instance, one can evaluate the probability whether two (appropriately scaled) security prices (for example, the price of an airline and a price of an automobile producer) can be represented by the same stochastic process. Finally, the BCC-distribution can also be used in studying statistical power. For example, one can evaluate how different the GARCH(1,1) process and the Geometric Brownian Motion process are in finite samples.

Obviously, it is important to represent the pricing process well; so many tests have been devised for the purpose of specification testing relying on different assumptions. Starting from Lo and MacKinlay's (1988) variance ratio test, authors have described many important properties of empirical asset returns. These properties are often documented as stylized facts in comparison to the Gaussian Random Walk (GRW) hypothesis. Accordingly, we also include the GRW model in addition to the GARCH type of models. Unlike the "usual" tests mentioned earlier, the BCC-test is a fully non-parametric test. Therefore, it does not aim to test the variance as variance-ratio tests nor the conditional mean as done in the large number of market-efficiency studies as summarized by Fama (1991), nor a non-parametric test of the conditional mean as in Hong and Lee (2005), nor aim to test any higher moments as in Neuberger (2012). This test does not aim to test any specific property of the assumed data-generating process but tests the full specification altogether. For example, in the case of the Gaussian Random Walk, the null hypothesis encompasses both the mean and the variance. In the case of a GARCH processes, the null hypothesis encompasses the mean, the unconditional variance, the conditional variance and all other parameters describing the structure of the conditional variances. In fact, the data-generating process does not have to have any parameter at all; the BCC-test can be used to test fully non-parametric specifications as well.

Our test is basically a non-parametric specification test, or a method for model validation. One of the first articles in this relatively new field has been written by Ait-Sahalia (1996), who compared the marginal density estimator with the non-parametric sample-based alternative. His method has certain drawbacks in finite samples, as shown by Pritsker (1998), therefore, researchers have developed several alternative tests. The common method for these non-parametric tests is to devise and compare some measure for the parametric and the non-parametric specifications as well. Naturally, rejection occurs when the discrepancy between the two measures are substantial. Papers basically differ in how to

measure the discrepancy. Transition density based comparisons have been used, for example, by Gao and King (2004), and Hong and Li (2005). Alternatively, comparison based on the likelihood function has been proposed by Fan, Zhang and Zhang (2001), who developed a generalized likelihood ratio test for this purpose. Anderson (1993) suggested a comparison based on the spectral densities. Recently, Song (2012) recommended a comparison based on the infinitesimal operator. These methods have been reviewed, for example, by Fan (2005) or Zhao (2008). Our method differs from these non-parametric tests because we propose to measure the discrepancy between two specifications by comparing the number of boundary-crossing events.

The BCC-test has several desirable properties. Besides the usual favorable properties of non-parametric tests, our method is a non-asymptotic one, and therefore, it does not suffer from size-distortion. Also, this test is relatively straight-forward to implement. The BCC-test is more powerful than an alternative non-parametric test based on the probability integral transform, inspired by Hong and Li (2005) and implemented by Ghalanos (2013) (further referred as PIT-test), when testing the Unit-root hypothesis against a close-unit-root alternative. It is also much more powerful when comparing the GARCH(1,1) hypothesis with Gaussian error terms against other GARCH specification using the Student-t distribution. All these properties make the BCC test a convenient general diagnostic tool. Naturally, the BCC-test suffers from some minor drawbacks. As our method is a non-parametric one, it requires a relatively large sample-size. The BCC distribution is a discrete one, consequently, selecting the usual 1%, 5% and 10% as critical values is somewhat problematic, and we have to make use of the closest available discrete value. Also, for long time series of over 5000 observations, calculations may be computationally intensive.

The paper is organized as follows. The second section introduces and characterizes BCC-processes and connects them to the existing literature on stochastic processes. The third section begins by discussing three different estimation methods and continues by dealing with the most common problems of non-parametric specification testing. The fourth section first applies the BCC test to the classical problem of unit root testing by comparing its statistical power to several parametric and a non-parametric alternatives. We continue this section by studying the statistical power of the BCC-test when the null hypothesis is the GARCH(1,1) process and the alternative models are other commonly used pricing models. We finish this part by applying the BCC-test for specification testing of pricing models on some important American and Australian stock indexes. The last section concludes.

2. Theory of BCC-processes

This section is a non-technical discussion on the theory of BCC-processes. In particular, we do not prove the existence of the concepts we introduce, but we provide references for those readers who are interested in existential proofs. The main innovation here is a recursive algorithm, which allows us to calculate the BCC-distribution from the first exit time distribution. The theory of first exit times is more developed for continuous-time processes; therefore we begin the discussion with the continuous case, and then continue with the discrete one. We conclude by showing how these two cases are connected.

Definition 1 (boundaries): Let a_t and b_t be two measurable functions, $a_0 < X_0 < b_0$.

The literature differentiates between three kinds of boundaries.[1] The simplest case is the one of constant boundaries, where $a_t = a$ and $b_t = b$. Also, in certain applications, it may be useful to choose $a_t = f(t)$ and $b_t = g(t)$, where the boundaries are continuous functions of time. Finally, boundaries may be stochastic as well: $a_{t+1} = f(X_t|F_{t-1})$ and $b_{t+1} = g(X_t|F_{t-1})$, that is to say boundaries at time $t + 1$ depend on the value of the process at time t, which is unknown at time $t - 1$. Depending on the actual applications, boundaries may be chosen exogenously or endogenously, as a solution of the appropriate optimal (stochastic) control problem. For example, in statistical applications one may ask which boundary functions would maximize the statistical power of the Boundary Crossing Counting test. In this paper, we are not going to solve such problems under general settings because doing so would require a lengthy technical discussion and at this stage we do not see any significant additional benefits of going down this path. Interested readers may find the description of these methods, for example, in the book of Kirk, (2004). Instead, we are going to work with several exogenously given constant boundaries and carry out the calculations for several of these boundary pairs: $(a, b) = \{(a_1, b_1), (a_2, b_2), \dots (a_k, b_k)\}$. Thus we will be able to evaluate how much the results are driven by the choice of the exogenously given boundaries.

Definition 2 (first exit-time for continuous distributions): Let T^{UL} indicate the time in which process X_t first crosses either boundaries:

$$T^{UL} = \begin{cases} inf(t: X_t \notin (a, b)), if\ t\ is\ finite \\ \infty\ otherwise. \end{cases} \quad (2)$$

Moreover, let T^U indicate the time in which X_t first crosses the upper boundary without crossing the lower boundary before. Also, let T^L indicate the time in which X_t first crosses the lower boundary without crossing the upper boundary before. Note that the literature typically uses the T^{ab} notation, while we applied the T^{UL} convention in order to be consistent with the rest of the paper.

Throughout the paper, we are going to use upper case T to refer to a specific time period. In the continuous case, time is indicated by lower case t spanning from zero to T, that is $t \in (0, T]$. In the discrete case, lower case t indicates time periods ranging from zero to T, that is $t = (0, t_1, t_2, \dots T]_{1xn}$. We use n to indicate the number of observations. Finally, we are going to use i to indicate the number of boundary-crossing events. The notation t_i indicates the time or the time-period at which exactly i boundary-crossing event has occurred. The notation t_0 indicates the last time the process has been restarted. If the process has not been restarted, then $t_0 = 0$. Therefore, the notation $t \in (t_0, T^{UL})$ can be used to indicate the time passed since the process has been restarted for the last time.

Boundary classification may be found in Karlin and Taylor, (1981, p. 234, Table 62), where they differentiate between "regular", "absorbing", "natural" and "entrance" types. The type of boundary applied in our paper does not have a one to one correspondence to any of the classical case: they could be called 'restarting boundaries'. If one must classify, restarting boundaries are attainable and regular boundaries, where the process is restarted upon boundary-crossing events.

Assumption 1: Let T^{UL} be positive and finite.

In the standard literature on stochastic processes Assumption 1 is frequently a theorem derived from more elementary assumptions. The finiteness of first-exit time is a well-known property for martingales, as explained, for example, by Medvegyev, (2007). The typical proof for non-martingales is to convert the process to a martingale, as shown, for example, by Karlin and Taylor (1998). The non-zero property of the first exit time is only problematic if the limits of the boundaries are equal to the initial value of the process, a case which is not dealt with here. A similar problem for hitting times is discussed by Valov, (2009). As Assumption 1 is not elementary, therefore, it is important to highlight that it imposes restrictions on X_t, a_t and on b_t. Not all stochastic processes and corresponding boundaries will satisfy this assumption.

Definition 3 (First Exit Time distributions for Continuous processes): Let $fetc(a_t, b_t, C, t, \omega)$ (short for First Exit Time distribution, Continuous) be defined as a continuous probability distribution, describing the probability that the first exit time is t. In this definition, a_t and b_t represent the boundaries, C is used to indicate constant parameters such as the value of drift or volatility, t indicates time-dependence, while ω indicates stochastic dependence. Let us use superscripts for representing the type of distribution:

- $fetc^{U}(a_t, b_t, C, t, \omega)$ describes the probability that the first exit time is t and the boundary crossing event is an upper crossing one,
- $fetc^{L}(a_t, b_t, C, t, \omega)$ describes the probability that the time first exit time is t and the boundary crossing event is a lower crossing one and finally,
- $fetc^{UL}(a_t, b_t, C, t, \omega)$ describes the probability that the first exit time is t and the boundary crossing event is either and upper- or a lower crossing one.

Moreover, let us use $fetc(a_t, b_t, C, t, \omega)$ to refer to these three distributions at once, without any specific upper index. Formally, let $fetc^{U}(a_t, b_t, C, t, \omega)$ be defined as:

$$fetc^{U}(a_t, b_t, C, t, \omega) = p(T^{UL} = t = T^{U}) \tag{3}$$

Naturally, $fetc^{L}(a_t, b_t, C, t, \omega)$ and $fetc^{UL}(a_t, b_t, C, t, \omega)$ is defined likewise. Let us use upper case letters for the cumulative distribution function. Note, that the cumulative distribution functions $FETC^{U}(a_t, b_t, C, t, \omega)$ and $FETC^{L}(a_t, b_t, C, t, \omega)$ are both `defective', which is to say that their value does not necessarily converge to one as t gets arbitrarily large but may converge to a value smaller than one. $FETC^{UL}(a_t, b_t, C, t, \omega)$, on the other hand, is a proper cumulative distribution function, which is a direct consequence of Assumption 1.

$$\lim_{z\to\infty} \int_0^z fetc^{U}(a_t, b_t, C, t, \omega) + \int_0^z fetc^{L}(a_t, b_t, C, t, \omega) = 1 \tag{4}$$

Assumption 2: Let the functional form of the first exit time distributions be time-independent. Also, regarding the stochastic structure of the first exit time distribution, let us only allow the function to depend on the initial value of the process.

