



Research Report
Statistical Research Unit
Göteborg University
Sweden

**Similarities and differences between
statistical surveillance and certain
decision rules in finance**

David Bock

**Research Report 2003:4
ISSN 0349-8034**

Mailing address:	Fax	Phone	Home Page:
Statistical Research Unit	Nat: 031-773 12 74	Nat: 031-773 10 00	http://www.stat.gu.se/stat
P.O. Box 660	Int: +46 31 773 12 74	Int: +46 31 773 10 00	
SE 405 30 Göteborg Sweden			

SIMILARITIES AND DIFFERENCES BETWEEN STATISTICAL SURVEILLANCE AND CERTAIN DECISION RULES IN FINANCE

By David Bock

Statistical Research Unit
Göteborg University, Sweden

ABSTRACT

Timely decisions about when to trade assets are important in financial management. In this paper we consider prospective decision rules that aim at extracting early signals about what decision to make, e.g. sell an asset. The decision rules are based on prospective monitoring of a statistic in order to detect a regime shift, e.g. a turn in the trend of asset prices. In the finance literature there are several suggested prospective decision rules that aim at detecting a turn in the price, for example the Filter rule and the rules that use moving averages. Another approach that has been proposed is the hidden Markov model (HMM) approach, where the price level is assumed to change between an upward and a downward trend according to a Markov chain. An approach not often used in a financial setting is statistical surveillance, which deals with the theory and methodology of online detection of an important change in the underlying process of a time series as soon as possible after it has occurred. In this paper inferential differences and similarities between statistical surveillance, two variants of the Filter rule, a rule based on moving averages and an HMM approach are investigated. A new non-parametric and robust approach never used in financial settings is proposed. Further, the purpose is to enhance the use of proper evaluation, where the timeliness of the alarm is considered. Evaluation measures and optimality criteria commonly used in statistical surveillance are reviewed and compared with those generally used in a financial setting. The methods are evaluated on Hang Seng Index.

Key Words: HIDDEN MARKOV MODELS, SEQUENTIAL METHODS, TURNING POINT, SURVEILLANCE, TRADING RULES

Address for correspondence: David Bock, Statistical Research Unit, Göteborg University, Box 660, SE 405 30 Göteborg, Sweden
Email: David.Bock@statistics.gu.se

1 INTRODUCTION

The aim for investors in the financial markets is to maximize a utility function that reflects the expected wealth, e.g. the expected net return on transactions. In a prospective framework, we have to make optimal sequential decisions whether to make a transaction or not. The approach considered in this paper is to monitor an indicator with the aim of detecting the optimal time to stop the monitoring and trade a certain asset, for example shares of a stock index fund or currency. The indicator is a statistic associated with the pricing process of the asset of interest. It is reasonable that optimal times to trade coincide with important events such as changes (regime shifts) in the stochastic properties of the indicator, for example changes in parameters of the model of the process. Then, finding the optimal time to stop the monitoring and trade is equivalent to the timely detection of a regime shift. Different types of indicators, regime shifts and model specifications are briefly discussed in section 2. The statistical methods discussed below are general and can detect any regime shift in any variable (the level or variance of the price, etc).

According to one point of view, namely the efficient market hypothesis, the financial markets are arbitrage-free and hence there is no point in trying to determine the optimal transaction time. According to the "First fundamental asset pricing theorem", the market is arbitrage-free if and only if the asset prices are martingales (see Shiryaev (1999), p. 413). This implies that all information about the expected price one time unit in the future is reflected in the current price. Many agents in the financial markets reject the hypothesis and argue that historical data of financial markets contain information that is valuable for trading decisions.

In this paper we work under the assumption that there exist temporary trends in the financial markets and that it is possible to identify regime shifts (turning points) where the trend changes direction by utilizing incoming data.

Many non-linear approaches have been suggested where the regime is determined by observable variables and a change has been interpreted as a change in one or several parameters of a time series model. Smooth transition autoregressive (STAR) models (see e.g. Teräsvirta (1994) and Van Dijk et al. (2002)) is an example of a family of such models. Monitoring in order to detect regime shifts is different from modeling shifts. In monitoring the estimation of parameters is not the main issue, but rather to make repeated decisions in order to determine, at each time, whether a change in an important parameter, for example the expected value, has occurred or not. The inference situation is one of surveillance where the aim is quick and safe detection of a change. Repeated decisions are made, the sample size is increasing, and the null hypothesis is never accepted. Thus, the methodology of statistical surveillance is appropriate. For general reviews on statistical surveillance, see Frisé and de Maré (1991), Wetherhill and Brown (1991), Srivastava and Wu (1993), Lai (1995), Frisé and Wessman (1999) and Frisé (2003)). Statistical process control and monitoring are other names for methods with this goal. A short review of the theory and methodology of statistical surveillance is given in section 3.