Assumption 2 requires us to condition the distributions only on information available a-priori at time zero. Perhaps a more intuitive explanation is to say that as a result of Assumption 2, we are not allowed to update the distribution function at time t based on the information revealed in period 1.... $t-1$. That is to say, $fetc(a_t, b_t, C, t, \omega)$ cannot, for example, be the function of X_{t-1}. Also, as a consequence of this assumption, the functional form of the distribution function does not depend on how many boundary-crossing events we have observed. In other words, $fetc(a_t, b_t, C, t_i, \omega) = fetc(a_t, b_t, C, t_{i'}, \omega)$, where t_i indicates the time at which exactly i boundary crossing events have occurred, while $t_{i'}$ indicates the time at which exactly i' boundary crossing events have occurred. From what has been said above, we have restricted the functional form of the first exit time distribution to $fetc(a_t, b_t, C, t, X_0)$. From now on, we suppress the dependence on the boundaries, on constants and on the initial value for ease of notation.

Definition 4 (Restarting the process in continuous time): Let X_0 be a fixed or a stochastic initial value. Each time the stochastic process crosses the boundary, its value is "almost immediately" reset to this initial value X_0. Formally,

$$X_{T^{UL}+\varepsilon} = X_{T^{U}+\varepsilon} = X_{T^{L}+\varepsilon} = X_0 \tag{5}$$

where ε is arbitrarily small. The notation 'almost immediately' assumes that the process is reset to its initial value exogenously. Therefore, we know that at time T^{UL} the process is at $X_t \notin (a, b)$ and we know that at time $, T^{UL} + \varepsilon$ the process is at X_0, yet we do not deal with what happens in the time-period $(T^{UL}, T^{UL} + \varepsilon)$.

We postpone the somewhat technical discussion on how to calculate the first exit time distribution for a while and assume that we have somehow obtained the first exit time distributions and proceed by showing how first exit time distributions characterizes the BCC processes. The probability that no boundary crossing event occurred until time T is simply the probability that the first boundary crossing event occurred at a later time:

$$p_0^{UL}(T) = \int_T^z fetc^{UL}(t)dt = 1 - FETC^{UL}(T) \tag{6}$$

where z can be arbitrarily large. It can be shown by induction that the probability of exactly $k \in N > 0$ boundary crossing events occurring in time T is calculated as:

$$p_k^{UL}(T) = \int_0^T fetc^{UL}(t) * p_{k-1}^{UL}(T-t)dt \tag{7}$$

Now $p_0^{UL}(T)$ can be calculated directly from the first exit time distributions, $p_1^{UL}(T)$ can be calculated from $p_0^{UL}(T)$ and so on. Therefore, by applying Equation (7) recursively, we can characterize the BCC-distribution completely using the first exit time distributions. Calculating first exit time distributions and iteratively applying Equations (7) analytically is quite challenging even for simple stochastic processes.

This is, however, not a serious problem from a practical perspective as in practice, data is discretely sampled. Let us therefore continue with the discrete-time formulation of the problem.

Definition 5 (Exit-time for discrete distributions): Let T^{UL} indicate the time in which X_t first crosses either boundaries:

$$T^{UL} = \begin{cases} inf(t: a_k < X_k < b_k \; and \; X_t < a_t \; or \; X_t > b_t), t_0 \leq k < t, t \; is \; finite \\ \infty \; otherwise. \end{cases} \quad (8)$$

Moreover, let T^U indicate the time in which X_t first crosses the upper boundary without crossing the lower boundary before. Also, let T^L indicate the time in which X_t first crosses the lower boundary without crossing the upper boundary before. Note that Assumption 1 also applies to the discrete case as well, therefore, exit-time is assumed to be positive and finite.

Definition 6 (First exit time distributions for discrete processes): Let $fetd(a_t, b_t, C, t, X_0)$ (short for First Exit Time distribution, Discrete) be defined as a discrete probability distribution, describing the probability that the boundary-crossing event occurs exactly in the t^{th} interval. Similarly to the continuous case,

- $fetd^U(a_t, b_t, C, t, X_0)$ describes the probability that there is an upper boundary-crossing event exactly in the t^{th} interval,
- $fetd^L(a_t, b_t, C, t, X_0)$ describes the probability that there is a lower boundary crossing event exactly in the t^{th} interval and finally,
- $fetd^{UL}(a_t, b_t, C, t, X_0)$ describes the probability that there is an either upper or lower boundary crossing event exactly in the t^{th} interval.

Moreover, let us use $fetd(a_t, b_t, C, t, X_0)$ to refer these three distributions at once, without any specific upper index. Note again, that Assumption 2 also applies to the discrete case. Let us suppress the dependence on the boundaries, on constants and on the initial value for ease of notation. Formally $fetd^U(t)$ is defined as:

$$fetd^U(t) = p(a_k < X_k < b_k \; and \; X_t > b_t), t_0 \leq k < t, t \; is \; finite \quad (9)$$

Naturally, $fetd^L(t)$ and $fetd^{UL}(t)$ is defined likewise.

Restarting the process in discrete time is somewhat less straight-forward than restarting in continuous time. The first issue is related with where to restart the process. There are at least two possibilities: on the one hand, it is possible to restart the process at some fixed or random X_0:

$$X_{T^{UL}+1} = X_{T^U+1} = X_{T^L+1} = X_0 \quad (10)$$

However, it may also be reasonable to restart the process at $X_0 + (X_{T^U} - b_{T^U})$ for upper boundary-crossing, and at $X_0 + (X_{T^L} - a_{T^L})$ for lower boundary-crossing:

$$X_{T^U+1} = X_0 + \left(X_{T^U} - b_{T^U}\right) \; and \; X_{T^L+1} = X_0 + (X_{T^L} - a_{T^L}) \quad (11)$$

It is in fact the role of the specific application to decide which approach is more relevant. For statistical application, it is at the discretion of the researcher to decide where to restart and for simplicity, we always restart the process at zero. The second issue is related with the magnitude of $(X_{T^U} - b_{T^U})$ and $(X_{T^L} - a_{T^L})$ as it is possible that the stochastic process not only crosses the boundaries a_t or b_t, but it also crosses $(j * a_t)$ or $(j * b_t)$, where $j > 1$ and is a natural number. Again, several sensible solutions may arise. Probably the simplest one is to select the boundaries so that such an event occurs sufficiently rarely, so that it does not influence the results. Also, it may be possible to modify the updating rule and increment or decrease the counting process not by always 1 but by j. These solutions may all be reasonable and therefore it is the role of the specific application to decide which approach is more relevant. In order to simplify the exposition, we are going to assume that we are able to select the boundaries in such a way that $j = 1$.

Assumption 3: Let us assume that the probability that boundaries were crossed twice within one time interval is negligibly small:

$$p(X_k < b_k \; and \; X_t > 2b_t) = p(X_k > a_k \; and \; X_t < 2a_t) = 0, t_0 \leq k < t, t \; is \; finite \tag{12}$$

As a direct consequence of Assumption 3, if we have n observations, the maximum number of boundary crossing event is n. In this case we can characterize the boundary crossing counting distribution with the help of the following P matrix:

$$P^{UL}_{nx(n+1)} = \begin{bmatrix} p_0^{UL}(1) & p_1^{UL}(1) & . & . & . & p_n^{UL}(1) \\ p_0^{UL}(2) & p_1^{UL}(2) & . & . & . & p_n^{UL}(2) \\ . & . & . & p_i^{UL}(t) & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ p_0^{UL}(n) & p_1^{UL}(n) & . & . & . & p_n^{UL}(n) \end{bmatrix} \tag{13}$$

where again $p_i^{UL}(t)$ describes the probability that until period t, exactly i boundary-crossing events have occurred. The first exit time distribution in some period t ($0 < t \leq n$) is given by the t^{th} row of the P matrix. The columns of the P matrix are also meaningful: they describe the probability that exactly 1,2,...t period is needed for i boundary-crossing events to occur. The first column of the matrix can be calculated using the first exit time distribution just as in the continuous case:

$$p_0^{UL}(t) = \sum_{k=t+1}^{z} fetd^{UL}(k) = 1 - \sum_{k=1}^{t} fetd^{UL}(k) = 1 - FETD^{UL}(t) \tag{14}$$

where z can be arbitrarily large. Any other column of the P^{UL} matrix may be calculated recursively using:

$$P^{UL}(j-1)_{nx1} = F_2(j)_{nxn} * F_{1_{nx1}} \tag{15}$$

where F_1 is a static matrix and is composed of the first exit time distributions.

$$F_{1_{nx1}} = \begin{bmatrix} fetd^{UL}(1) \\ fetd^{UL}(2) \\ fetd^{UL}(3) \\ . \\ . \\ . \\ fetd^{UL}(n-1) \\ fetd^{UL}(n) \end{bmatrix} \tag{16}$$

Now $F_2(j)$ matrix can be expressed recursively, using probabilities obtained in the previous steps, and is calculated as:

$$F_2(j)_{nxn} = \begin{bmatrix} p_{j-1}^{UL}(0) & p_{j-1}^{UL}(-1) & . & . & . & 0 \\ p_{j-1}^{UL}(1) & p_{j-1}^{UL}(0) & . & . & . & 0 \\ p_{j-1}^{UL}(2) & p_{j-1}^{UL}(1) & . & . & . & 0 \\ . & . & . & p_{j-1}^{UL}(n-(n-r+c)) & . & . \\ . & . & . & . & . & . \\ p_{j-1}^{UL}(n-2) & p_{j-1}^{UL}(n-3) & . & . & . & 0 \\ p_{j-1}^{UL}(n-1) & p_{j-1}^{UL}(n-2) & . & . & . & p_{j-1}^{UL}(0) \end{bmatrix} \tag{17}$$

A general term of $F_2(j)$ matrix is given by $p_{j-1}^{UL}(n-(n-r+c))$, where n is the number of observations, r is the row number, c is the column number and the following conventions are respected:

- if $j-1 > n-(n-r+c)$ then $p_{j-1}^{UL}\big(n-(n-r+c)\big) = 0$
- if $j-1 = n-(n-r+c) = 0$ then $p_{j-1}^{UL}\big(n-(n-r+c)\big) = 1$

Formula (15) can be verified by induction. Based on Equations (14)-(17), the BCC-distribution can fully be characterized by solving for $fetd^{UL}(t)$. The advantage of this matrix-formulation over the brute-force combinatorial calculation is the reduction in the calculation-complexity: while the brute-force combinatorial method would require n^3 steps, the matrix-formulation described above reduces the required number of steps to n^2, and thus the algorithm becomes, although slow, yet feasible. For example in the case of 20 years of daily observed security data (approximately five thousand observations), the runtime[2] was approximately 60 minutes using the matrix formulation above, while interpolation suggests that the brute-force combinatorial method would require 200 days.

The BCC^L and the BCC^U distribution can be characterized using BCC^{UL}. Let us show this for the case of upper-crossing distribution, as the case of lower-crossing can be solved in a similar manner. Let us introduce the following concept:

Definition 7 (upper-boundary crossing probability): Let the probability that the stochastic process reaches the upper boundary before hitting the lower boundary for all $t < T^{UL}$ be defined:

[2] The calculations were done on a personal computer with a dual-core processor and 3 GB RAM.

$$P\left(X_{T^{UL}} = b_t \middle| b_k > X_k > a_k\right) = p(X_0), t_0 \leq k < t, t\ is\ finite \tag{18}$$

Let us suppress the dependence on the initial value of the process for simplicity. Also let us introduce the notation q for lower-boundary crossing probability, that is $q = 1 - p$. Similarly to the case of first exit time distribution, we will need to assume that the upper-boundary crossing probability does not depend on how many times the process has been restarted.

Assumption 4: Let the functional form of the upper-boundary crossing probability be time-independent. Also, as for the stochastic structure of the upper-boundary crossing probability, let us only allow the function to depend on the initial value of the process.

Let us assume[3] that that we have somehow obtained the upper boundary crossing probability, and proceed by showing how upper boundary crossing probability characterizes BCC^U. For some period t, the upper boundary crossing can be calculated as:

$$P^U(t)_{1x(n+1)} = P^{UL}(t)_{1x(n+1)}\left[Q_{(n+1)x(n+1)} .* PA_L(n+1)\right] \tag{19}$$

Where $.*$ indicated element by element multiplication, $PA_L(n+1)$ is the absolute value of the lower triangular Pascal matrix with $n+1$ rows and columns and finally, $Q_{(n+1)x(n+1)}$ matrix is constructed as follows:

$$Q_{(n+1)x(n+1)} = \begin{bmatrix} 1 & 0 & . & . & . & 0 \\ p & q & . & . & . & 0 \\ p^2 & p*q & . & . & . & 0 \\ . & . & . & p^{r-c}q^{c-1} * D(r,c) & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ p^n & p^{n-1}*q & . & . & . & q^n \end{bmatrix} \tag{20}$$

A general term of Q matrix is given by $p^{r-c}q^{c-1} * D(r,c)$, where r is the row number, c is the column number and $D(r,c)$ is defined as follows:

- $D(r,c) = 0$, if $r < c$
- $D(r,c) = 1$, if $r \geq c$

Equations (20) can be proven by induction.

Let us finish this section by briefly discussing the existing results on first exit times, as well as highlighting a few ideas on how these results may be used to characterize BCC processes analytically. The analytical solution can largely be facilitated by assuming constant boundaries. The cases of non-constant or stochastic boundaries are somewhat more challenging and to our knowledge these cases have not been solved in the literature. One possible solution is to adopt the techniques which have been

[3] Interested reader may find an excellent description on how to use the scale function in order to calculate upper-boundary crossing probabilities in Karlin and Taylor (1981). A less technical discussion focusing on Geometric Brownian Motion can be found in Karlin and Taylor (1998).

used to solve for first passage times (the case of one boundary) and are described in Valov (2009) or in Redner (2001).

The first-exit time distribution functions as well as their Laplace-transform can be found in the handbook of Borodin and Salminen (2002) for several processes, such as for the Brownian Motion with no drift (p. 212), for Brownian Motion with drift, (p.309), for Bessel process of order 0.5 (p.309) and finally for Geometric Brownian Motion (p.627). In page 109, the authors also discuss briefly a general method for solving for these distributions. More detailed proof for the Geometric Brownian Motion is available in Lin's paper (1998).

The typical procedure for deriving the first exit time distribution, the cumulative first exit time distribution and their Laplace Transform is the following. First, by subtracting the expected value from the original stochastic process, we obtain a martingale. If the initial value of the martingale is known, then we can express the expected value of the martingale at the first exit time. This is so because according to the Optimal Sampling Theorem (or Doobs lemma), the expected value of a martingale conditioned on the information available at time zero is equal to its initial value at any time-period. By rearranging this expected value, we can obtain the Laplace transforms of the first exit times. The probability distribution functions can then be derived by inverting these Laplace transforms, and finally, the cumulative probability distribution functions can be derived by integration. Once the formula for the first-exit time distribution has been derived, we can calculate the corresponding discrete first exit time distributions.

$$fetd(k) = \int_{t_{k-1}}^{t_k} fetc(t)dt\,, k = 1 \dots n \tag{21}$$

Once the discrete first-exit time distribution is known, the iterative procedure described by Equations (14)-(17) can be used.

3. Estimating BCC distributions

This section focuses on how to actually estimate BCC-distributions. As the BCC distribution has not been introduced in the literature beforehand, we cannot compare our results against standard benchmarks. Instead, we calibrate our method by developing and comparing three different estimation procedures. Throughout this section we use n to indicate the simulated or the actual sample size and m to indicate the number of simulated sample paths or the number of available actual data series. Moreover, we introduce the count operator #, which essentially counts the number of cases meeting certain criteria.

3.1 Estimation methods

Method of Direct Estimation resulting in the Direct BCC-distribution $(Y_T(D))$: The following fairly straightforward method uses the fact that the number of boundary crossing events can simply be counted based on the data:

1. Apply the boundary-functions a_i and b_i for all $i = 1 \dots n$.
2. Restart the process at each boundary-crossing event. If the stochastic process is not restarted at fixed or stochastic X_0, but it is restarted at $X_0 + (X_{T^U} - b)$ or at $X_0 + (X_{T^L} - a)$, then the process should be started at $X'_0 \sim D\big(X_0 + q * (X_{T^U} - b) + (1 - q) * (X_{T^L} - a)\big)$, where D indicates distribution and q is the upper-boundary crossing probability. Alternatively, the counting process can be started after the first boundary-crossing event has occurred.
3. For each sample path, count the number of upper- $[y_1^U, y_2^U, \dots y_m^U]$ and lower boundary crossing events $[y_1^L, y_2^L, \dots y_m^L]$. The total number of boundary-crossing events can then be described as: $[y_1^{UL}, y_2^{UL}, \dots y_m^{UL}] = [y_1^U, y_2^U, \dots y_m^U] + [y_1^L, y_2^L, \dots y_m^L]$.

The BCC-probabilities may then be estimated for $T < n$ as:

$$p_k^U(T)_{\#} = \frac{\#(y_i^U = k)}{m};\ p_k^L(T)_{\#} = \frac{\#(y_i^L = k)}{m};\ p_k^{UL}(T)_{\#} = \frac{\#(y_i^{UL} = k)}{m} \tag{22}$$

If we are willing to assume that $\mathrm{p_k}(T)$ is a smooth distribution, it may be appropriate to apply smoothing algorithm on $\mathrm{p_k}(T)_{\#}$. The advantage of the method of direct estimation is in its simplicity. In particular, it does not require any specific assumption on the boundary structure. On the other hand, this method often produces only a single point-estimate, as in actual time-series data ($m = 1$) we can typically observe only one realization of the stochastic process. Also, this method is relatively less precise on rare events. For example, in order to have approximately 100 observations for estimating the critical value for the 1% significance level, it is advisable to have 100 000 or more realizations of the stochastic process, which may be computation- and memory intensive in many cases.

Recursive iteration using analytical first exit time distribution resulting in the Analytical BCC-distribution $(Y_T(A))$: The following method is a direct implementation of the algorithm described by equations (14)-(17):

1. Calculate the $fetd^U(k), fetd^L(k), fetd^{UL}(k), k = 1 \dots n$ based on the analytical first exit time distributing described in Equation (21).
2. Carry out the iteration described in Equations (14)-(17).

This method can only be used when Assumption 2 on the probability distribution of the first exit times is satisfied. Also, we should be able to evaluate the first exit time distributions analytically, which so far, have only been derived for the case of constant boundaries and only for a few stochastic processes.

Recursive iteration using estimated first exit time distribution resulting in the Recursive Estimated BCC- distribution $(Y_T(RE))$: If the stochastic process is either not described parametrically or the first exit time distribution is unavailable, then it is still possible to approximate BCC processes by first estimating the first-exit time distributions from actual or from simulated data, and then by carrying out the iteration described in Equations (14)-(17):

1. Apply the boundary-functions a_k and b_k for all $k = 1 \dots n$ on the available data (assuming we have $m = 1, \dots j$ data-series each having lengths n_j) or on the simulated data of m sample paths, each having lengths n.
2. Restart the process at each boundary-crossing event. If the stochastic process is not restarted at fixed or stochastic X_0, but is restarted at $X_0 + (X_{T^U} - b)$ or at $X_0 + (X_{T^L} - a)$, then the process should be started at $X'_0 \sim D\big(X_0 + q * (X_{T^U} - b) + (1 - q) * (X_{T^L} - a)\big)$, where q is the upper-boundary crossing probability. Alternatively, the counting process can be started after the first boundary-crossing event has occurred.
3. Estimate first exit time distribution by collecting the upper boundary crossing times, $\left[t_1^U, t_2^U, \dots t_j^U\right]_{1x\left(\sum_{j=1}^m Y^U(n_i)\right)}$, the lower boundary-crossing times $[y_1^L, y_2^L, \dots y_m^L]_{1x\left(\sum_{j=1}^m Y^U(n_i)\right)}$ and finally, the total number of boundary-crossing times $\left[t_1^{UL}, t_2^{UL}, \dots t_j^{UL}\right]_{1x\left(\sum_{j=1}^m Y^{UL}(n_i)\right)}$. Note that these matrixes have stochastic lengths. Using these observations, the probability distribution for the first exit time is estimated as:

$$fetd_{\#}^{U}(t) = \frac{\#(t_i^U = t)}{\sum_{j=1}^m Y^{UL}(n_i)}; fetd_{\#}^{L}(t) = \frac{\#(t_i^L = t)}{\sum_{j=1}^m Y^{UL}(n_i)}; fetd_{\#}^{UL}(t) = \frac{\#(t_i^{UL} = t)}{\sum_{j=1}^m Y^{UL}(n_i)} \quad (23)$$

4. Carry out the iteration described in Equations (14)-(17).

In case we are willing to assume that fetd(t) is a smooth function, then it may be appropriate to apply some kind of smoothing algorithm on fetd$(t)_{\#}$. This is especially useful as the pure counting estimation reduces the degrees of freedom, which can be problematic. The main advantage of this method lies in its relative accuracy in case of rare events, which is of course highly important for statistical specification testing. Intuitively, this method uses more information as in this case we not only consider the number of boundary crossing events, but also the timing of the crossings. In addition to that, this method can be used without a parametric specification based on the data only.

Finally, if the process is not defined parametrically, then the second assumption cannot be verified directly, and it is advisable to test for the validity of this assumption. Cross-sectional and time series wise homogeneity can, for example, be tested by dividing the sample data into two separate sub-samples, and then by estimating the first exit time distribution as well as the boundary crossing counting distribution separately in these two sub-samples. The Kolmogorov-Smirnov statistics then can be used to quantify the distance between the two distribution functions. Alternatively, K-sample Anderson–Darling test may be used for testing whether the first-exit times, or the boundary crossing counts in the two sub-samples come from the same population. Of course, this method may be complemented by testing for structural breaks.

3.2 Calibrating estimation methods

As explained in the review of Fan (2005) and in the corresponding commentaries of Phillips and Yu (2005) and of Sørensen (2005), the most important problems of non-parametric statistical tests are finite-sample and sampling biases. This latter one occurs when the null hypothesis is a continuous stochastic process to be tested on discretely sampled data. The three different methods devised in the previous section provide us an opportunity to measure and quantify the magnitude of these issues. In most practical cases, only one or two methods are available. Yet in the case of the Brownian motion to be analyzed shortly, we can estimate the BCC distribution using all three methods. Therefore, the next section is devoted to the case of Brownian motion. The potential issues are summarized in Table 1.