Measures used for evaluating how well a model describes data may be inappropriate for evaluating monitoring methods. The mean squared error (MSE) or the Brier probability score (also referred to as the quadratic probability score) are often used to describe the goodness of fit of a model. In a monitoring situation, the timeliness of an alarm signal, that is the relation between the time of alarm and the time of the change, is important. MSE or the Brier probability score does not take into account the order of the observations and hence do not reflect timeliness. Furthermore, Leitch and Tanner

(1991) argue that measures like the MSE is not necessarily closely related to the profitability of forecasting based decisions. Therefore, conventional measures of goodness of fit may not be appropriate in this setting.

Since timeliness is crucial in a trading setting, the data should not be analyzed in a retrospective setting but in a prospective framework where incoming data are analyzed online and sequential trading decisions are made. However, a lot of work concerning detecting regime shifts in financial time series has been made in a retrospective setting, see e.g. Wichern et al. (1976), Hsu (1982), Inclan and Tiao (1994) and Chen and Gupta (1997). However, we will not consider a retrospective setting.

Prospective decision rules used in finance in order to find optimal times to trade are often referred to as technical trading rules (or technical analysis, see e.g. Pring (1985) and Edwards and Magee (1992)). The general goal, according to Lo (2000), is to detect regularities in the time series of prices by extracting nonlinear patterns from noisy data. Whether prospective decision rules can generate profits that are higher than would be expected if the efficient market hypothesis holds has been investigated extensively. Whereas e.g. Fama and Blume (1966) and Szakmary and Davidson III (1999) conclude that profits cannot be made when transaction costs are taken into account, many studies give support for the profitability of the prospective framework, see e.g. Sweeney (1986), Sweeney (1988), Brock et al. (1992), Levich and Thomas (1993) and Neely et al. (1997). Though these studies use a prospective setting, no reference is made to the statistical surveillance framework. The decision rules (or trading rules) discussed in this paper are those which include an inferential aspect regarding a change in the process.

Neftci (1991) points out that many of the decision rules referred to as technical trading rules are *ad hoc* and the statistical properties of the methods are often unknown. A consequence of this, as argued by Dewachter (1997) and Dewachter (2001), is that the statistical source of any possible profitability is not identified. A methodology that is perfectly suited for the prospective decision making situation in trading is statistical surveillance. Since the properties of methods of statistical surveillance have been investigated extensively, an integration of the theory and methodology of statistical surveillance and financial decision rules may be fruitful. Statistical surveillance has, however, not been used a lot, perhaps as a result of the cultural differences, e.g. linguistic barriers between financial decision rules and statistical surveillance. The latter has mostly been considered in medical and quality control settings.

Lam and Yam (1997) claim to be the first to attempt to link up methods of statistical surveillance with those of technical trading rules. Previous to Lam and Yam (1997), Theodossiou (1993) applied a method of statistical surveillance for detecting business failures. Recent studies that relate statistical surveillance to financial applications are Severin and Schmid (1999), Schipper and Schmid (2001a), Schipper and Schmid (2001b), and Steland (2002). These papers will be discussed in section 4.1.

The purpose of this paper is to investigate the inferential differences and similarities between some methods of statistical surveillance and some prospective decision rules used in finance. The decision rules used in finance are here represented by two different variants of the Filter Rule suggested by Alexander (1961) and Lam and Yam (1997), two rules based on moving averages and a rule that use a hidden Markov model (HMM). Furthermore, the purpose is to enhance the use of proper evaluation. Evaluation measures and utility functions commonly used in statistical surveillance are reviewed and compared with those generally used in financial settings.

The performance of decision rules can be evaluated in different ways, for example by means of different case studies or by a simulation study. A common feature in the finance literature on evaluation of decision rules is the use of different case studies. In

this paper evaluation is made by one set of real data, the Hang Seng Index. The pros and cons of evaluation by case studies as compared with evaluations by the stochastic properties of a model are discussed.

The plan of this paper is as follows. Examples of indicators, regime-shifts, models and events to be detected are discussed in section 2 whereas section 3 briefly describes the theory and methodology of statistical surveillance. In section 4 some of the recent studies that relate statistical surveillance to financial applications are described and similarities and differences between some methods proposed in the literature on finance and some methods of statistical surveillance are discussed. Some of the methods are compared in section 5 by means of a case study. Concluding remarks are given in section 6.

2 REGIME SHIFTS, INDICATORS AND MODEL SPECIFICATION

Examples of different indicators and regime shifts to be detected are given in section 2.1. However, for all types of indicators and regime shifts the problem can be treated in a general framework. But in order to make the discussion more concrete we will limit the discussion to a specific problem. We will consider the problem of timely detection of a turning point in the trend of a univariate cyclical process. This type of regime shift is often considered in the literature on financial decision rules due to the assumption of trends in the financial markets. The specification of models when the regime shift is a turning point is made in section 2.2 and the specification of events to be detected is made in section 2.3.

2.1 *Examples of regime shifts and indicators*

The stochastic process under surveillance is univariate or multivariate. The indicator can be constructed from one or several processes that are leading with respect to the variable of interest, e.g. the asset price level. The indicator can e.g. be calculated from prices of other assets. For trading based on indicators that are believed to lead, see e.g. Boehm and Moore (1991), Moore et al. (1994) and Brooks et al. (2001). In Boehm and Moore (1991) and Moore et al. (1994), a leading index is monitored in order to extract early signals of turning points in order to trade stock indices.