	Direct BCC-distribution	Analytical BCC- distribution	Recursively Estimated BCC-distribution
Assumptions			
(2) Time-independent first exit time distribution	Not required	Required	Required
(3) Only one boundary-crossing	Required	Not required	Required
Common statistical issues			
- Finite-sample bias	Present in small sampled	Not present	Present as an estimation error
- Sampling bias	Present, can be solved by simulating extreme points.	Not present	Present, can be solved by simulating extreme points.

Table 1: What are the common statistical issues and potential violations of the assumptions when estimating Boundary-Crossing Counting distributions?

In the case of the Analytical BCC-distribution, the only assumption required is the time-independent first exit time distribution. For the Brownian motion we know this is true. Therefore, in this case we are able to estimate the true non-asymptotic BCC-distribution with arbitrarily precision. The other two methods both suffer from sampling bias due to unregistered boundary crossings. In addition to that, the Direct BCC distribution suffers from small-sample bias even if the cross-section is large. Finally, the validity of the second and third assumptions can also be tested by comparing the different BCC distributions to each other.

Without loss of generality, we can assume that the standard deviation of the process is normalized to one: Simply we use one standard deviation as a measurement unit. Using the notations introduced above, we compare

$$Y_T^{UL}(X_t, a = -6, b = 6, A); Y_T^{UL}(X_t, a = -6, b = 6, RE); Y_T^{UL}(X_t, a = -6, b = 6, D), \tag{24}$$

that is to say, we are going to compare the analytical, the recursive estimated, and finally the direct boundary crossing counting distribution generated over the standard Brownian motion by constant boundaries -6 and +6. Note that the choice of the boundary at this stage is arbitrary, and we are going to show BCC distributions for other constant boundaries if the choice of the boundaries influences the results.

Let us begin by analyzing the two iterative methods. As they only differ in the way first exit time distribution is obtained, let us focus first on the first exit time distributions. In order to obtain the analytical solution, let us assume that X_t is generated by sampling from a continuous Wiener process, using some constant sampling frequency t:

$$X_t = X_{t-1} + \mu, \mu \sim N(0,1) \tag{25}$$

As for the analytical BCC distribution, the first exit time distribution can be expressed using the following theta functions (Borodin and Salminen, 2002, p. 640, notation slightly modified):

$$ss_t(v,z) \approx \sum_{k=-k^*}^{k^*} \frac{z - v + 2k * z}{\sqrt{2 * \Pi} * t^{3/2}} e^{\frac{-(z-v+2k*z)^2}{2*t}}, v < z \tag{26}$$

Also,

$$cc_t(v,z) \approx \sum_{k=-k^*}^{k^*} (-1)^k * \frac{v + z + 2k * z}{\sqrt{2 * \Pi} * t^{3/2}} e^{\frac{-(v+z+2k*z)^2}{2*t}}, v < z \tag{27}$$

Where in both cases the approximation improves as k^*goes to infinity. In our calculations, we have used k^* = 1000 although k^* = 50 already provides decent approximation. Using these functions, the first exit time for upper or lower crossing can be calculated as:

$$fetd^U(t) \approx ss_t(X_0 - a, b - a), fetd^L(t) \approx ss_t(b - X_0, b - a) \tag{28}$$

Note that $X_0 - a < b - a$ and $b - X_0 < b - a$, therefore, conditions in formula (26) are fulfilled. As for the case of upper and lower crossings both:

$$fetd^{Ul}(t) \approx cc_t\left(\frac{b + a - 2X_0}{2}, \frac{b - a}{2}\right) \tag{29}$$

Also note that $\frac{b+a-2X_0}{2} < \frac{b+a-2a}{2} = \frac{b-a}{2}$, therefore, the condition in formula (27) is fulfilled. Let us also solve for the first exit time distribution by estimating[4] it from simulated data. For this purpose, we have simulated 1000 sample paths, each having lengths of 2520 observations. Estimations were done using equations (23).

Sampling issues

The estimated first exit time distributions have the shape of the Inverse-Gaussian distribution (Feller, 1971, p.52 or Lin, 1998) as in Figure 3, while BCC distributions have the shape of a normal distribution,

[4] Simulations were carried out using Matlab, empirical histogram was obtained by the counting estimator as described in equations (24). The probabilities for rare events were estimated using normal kernel.

yet they are not, as BCC distributions are discrete.[5] The diagram also reveals that the estimated- and the analytical first exit time distributions are not equal, the difference appears to be significant.

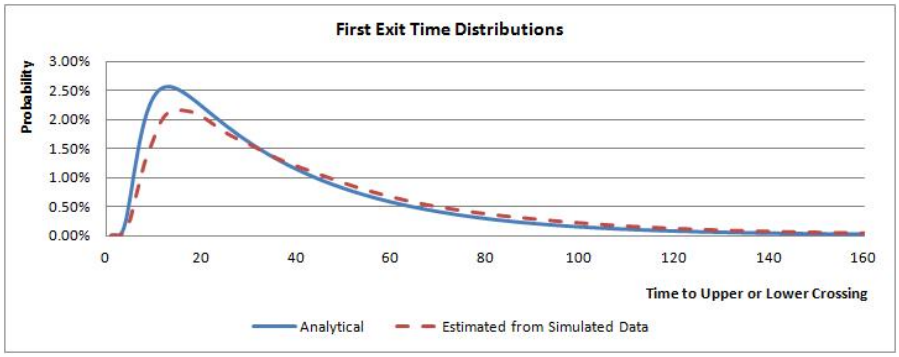

Figure 3: Analytical and estimated first exit time distribution of the standard Brownian motion, for boundaries a = -6; b =6.

The estimated first exit time distribution tends to underestimate the probability of shorter exit times and overestimate the probability of longer exit times. The estimation does not improve as the sample size increases therefore this is not a small-sample bias. What causes this difference? It is due to sampling: the analytical solution treats time as a continuous variable, while the simulated solution works with discretely sampled observations. In this latter case, the value of the random variable between the two observations is unknown. It may very well be possible that the random variable crosses the boundaries between the two observations, yet at the moment of observation, the random variable is no longer outside the boundaries. Such unobserved crossings are not registered as boundary-crossings in case of discrete sampling.

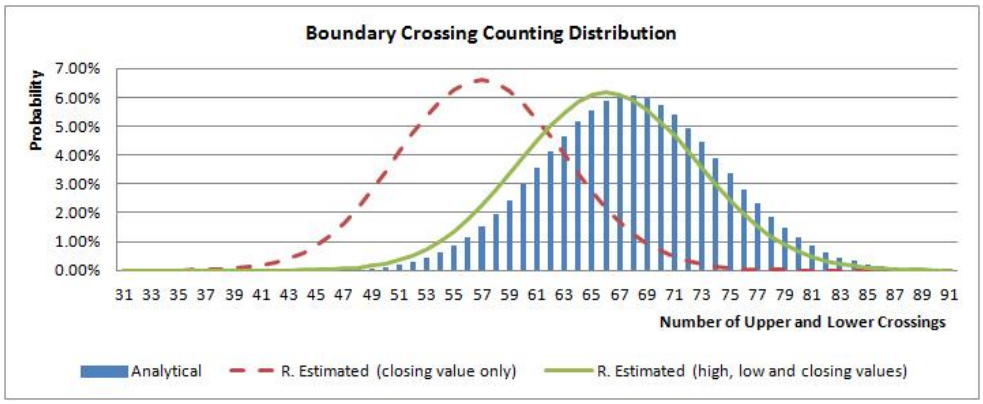

Figure 4: Reducing bias by simulating minimum and maximum values. Analytical and Recursive Estimated boundary crossing counting distribution for 2520 observations, upper and lower crossing both, for standard Brownian motion and constant boundaries -6 and 6 or using the short notation introduced above $Y_{2520}^{UL}(BM(0,1), a = -6, b = 6, A)$ *and* $Y_{2520}^{UL}(BM(0,1), a = -6, b = 6, RE)$

[5] Still, we sometimes use solid lines in the diagrams, so that readers can differentiate between the various distributions.

The simulated first exit time distribution therefore underestimates the probability of shorter first exit times, and overestimates the probability of longer first exit time, as shown in Figure 3. As the simulated first-exit time distribution is biased, naturally the simulated BCC distributions are also biased, as shown in Figure 4.

The solution is to simulate not only the sampling value in discrete intervals, but also to simulate the minimum and the maximum value between the sampling intervals. In Figure 4, the Recursively Estimated BCC-distribution is calculated based on simulated data under the usual sampling method. In addition to that, we have calculated a second RE-BCC-distribution that is based on minimum and maximum values as well. In this case, we have increased the sampling frequency by hundred. That is to say, instead of generating 2 520 observations using mean zero and unit standard deviations, we have simulated 252 000 observations using mean zero and standard deviation of 1/100. Then for each time-period we have selected not only the closing value, but also the minimum and the maximum value[6] as well. The first simulated BCC distribution underestimates the number of boundary crossing events. This is in-line with the explanation given above: the simulated distribution only counts those boundary crossing events where the random variable remains outside of the boundaries at the moment of sampling. On the other hand, the bias in the second RE-BCC distribution is largely reduced. Therefore, sampling bias can be substantially reduced by taking into account the minimum and the maximum values as well.

In practice, the data is typically sampled discretely. In this case we have to assume that the data generating process under the null hypothesis and the data generating process under the alternative hypothesis has been sampled using the same sampling mechanism. In this case the sampling bias influences the two BCC distributions in a similar manner. Therefore, if the BCC distribution under the null hypothesis and the BCC distribution under the alternative hypothesis are considerably different, then the two data-sets are unlikely to be generated by the same data-generating process. Consequently, we can still use the BCC-distribution for specification-testing of data-generating processes: We simply have to compare two equally biased distributions. In some applications – especially in the domain of finance – sampling is done differently: for instance, in case of financial data[7] typically the minimum and the maximum values are also observed. In this case, we can decide to make use of as well as simulate the minimum and the maximum values and work with the true distributions. Alternatively, we can rely on the closing prices only and compare two equally biased BCC-distributions.

[6] We acknowledge that this method underestimates the maximum value and overestimates the minimum value. There are more sophisticated methods for generating extreme prices as explained, for example, in Mcleish (2002). As here we use extreme values for illustrative purpose only, we would not substantially benefit from using more sophisticated models.

[7] The typical structure for daily financial prices is to have open, high, low and closing prices.

Small sample bias

Let us continue our analysis by comparing the Analytical BCC distribution with those obtained using the method of Direct Estimation using equations (22). For the estimation, we have simulated 5000 sample paths with sample lengths of 252 observations, which is approximately 1 year of observations in financial price-series. We have simulated minimum and maximum values as well. As each sample path results in one observation, overall we have 5000 observations. The direct estimation for the BCC distribution is essentially a normalized, non-smoothed histogram based on these 5000 observations, which is compared with the Analytical BCC distribution. The diagram below reveals that the analytical and the direct distribution are not equal. The difference between the two distributions is significant. The direct estimation tends to overestimate the probability of fewer boundary crossing events and underestimate the probability of more boundary crossing events.

The difference between the two distributions is due to the fact that we do not take into account the evolution of the stochastic process after the last boundary-crossing event. In other words, we do not observe where and when the stochastic process crosses the boundaries after it has been restarted for the last time. Naturally, as sample size increases, the role of this last observation boundary-crossing diminishes. For boundaries of five standard deviations, a 1000 observation almost completely eliminates the difference between the Direct and the Analytical or the Recursive Estimated distributions in case of the BM(0,1) process. In the actual specification testing, we use at least 10 years of data, therefore small sample bias is not likely to influence our results.

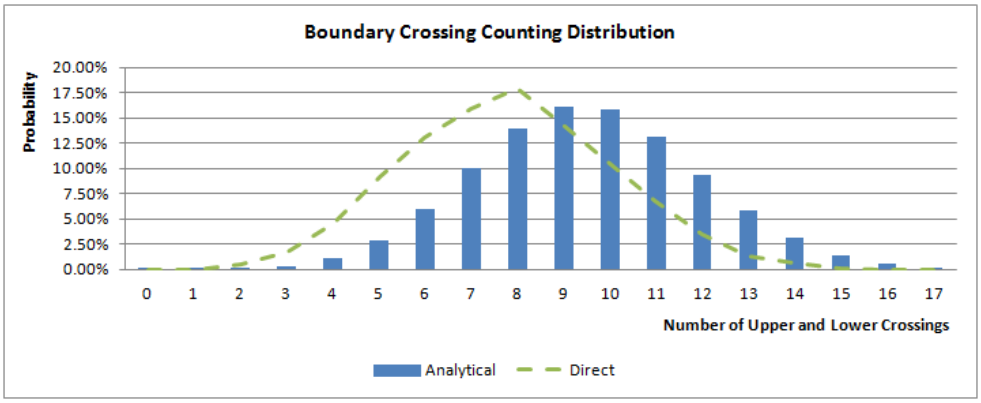

Figure 5: Small sample bias of the Direct method. Analytical and Direct boundary crossing counting distribution for 252 observation, upper and lower crossing both for standard Brownian motion and constant boundaries -5 and 5 or using the short notation introduced above $Y_{252}^{UL}(BM(0,1), a = -5, b = 5, A)$ *and* $Y_{252}^{UL}(BM(0,1), a = -5, b = 5, D)$

In any case, BCC-test is more of a large sample tool; therefore, this issue is unlikely to cause serious problems in practical applications.

Issue of selecting critical values and the convention of closest value

The next issue is related with selecting critical values from a discrete distribution. Often test-statistics have continuous distributions; therefore, selecting the desired critical values resulting the usual probability for type-1 errors (1%, 5%, 10%) is relatively straightforward. BCC-distributions, on the other hand, are discrete distributions; thus we are unable to find the exact critical values matching the desired type-1 error probabilities. For example, in Figure 5, $P\left(Y_{252}^{UL}(BM(0,1), a = -5, b = 5, A) \leq 5\right) = 4\%$ and $P\left(Y_{252}^{UL}(BM(0,1), a = -5, b = 5, A) \leq 6\right) = 8.9\%$. Consequently, if we seek a critical value with probability of type-1 error close to 5%, then we have to choose either 5 as a critical value with probability of type-1 error of 4%, or 6 as a critical value with the probability of type-1 error of 8.9%. This is rather a matter of convention and we choose critical values so that the type-1 errors are closest to 5%.

Violation of assumptions

Finally, deviation between BCC-distributions may be due to the violation of Assumptions 2-3. Figure 6 below showing the direct and the recursive estimated BCC-distributions, highlights the importance of boundary selection. If we place the boundaries at 3 standard deviation distances, Assumption 3, stating that the probability of boundaries crossed twice within one time interval is negligible, is violated. Note that if boundaries are chosen to be wider - for example, five standard deviations - then as Assumptions 3 gets stronger, the difference between the estimated recursive and the direct distribution gets reduced to a great extent.

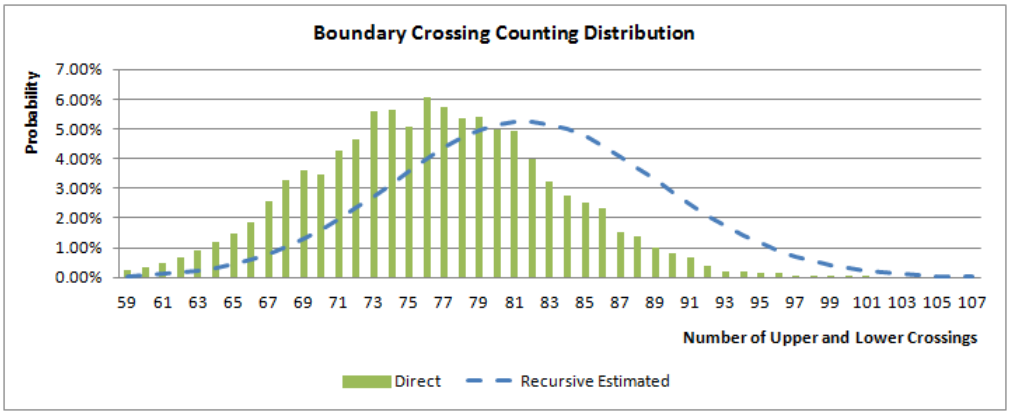

Figure 6: Recursive Estimated and Direct boundary crossing counting distribution for 2520 observation, upper and lower crossing both, for standard Brownian motion and constant boundaries -3 and 3 or using the short notation introduced above $Y_{252}^{UL}(BM(0,1), a = -3, b = 3, RE)$ *and* $Y_{252}^{UL}(BM(0,1), a = -3, b = 3, D)$

This observation is particularly important when dealing with financial data, as due to the presence of conditional heteroscedasticity, large variations occur more frequently. Therefore, we worked with 8 standard deviation distances, a level at which the problem described above does not manifest.

Further remarks

From the above, we can conclude that generally, the outcomes of the three different estimation methods are three different BCC-distributions. Therefore, in statistical applications, we have to be careful when comparing BCC distributions that are not estimated using the same method. To put it another way, it is advisable to apply the `ceteris paribus' principle in the sense that when testing the equality of two data generating process by comparing their BCC distributions, we have to keep everything else constant: we have to try to use the same estimation method as well as the same lower and upper boundaries. Unfortunately, this is not possible most of the time, as in actual data we typically have only one realization of the stochastic process. Therefore, we typically compare a single direct estimate with a recursively estimated distribution. The following general estimation procedure minimizes the risk that our inference is driven by statistical errors:

- Calculate the BCC-distribution under the null hypothesis using the method of recursive estimation as well as using the method of direct estimation.
- If, under the null hypothesis, the two estimation procedures result an essentially identical distribution, we can use the more accurate RE-BCC-distribution for inference.
- On the other hand, if the two estimation procedures does not yield similar distributions, then either we have to adjust the settings by, for example, increasing the magnitude of the boundaries, or we have to simulate a large amount of data at a level where the histogram-based direct estimation results sufficient accuracy at the desired significance level.
- In both cases, we would mostly carry out the test by comparing the simulated distribution with the direct point-estimate of the actual data.

Also, for applications when not only the distance between two BCC distributions but the BCC distribution itself is of interest, we have to keep in mind that the `naïve' estimation of the BCC distribution may be biased due to sampling error, which can be corrected by using minimum and maximum values for the estimations. Nevertheless, the BCC-method is a non-asymptotic one; therefore, we are often able to quantitatively estimate the magnitude of these discrepancies, which is not always possible when applying asymptotic results.

4. Statistical specification testing

We begin our analysis by showing how to apply our method for unit-root testing. As the nature of time-series is one of the most fundamental questions in economics and in finance, starting from Dickey and Fuller (1981) and Phillips and Perron (1988), econometricians have devised a large number of tests for identifying unit-roots. As reviewing these tests would require a separate article, and it has already been done, for example, by Phillips and Xiao (1998), we do not attempt to provide a comprehensive review here. Rather we threat the most simple (and hence typically unrealistic) case of a single time-series with normally distributed *IID* errors in detail. Our goal here is to demonstrate how a new statistical tool

operates in an environment that is familiar to most researchers. We continue our analysis with financial applications, and finally we conclude by a small remark on the mean-variance framework.

4.1 Unit Root Testing

Let us assume that we want to differentiate between a unit-root process and a near-unit-root process.

$$X_t = \alpha * X_{t-1} + \mu, \mu \sim N(0,1), H_0: \alpha = 1, H_1: \alpha \neq 1 = 0.9 \quad (30)$$

The alternative hypothesis typically assumes that $\alpha < 1$, while here we could test for explosive ($\alpha > 1$) null-hypothesis as well. Nevertheless, we chose the non-explosive case so as to be able to compare our results to common unit-root procedures.

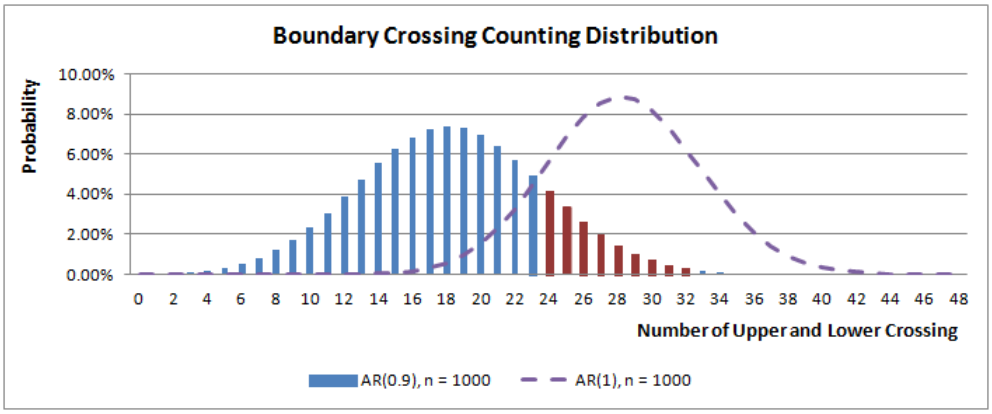

Figure 7: Testing Unit-root process against near unit-root process using boundaries of five standard deviations. We compare $Y_{1000}^{UL}(AR(1), a = -5, b = 5, RE)$ *and* $Y_{252}^{UL}(AR(0.9), a = -5, b = 5, RE)$.

The logic of the test (also illustrated in Figure 7) is as follows: If we were to collect 1 000 observations from the unit root process, then in approximately 5% (6.1%) of the cases we would observe 24 or fewer boundary-crossing events. On the other hand, in case of an AR(0.9) process, we would observe 24 or more boundary-crossing events and would falsely accept the unit root hypothesis in 33.6% of the cases. Naturally, if we were to test the explosive alternative hypothesis, we would do a similar one-sided test on the right-hand-side of the distribution.

Our results are compared with several parametric unit-root tests. In particular, we compared the BCC test with the Augmented Dickey-Fuller test or ADF test (Dickey and Fuller, 1982), the Phillips-Peron test (Phillips and Perron, 1988), the variance ratio test (Lo and MacKinlay, 1988), and finally the KPSS test (Kwiatkowski, Phillips, Schmidt and Shin, 1992). We implemented these tests using the corresponding build-in Matlab functions. We use 5% nominal significance level. The null-hypothesis is the unit-root hypothesis except for the KPSS test, where the null hypothesis is the trend-stationarity one. The logic of our analysis is as follows.

- We simulate 1 000 AR(1) sample paths and then carry out the tests described above for various sample sizes. In case of the first three tests, the number of rejections divided by 1000 provides

us with the actual size. In case of the KPSS test, the number of acceptance divided by 1000 provides us the actual size.

- We also simulate 1000 AR(0.9) sample paths and then carry out the tests described above for various sample sizes. In case of the first three tests, the number of acceptances divided by 1000 provides us with the probability of type 2 errors. In case of the KPSS test the number of rejections divided by 1000 provides us the probability of type 2 errors.

In addition to that, we have also implemented a non-parametric test inspired by Hong and Li (2005), which tests the joint hypothesis of i.i.d and uniformity for a series of probability integral transformed (PIT) data. We have implemented this test using the "HLTest" function of Ghalanos (2013) rugarch package using the default settings provider in the author's web page. More precisely:

- We have obtained the cumulative distribution function for the model under the null hypothesis using simulation based on 1000 sample paths each having 2520 observations.
- We have applied this function to transform the residuals generated by the models under the alternative hypothesis so as to obtain the generalized residuals.
- We substituted these generalized residuals to the HL_function of the rugarch package.
- The actual size of the test was obtained by dividing the number of rejections of the model under the null hypothesis by the total number of sample paths.
- The probability of type 2 errors was obtained by dividing the number of acceptances on the model under the alternative hypothesis by the total number of sample paths.

We have used the PIT function of the first differences. As the test is based on the probability integral transform, we refer to this test as PIT-test.

Unit Root tests at 5% Nominal Significance level, Data Generating Process is AR(0.9)					
	Sample Size				
	100	252	756	1 000	1 764
Power of Parametric tests					
- ADF	44.9%	84.8%	99.5%	100.0%	100.0%
- KPSS	5.7%	2.8%	1.5%	1.6%	1.3%
- Phillips Perron	44.9%	84.8%	99.5%	100.0%	100.0%
- Variance ratio	6.5%	11.2%	27.5%	35.8%	56.0%
Size of Non-parametric tests					
- PIT test(first difference)	18.1%	17.9%	16.9%	16.0%	16.8%
- BCC(a = -5, b = 5, RE, p*=20%)	14.0%	15.3%	20.9%	19.5%	20.4%
- BCC(a = -5, b = 5, RE, p*=5%)	2.4%	6.6%	5.4%	6.1%	4.6%
- BCC(a = -4, b = 4, RE, p*=5%)	2.9%	5.8%	4.4%	5.8%	5.3%
Power of Non-parametric tests					
- PIT test(first difference)	15.00%	20.00%	50.40%	61.00%	86.50%
- BCC(a = -5, b = 5, RE, p*=20%)	11.50%	35.84%	76.66%	83.43%	95.31%
- BCC(a = -5, b = 5, RE, p*=5%)	0.00%	22.31%	53.64%	66.35%	83.94%
- BCC(a = -4, b = 4, RE, p*=5%)	2.71%	17.31%	39.34%	53.55%	74.15%

Table 2: Statistical power at 5% nominal significance level when the data generating process is AR(0.9).

The first part of Table 2 reveals that the BCC test is generally less powerful than the ADF and the Phillips-Peron tests, but it is more powerful than the KPSS and the Variance ratio test. The lower power of the KPSS test may partially be explained by the fact that the actual size of the test is much lower than the nominal size, which is also noted by, for example, Caner and Kilian (2001). On the other hand, the Variance ratio test has been primarily designed to identify non-normal disturbances (for example heteroscedasticity, conditional dependence in the variance structure), therefore, it is not primarily aimed to differentiate between the AR(0.9) and the AR(1) processes.

This finding is in-line with our expectations: The BCC-test is a non-parametric one, encompassing the full specification, not just the first or second moments. Intuitively, such general test should perform relatively well under various circumstances, yet it is likely that a specific test derived for a specific problem performs better on that specific problem than a general test.

The second part in Table 2 reveals that the PIT-test suffers from a certain size-distortion. Therefore, for the sake of comparison, we have calculated the power of the BCC-test for an actual size comparable to the PIT-test and for the usual size as well. Hence, we can conclude that in case of unit-root testing the BCC-test is more powerful than the PIT-test if similar test-size is applied or it is equally powerful, yet suffers from significantly less size-distortion if the usual size is applied. Therefore, the BCC-test has higher discriminative power than the alternative non-parametric test. The last two rows of Table 2 show the statistical power under different upper and lower boundaries at 5% nominal significance level. We find that the power of the BCC-test increases as the boundaries get wider. This result is again line with our intuition, as stationary and unit-root processes differ in how likely they are to take extreme values. If we choose narrow lower and upper boundaries, we have to restart the process frequently, and thus we do not allow this difference to manifest. On the other hand, if we choose too large boundaries, then we do not observe enough boundary-crossing events. The statistical power appears to be a convex function of the magnitude of the boundaries. Indeed, simulation reveals that for a sample size of 1000 observations, the power of the unit-root test is generally highest when using 5-7 standard deviation distances.

4.2 Specification testing for financial time-series

Let us continue our analysis by testing commonly used price models. The Black-Scholes model (Black and Scholes, 1973.) is probably the most widespread model, which assumes that prices follow a Geometric Brownian Motion (further referred as GBM). Its popularity is due to its simplicity, yet it also captures important characteristics of the price process, namely the linear growth of the variance of the return fluctuations with timescale. As a drawback, the Black-Scholes model fails to characterize the volatility structure of the market returns. The most common approach to correct for this drawback is to apply ARCH/GARCH models first introduced by Engle (1982). We are going to consider many different GARCH specifications.

We begin our analysis by studying how different these models are in finite samples. For this purpose, we compare the statistical power of the BCC-test and the PIT-test when the null hypothesis is Bollerslev's GARCH(1,1) model using Gaussian distribution for the error term.

Base-line model: GARCH(1,1) with Gaussian Conditional Distr. Estimated on DJI(2002-2012)					
	Sample Size				
	100	252	756	1764	2520
Actual size					
- PIT Test based on first difference	22.0%	20.4%	19.9%	19.6%	21.5%
- BCC Test, Lower and Upper Crossing Both	13.2%	7.6%	6.6%	5.6%	5.3%
Power againts Geometric Brownian Motion					
- PIT Test based on first difference	19.4%	17.2%	16.2%	19.0%	15.2%
- BCC Test, Lower and Upper Crossing Both	11.6%	6.3%	14.0%	28.8%	38.6%
Power againts Garch(2,2), Gaussian					
- PIT Test based on first difference	20.6%	20.2%	19.6%	21.2%	20.4%
- BCC Test, Lower and Upper Crossing Both	11.5%	6.3%	5.2%	4.4%	4.3%
Power againts Garch(1,1), Student-t					
- PIT Test based on first difference	20.8%	18.3%	21.7%	18.9%	23.2%
- BCC Test, Lower and Upper Crossing Both	32.5%	25.9%	31.1%	39.2%	44.1%
Power againts E-Garch, Gaussian					
- PIT Test based on first difference	22.2%	19.0%	21.4%	21.2%	20.2%
- BCC Test, Lower and Upper Crossing Both	18.2%	11.1%	15.9%	24.0%	29.0%
Power againts E-Garch, Student-t					
- PIT Test based on first difference	18.6%	20.6%	18.8%	20.4%	21.9%
- BCC Test, Lower and Upper Crossing Both	35.6%	29.3%	35.4%	44.8%	50.4%
Power againts GJI, Gaussian					
- PIT Test based on first difference	22.0%	24.9%	22.8%	22.6%	22.8%
- BCC Test, Lower and Upper Crossing Both	11.1%	6.3%	5.5%	5.4%	6.2%
Power againts GJR, Student-t					
- PIT Test based on first difference	19.0%	21.2%	23.6%	23.6%	21.6%
- BCC Test, Lower and Upper Crossing Both	29.4%	22.1%	23.7%	26.7%	28.6%

Table 3: The actual size and statistical power of the PIT-test and the BCC-test. When calculating the BCC-test, we have used seven standard deviation distances and have calculated the BCC distribution using the recursive estimation method.

We consider several alternative hypotheses: The first alternative is the Geometric Brownian Motion. The second alternative is the GARCH(2,2) model, allowing for more sophisticated conditional heteroscedasticity structure. The third specification is the GARCH(1,1) model, using the Student-t distribution with eight degree of freedom for modeling the error term. The fourth and the fifth

specification uses Nelson's (1991) exponential general autoregressive conditional heteroscedasticity model (further referred as E-Garch), allowing the error term to be distributed under Gaussian and Student-t distributions with eight degree of freedom. Finally, the last two specification is based on Glosten, Jagannathan and Runkle (1993)'s specification (further referred as GJR model) that includes leverage terms for modeling asymmetric volatility clustering. In this setup, large negative changes are more likely to be clustered than positive changes. Each model was estimated using the closing price of the Dow Jones Industrial average between 2002 and 2012. Therefore, all specification models the same time-period, and they only differ in how to describe this very same time-period. Each model was implemented using Matlab's Econometric Toolbox.

The first two rows of Table 3 reveal that the PIT-test suffers from a certain size-distortion, while the BCC-test – being a non-asymptotic method – has an actual size very close to the 5% level. The PIT-test is typically unable to differentiate between the various models. The BCC model appears to be useful when differentiating between models under different error-distributions. As error-distribution is one of the most fundamental questions in modeling financial time series, the BCC-test appears to be a useful diagnosis test in this regard. On the other hand, the BCC-test is not useful for finding the right structure of the conditional heteroscedasticity, as it is unable to differentiate, for example, between the GARCH(1,1) and the GARCH(2,2) specifications. Likely, these models do not differ in boundary-crossing behavior, rather, they differ in what happens between the boundary-crossings.

Let us continue our analysis by testing these models on actual financial data. Although BCC-test could theoretically be used for evaluating specifications including high, low and closing prices, most articles use closing-prices only. In order to maintain comparability to the rest of the literature, we will also rely on closing prices. In case of daily closing prices, sampling in actual data is done by the opening hours of the stock exchanges. As Figure 4 highlights, BCC distributions are somewhat influenced by sampling, therefore, let us briefly review this issue.

In the data, we do not have observations for weekends and other holidays. Many sensible approaches may arise as analyzed in more details by, for example, Oscar Jordá (undated) in a working paper. One possible view may be the concept of "economic time", which roughly assumes that time stops in weekends and holidays. From this point of view, the time-difference between Friday and the coming Monday is the same as the time-difference between Monday and Tuesday. Alternatively, one may work using physical time: in that case the time difference between Friday and the coming Monday is three times the time difference between Monday and Tuesday.

The Black-Scholes framework predicts that the expected first and second moment are both proportional with time. If time is measured as economic time, the first and second moment of log returns should be independent of the physical time. If time is measured as physical time, the first and the second moment should be proportional to the physical time passed between the two observations.

Daily log returns of the Dow Jones Industrial Average, 1928-2012			
Physical time between two observations (days)	# of observations	Mean log return	Variance of log return
All	20868	0.019%	0.014%
1	16211	0.042%	0.012%
2	322	0.058%	0.017%
3	3865	-0.067%	0.019%
4	461	-0.116%	0.021%

Table 4: Relationship between the physical time passed between two observations and the mean and the variance of the log returns

The short summary table above cannot fully confirm either approach. The first column of the table shows the physical time passed between two observations. The average log-return for observations when three days passed between two observations appear to have lower log-returns than for those days when only one day passed between two observations. This effect (often called `Monday effect') is documented by French (1980) as typically describing the return between Friday and Monday. On the other hand, the variance appears to be a positive function of physical time, yet it is not proportional: 3 times increases in physical time do not lead to a 3 times increase in variances. To conclude, either simple approach appears to capture sampling perfectly. Incorporating a more complicated sampling assumption, however, would increase the complexity of the models, and we do not see any significant additional benefits of going down this path at this stage. Therefore, we are going to choose the `economic time hypothesis', as it appears to be somewhat more plausible than the physical time hypothesis.

Assumption 5: Time is non-stochastic[8] `economic time'.

Based on Assumption 5, we are going to use constant sampling when simulating prices with the financial models described above. We are going to compare the number of boundary crossing events in these simulated prices with the number of boundary crossing events observed in the closing prices of the Dow Jones Industrial Average (DJI) index,[9] of the Nasdaq Composite Index (IXID), and finally of the Australian index (AORD). More precisely,

- we begin by selecting 10-20 years of data from these indexes.
- In case of the Geometric Brownian Motion we carry out the standard maximum likelihood estimates (detailed, for example, in Gourieroux and Jasiak, 2001) so as to obtain parameter estimates for the Black-Scholes model. Likewise, we use Matlab's econometrics toolbox in order

[8] The non-stochastic nature of time is by far not trivial and there are models using stochastic trading times (Mandelbrot, 2004, p.207).

[9] We mostly work with stock indexes rather individual stocks in order to reduce the number of practical issues related with stock prices for instance the problem of jumps (gaps), survival biases, etc.

to obtain estimates for the various GARCH model. Therefore, we compare in-the-sample fits and not out-of-sample forecasts.

- We use these estimates to generate 1000 sample paths, each having 2520 or 5040 observations.
- Using these simulated data, we estimate BCC distributions using the method of direct estimation and the method recursive estimations. We do not use the analytical method, as the first exit time distribution for the GARCH process has not yet been derived.
- We compare the recursive estimated BCC distribution with the actual number of boundary crossing events observed in the data and obtained using the method of direct counting.

We select the boundaries to be at eight standard deviation distances, as at this level, the direct and the recursive estimated BCC-distributions are highly similar under the null hypothesis. That allows us to use the recursive estimated BCC distribution for calculating the critical values. The first row of Table 5 shows the actual number of boundary crossing events. This is followed by lower and upper critical values at 10% significance level for the different models and the corresponding p* value for the data. For example, in last column of the GARCH(1,1) specification, we can see that with approximately 90% probability we should observe between 90 and 139 boundary crossing events. On the actual data, we only observe 70 boundary crossing events. Observing 70 or fewer boundary-crossing events in case of simulated GARCH(1,1) data has the probability of 0.2%, therefore we reject this specification for modeling the Australian Stock Index. The first column of Table 5 is based on 2520 observations (approximately 10 years of data), and this sample is not sufficiently large to reject any of the specifications. This is not surprising, as non-parametric methods often require large amounts of data. The other three columns are based on 5040 observations (approximately 20 years of data), and each column rejects one specification at the usual significance level. It may be somewhat surprising that the Geometric Brownian Motion is accepted as specification for all four cases. Note that this result is very much driven by the choice of boundaries. If we were to choose five or six standard deviation distances, the GBM model would be typically rejected. This finding is in line with the 4^{th} stylized fact (Aggregational Gaussianity) of Cont (2001), who points out that the Gaussian nature of log-returns increases as the time scale over which returns are calculated increases. To put it another way, the Geometric Brownian Motion may very well be an adequate formulation when dealing with monthly data. However, the same formulation would be highly inadequate when dealing with intraday data. Perhaps that is why financial practices relying on frequent readjustments trades – such as dynamic hedging in option pricing – typically reject the GMB model. On the other hand, the theory of stochastic portfolio allocation uses the GBM hypothesis much more commonly.

Also, we frequently reject the GARCH(1,1) model, which is somewhat surprising, as out-of-sample forecast based comparisons sometimes prefer this specification over more sophisticated alternatives, as shown by Hansen and Lunde (2005). Note that not all studies shares this view: in the past years authors have produced hundreds of out-of-sample forecast based comparisons, reviewed by Poon and Granger

(2003) with inconclusive results. Here, we rely on an in-the-sample comparison which naturally favors more sophisticated models.

In-the Sample Fit of Common Price-models on Major Stock Indexes (critical values at 10% nominal size and p* values)				
	DJI 2002 - 2012 5040	DJI 1992 - 2002 5040	IXID 1992 - 2002 5040	AORD 1992 - 2002 5040
Observed number of boundary crossing events - Total number of crossings	31	59	69	70
Geometric Brownian Motion - BCC_all(a = -8, b = 8,RE), lower critical value - BCC_all(a = -8, b = 8,RE), upper critical value - Corresponding p value	26 41 60.3%	57 78 19.3%	57 79 81.2%	56 78 68.3%
Garch(1,1), Gaussian - BCC_all(a = -8, b = 8,RE), lower critical value - BCC_all(a = -8, b = 8,RE), upper critical value - Corresponding p value	17 36 38.2%	59 91 11.5%	47 80 56.9%	90 139 0.2%
Garch(2,2), Gaussian - BCC_all(a = -8, b = 8,RE), lower critical value - BCC_all(a = -8, b = 8,RE), upper critical value - Corresponding p value	16 33 24.6%	63 102 **4.5%**	46 77 38.9%	51 79 51.9%
Garch(1,1), Student-t - BCC_all(a = -8, b = 8,RE), lower critical value - BCC_all(a = -8, b = 8,RE), upper critical value - Corresponding p value	21 50 65.0%	48 81 63.8%	153 270 0.0%	58 94 63.0%
E-Garch, Gaussian - BCC_all(a = -8, b = 8,RE), lower critical value - BCC_all(a = -8, b = 8,RE), upper critical value - Corresponding p value	23 41 88.4%	61 89 5.3%	58 87 76.1%	58 84 88.5%
E-Garch, Student-t - BCC_all(a = -8, b = 8,RE), lower critical value - BCC_all(a = -8, b = 8,RE), upper critical value - Corresponding p value	22 41 94.9%	55 82 27.9%	55 86 93.8%	55 82 80.3%
GJI, Gaussian - BCC_all(a = -8, b = 8,RE), lower critical value - BCC_all(a = -8, b = 8,RE), upper critical value - Corresponding p value	16 34 26.8%	46 73 98.9%	42 72 17.3%	51 80 59.9%
GJR, Student-t - BCC_all(a = -8, b = 8,RE), lower critical value - BCC_all(a = -8, b = 8,RE), upper critical value - Corresponding p value	19 44 95.3%	47 78 78.3%	59 103 39.8%	56 89 85.4%

Table 5: How well do common price models represent major stock indexes? The BCC-test is based on 8 standard deviation distances; critical values are calculated for 10% probability of type-1 error.

The simple GARCH model is problematic because it does not take into account the asymmetry in upper and lower movements and hence overestimates the volatility for upper movements. This generates too many boundary-crossing events. Overestimation can be eliminated by either using the E-Garch specification with the Student-t distribution for the error terms, or by accounting for the asymmetric nature of financial prices using the BJR model. Therefore, we provide some evidence that without taking into account the asymmetry in the conditional variance structure, the probability of large upper-direction changes may be overestimated, which is detected by the BCC-test. The importance of asymmetry in financial markets is supported by other studies as well. For example Chou (2005) points out that the role of asymmetry is not negligible when dealing with price-ranges. Therefore, our findings contribute to the existing literature in this regard as well.

4.3 How good is the mean-variance framework?

Let us conclude the paper with an experiment regarding the mean-variance framework. This fundamental cornerstone of finance implies that in order to evaluate an investment, it is sufficient to consider the profitability of the investment typically measured as the first moment, and the risk of the investment typically measured as the variance of log price changes. The question may arise: Is it really sufficient to consider only these two factors? Testing this assumption would require a separate paper, yet BCC-processes may shed some light on this issue. One way to quantify this assumption is to say that if we standardize a log-return series (by deducting their mean and by dividing them by their standard deviation), the resulting random variable should be a unit-root process with zero mean and unit standard deviation. Is this really the case? We are going to experiment with this hypothesis using the following method:

- Let us select 20 years of daily log returns ranging from 1992-2012 for major stock indexes such as the Dow Jones Industrial Average (USA), S&P500 (USA), NASDAQ Composite (USA), CAC40 (France), DAX (Germany), Nikkei 225 (Japan), FTSE 100 (United Kingdom), and finally AORD (Australia, All Ordinary). These eight stock indexes cover a wide range of countries and industries, although we lack, for example, emerging markets.
- Let us normalize (standardize) the log-return of these indexes. When normalizing, we estimate the mean and the standard deviation separately for each time series.
- Let us compare the BCC distribution of these normalized log-returns and of simulated Gaussian returns for symmetric, constant boundaries at seven standard deviation distances. Distributions are estimated using the method of recursive estimation.

Figure 8 below reveals that the BCC distribution of the normalized log-returns does not equal the BCC-distribution generated by the standard normal distribution. In particular, the standard normal distribution predicts that the number of boundary crossing events should be between 62 and 94 observations with 95% probability. In case of the normalized log-returns, in approximately ¼ of the cases (23.1% of the cases), the observed numbers of boundary crossing events are either higher or lower. The

difference between the two distributions is significant, which can be tested by either applying Kolmogorov-Smirnov test or by simulations.

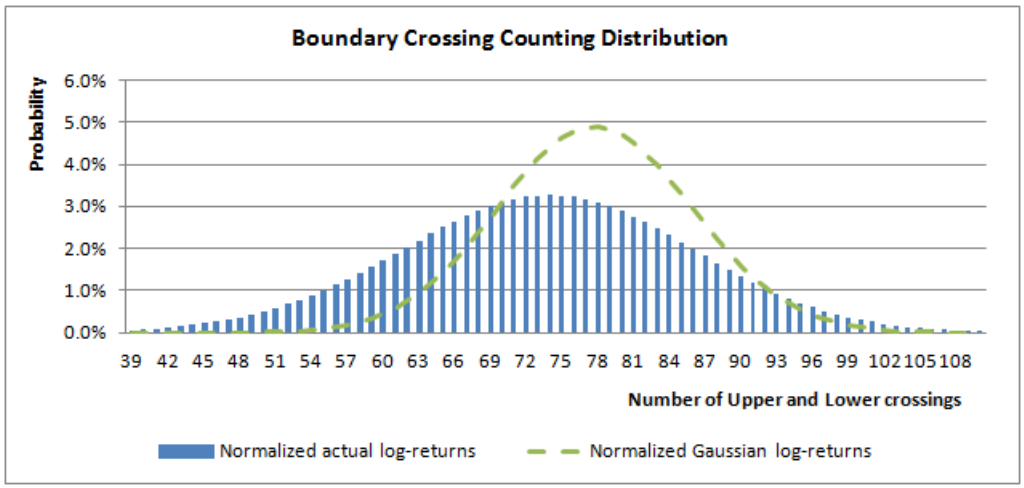

Figure 8: Estimated recursive boundary crossing counting distributions under constant symmetric boundaries of seven standard deviation distances. $Y_{5040}^{UL}(StockIndexes(normalized), a = -7, b = 7, RE)$, $Y_{5040}^{UL}(BM(0,1), a = -7, b = 7, RE)$

The BCC-test in this regard is a true omnibus type test. It reveals that for some reason, the Central Limit Theorem does not fully apply to log-return series even at relatively large sample size. Nevertheless, it reveals little on why this may be the case. In any case, Figure 8 highlights the richness of financial data.

5. Summary

In this paper we introduced and characterized a new stochastic process which counts how many times the original stochastic process crosses boundaries. We have discussed and calibrated three different estimation procedures for estimating Boundary-crossing Counting distributions. We have used this distribution in model validation and statistical specification testing. For unit root testing, we have shown that BCC-tests are less powerful than specific parametric unit-root tests but more powerful than other parametric and a probability integral transform based non-parametric tests. Also, the BCC-test appears to be a reasonably powerful tool for differentiating between commonly used GARCH-type models when the distribution of the error term is different. When doing statistical specification testing of major stock-indexes, the BCC-test often rejects the hypothesis that these indexes were generated by GARCH(1,1) models, while it always accepts the leveraged GARCH or GRJ specification, pointing out the importance of asymmetry in modeling financial time series.

Let us conclude this paper by putting our method into perspective. Our approach fundamentally differs from the usual approach because it takes a different perspective. The vast majority of studies analyze how certain statistics behave within a pre-defined amount of time: how daily log returns behave, how monthly volatilities behave, etc. In this paper we chose an inverse approach by fixing some pre-defined level of change and by analyzing how much time it take for trajectories to cross that pre-defined level.

So in a sense, we are analyzing time. Intuitively, looking at a phenomenon from a different angle should be enriching, yet the findings should have complementary characteristics and should not be completely against the findings of the usual perspective, which has happened in our case as well. Finally, let us remark that Boundary-crossing Counting Processes potentially have many other statistical and non-statistical applications in the domain of Finance and Management Sciences.

Bibliography

Ait-Sahalia, Y. [1996]: Testing Continuous-Time Models of the Spot Interest Rate, *Review of Financial Studies, 9, 385-426.*

Anderson, T. W. [1993]: Goodness of Fit Tests for Spectral Distributions, *The Annals of Statistics 21,* no. 2 830–847. doi:10.1214/aos/1176349153.

Bachelier, L. [1900]: Theorie de la speculation, Gauthier-Villars.

Black, F. and Scholes, M. [1973]: The pricing of options and corporate liabilities, *Journal of Political Economy,* Vol.81, pp.637-654.

Bollerslev, T. [1986]: Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics,* Vol. 31, 1986, pp. 307–327.

Borodin, Andrei N. and Salminen, P. [2002]: Handbook of Brownian Motion: Facts and Formulae, Birkhauser.

Caner, M. and Kilian, L. [2001]: Size distortions of the null hypothesis of stationarity: evidence and implications for the PPP debate, *Journal of International Money and Finance,* 20, 639-657.

Chou, Ray Y. [2005]: Modeling the Asymmetry of Stock Movements Using Price Ranges, *in Dek Terrell, Thomas B. Fomby (ed.) Econometric Analysis of Financial and Economic Time Series (Advances in Econometrics, Volume 20),* Emerald Group Publishing Limited, pp.231-257.

Cont, R. [2001]: Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues, *Quantitative Finance 1*, no. 2: 223–236.

Dickey, D.A. and Fuller, W.A. [1981]: Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, *Econometrica,* 49, 1057-1072.

Engle, R.F. [1982]: Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica,* 50, 987-1007.

Fama, Eugene F. [1991]: Efficient Capital Markets: II. *The Journal of Finance,* Volume 46, Issue 5, pages 1575–1617.

Fan, Jianqing, Zhang, Chunming and Zhang, Jian [2001]: Generalized Likelihood Ratio Statistics and Wilks Phenomenon. *The Annals of Statistics,* 29, no. 1 153–193. doi:10.1214/aos/996986505.

Fan, Jianqing [2005]: A Selective Overview of Nonparametric Methods in Financial Econometrics. *Statistical Science,* 20, no. 4. doi:10.1214/088342305000000412.

Feller, William [1968]: An Introduction to Probability Theory and Its Applications, Vol. 1. 3rd ed., Wiley

Feller, William [1971]: An Introduction to Probability Theory and Its Applications, Vol. 2. 1st ed., Wiley

French, Kenneth R. [1980]: Stock returns and the weekend effect, *Journal of Financial Economics,* 8, 55-70

Gao, Jiti, and King, Maxwell [2004]: Adaptive testing in continuous-time diffusion models. *Econometric Theory,* 20, no. 05: 844–882. doi:10.1017/S0266466604205023.

Ghalanos, Alexios [2013]: rugarch: Univariate GARCH models. R package version 1.2-2. Downloaded from http://www.inside-r.org/packages/cran/rugarch/docs/HLTest

Glosten, L. R., Jagannathan, R. and Runkle, D. E. [1993]: On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *The Journal of Finance,* Vol. 48, No. 5, 1993, pp. 1779–1801.

Gourieroux, Christian, and Jasiak, Joann [2001]: Financial Econometrics: Problems, Models, and Methods. *Princeton University Press*

Hansen, Peter R., and Lunde, Asger [2005]: A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)? *Journal of Applied Econometrics,* 20, no. 7: 873–889. doi:10.1002/jae.800.

Hong,Y. and Li, H. [2005]: Non parametric specification testing for continuous-time models with applications to term structure of interest rates*. Review of Financial Studies, 18, 37–84.*

Hong, Y. and Lee, Yoon-Jin [2005]: Generalized Spectral Tests for Conditional Mean Models in Time Series with Conditional Heteroscedasticity of Unknown Form, *Review of Economic Studies,* 72, (2): 499-541.

Jordá, Oscar [undated]: Economic Time. http://www.econ.ucdavis.edu/faculty/jorda/papers/Time.pdf , *Working paper*

Karlin, Samuel and Taylor, Howard M. [1981]: A Second Course in Stochastic Processes, 1st ed., *Academic Press*

Karlin, Samuel and Taylor, Howard M. [1998]: An Introduction to Stochastic Modeling, 3rd ed., *Academic Press*

Kirk, Donald E. [2004]: Optimal Control Theory: An Introduction, Mineola, N.Y: Dover Publications,

Kolmogorov, A. [1931]: Uber die analytischen Methoden in der Wahrscheinlichkeitsrechnung, *Math. Anal.,* 104, 415-458.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. [1992]: Testing the null hypothesis of stationarity against the alternative of a unit root, Journal of Econometrics, 54, 159-178.

Lin, S. X. [1998]: Double Barrier Hitting Time Distributions with Applications to Exotic Options, *Insurance: Mathematics and Economics,* 23, no. 1: 45–58.

Linetsky, V. [2004]: Computing hitting time densities for CIR and OU diffusions: applications to mean-reverting modes, *The Journal of Computational Finance,* 7, 4.

Lo, A.W. and MacKinlay, A.C. [1988]: Stock market prices do not follow random walks: evidence from a simple specification test, *Review of Financial Studies,* 1(1): 41-66 doi:10.1093/rfs/1.1.41.

Mandelbrot, Benoit B. [2004]: The (mis)behavior of Markets: A Fractal View of Risk, Ruin, and Reward, *Basic Books*, New York

Mcleish, Donald L. [2002]: Highs and lows: Some properties of the extremes of a diffusion and applications in finance, *The Canadian Journal of Statistics,* Volume 30, Issue 2, p. 243-267.

Medvegyev, Péter [2007]: Stochastic Integration Theory, *Oxford Graduate Texts in Mathematics*

Neuberger, Anthony [2012]: Realized Skewness, *Review of Financial Studies,* 25, (11): 3423-3455 first published online September 28, 2012 doi:10.1093/rfs/hhs101

Nelson, D. B. [1991]: Conditional heteroskedasticity in asset returns: A new approach, *Econometrica,* 59, 347-370.

Phillips, P.C.B. and Perron, P. [1988]: Testing for a Unit Root in Time Series Regression, *Biometrika*, 75, 335–346.

Phillips, P. C. B. and Xiao, Zhijie [1998]: A Primer on Unit Root Testing, *Journal of Economic Surveys*, Wiley Blackwell, vol. 12(5), pages 423-69.

Poon, S. and Granger, C. [2003]: Forecasting variance in financial markets: a review, *Journal of Economic Literature,* 41: 478–540.

Primožič, Tom [2011]: Estimating expected first passage times using multilevel Monte Carlo algorithm, New College, University of Oxford.

Pritsker, M. [1998]: Nonparametric Density Estimation and Test of Continuous Time Interest Rate Models, *Review of Financial Studies*, 11, 449–487.

Redner, Sidney [2001]: A Guide to First-Passage Processes, Cambridge University Press

Song, Zhaogang [2012]: A Martingale Approach for Testing Diffusion Models Based on Infinitesimal Operator. *Journal of Econometrics,* 162, no. 2 189–212. doi:10.1016/j.jeconom.2010.12.005.

Valenti, Davide, Bernardo, Spagnolo, and Giovanni, Bonanno [2006]: Hitting Time Distributions in Financial Markets, *Quantitative Finance Papers*, arXiv.org, no. physics/0608201

Valov, Angel V. [2009]: First Passage Times: Integral Equations, Randomization and Analytical Approximations, *Phd Thesis*, University of Toronto

Wilmott, P. [1998]: Derivatives: The Theory and Practice of Financial Engineering. Chichester, West Sussex, England; New York: Wiley

Zhao, Zhibiao [2008]: Parametric and Nonparametric Models and Methods in Financial Econometrics. *Statistics Surveys,* 2, 1–42. doi:10.1214/08-SS034.

Historical data was downloaded from finance.yahoo.com using the following tickers: Down Jones Industrial Average - DJI; Nasdaq Composite – IXIC; SP500 – GSPC; CAC40 (France) – FCHI; DAX (Germany) – GDAXI; Nikkei (Japan) - N225; FTSE (United Kingdom) – FTSE; Australia, All Ordinaries – AORD.