When several processes are monitored, there are two obvious ways to simplify the multivariate situation. One is to reduce the processes into one process by some summary statistic. Wessman (1998) demonstrates that the minimal sufficient statistic for detecting regime shifts in several variables with the same change point (or known time-lag) is univariate. The other approach is to monitor each process separately and then combine the information. For reviews on multivariate surveillance, see Wessman (1999) and Frisén (2003).

Risk is important in financial analysis, e.g. in portfolio management and in option pricing. A sudden increase in the risk level might hence be the regime shift to detect. The estimated variance can be interpreted as a measure of volatility, which itself is a measure of risk. In e.g. Schipper and Schmid (2001b) the monitoring of the variance is considered in a financial context. In the literature on surveillance of the variance, standard methods, usually used for monitoring the process mean, are applied on different transformations of the variance (Frisén (2003)). Two examples are the

standard deviation (von Collani and Sheil (1989)) and the logarithmic of the variance (Chang and Gan (1995)). It has been found that a large negative change in the price is often followed by an increased variance (see Franses and van Dijk (2000), p. 16). This phenomenon is called the Leverage effect (Black (1976)). In Schipper and Schmid (2001a) the aim is to simultaneously detect an additive outlier and a changed variance.

According to market microstructure theory, see e.g. O'Hara (1995), the time between events associated with assets reveal information about the state of the market. Thus, the arrival time or the counting process of transactions could be an indicator to be monitored in order to detect a change in the intensity parameter. In Vardeman and Ray (1985) and Gan (1998) exponentially distributed arrival times are under surveillance. However, a realistic model of the arrival times of transactions is complicated and an exponential distribution might be too simple, see e.g. Engle and Russell (1998) and Zhang et al. (2001). In Rydberg and Shepard (1999) a model of the counting process of transactions is proposed. A Poisson distribution is used where the intensity has a structure similar to a generalized autoregressive conditional heteroscedasticity (GARCH) process. In Sonesson and Bock (2003) an optimal surveillance method for the arrival time and the counting process of an adverse health event is derived under the assumption of a Poisson process that is homogeneously conditional on the states.

Associated with each transaction, marks are often available and are said to be related to the intensity (see Easley and O'Hara (1992) and Dufour and Engle (2000)). Examples of marks are the volume that is being traded, the level of the ask, bid and settling price and which traders that took part in the trade. A monitoring system that incorporates both the arrival times (or the counting process) and several marks might be treated in the context of multivariate surveillance.

Another example of an indicator is the residuals of a time series model where the change is in the stochastic properties of the residuals. Yet another example is a turning point in a cyclical process to be described below.

2.2 Models used when the regime shift is a turning point

A turning point, i.e. a change in the monotonicity in a cyclical process is the regime shift considered for exemplification here. Examples of such processes are the price level of the asset of interest or the price level discounted by the opportunity cost rate that is the return you could earn on alternative investments of similar risk. The indicator that is being monitored could be a process Y , e.g. the price of an asset, or some transformation, e.g. the logarithmic transform or the first difference. For simplicity the process under surveillance is generally denoted X , where X can be achieved by a transformation ($\ln Y$, $Y(t)-Y(t-1)$ or $\ln Y(t)-\ln Y(t-1)$). The situation under study is one where we assume a linear model for X ,

$$X(t)=\mu(t)+\varepsilon(t) \quad (1)$$

where $\mu(t)$ is the trend cycle and $\varepsilon(t)\sim\text{iid } N[0, \sigma^2]$, $t=1, 2, \dots$

The assumption in (1) is in general too simple for financial data and extensions could be motivated by the features of the data or economic theory. Model (1), however, is used here to emphasize the inferential issues. The problem of autocorrelation is briefly discussed in section 5.3.3. Additional problems of seasonality, multivariate processes and additional trends in a setting of surveillance is discussed in Andersson et al. (2002a). Time series of asset prices are known to be severely heteroscedastic, which

should be taken into account by assigning the available observations at time t , $\{x(1), x(2), \dots, x(t)\}$, different weights in the analysis.

The aim is to detect a change from an upward trend to a downward trend, i.e. a turn. When $X(t)$ is an undifferentiated series, then the aim is to detect a turn in μ . The examples and discussion henceforward will be for a peak (the best time to sell an asset), but the results also hold for detecting a trough (the best time to buy an asset). For the vicinity of a peak, μ is monotonic within each regime. That is

$$E[X_t]=\mu_t: \begin{cases} \mu^D: \mu(1) \leq \dots \leq \mu(t), & t < \tau \\ \mu^{Cr}: \mu(1) \leq \dots \leq \mu(\tau-1) \text{ and } \mu(\tau-1) \geq \dots \geq \mu(t), & t \geq \tau \end{cases} \quad (2)$$

where $X_t=\{X(1), \dots, X(t)\}$ and τ is the unknown time of the change. With a random time of occurrence, μ is a stochastic process.

Both nonparametric and parametric specifications will be considered. The parametric specification of μ is a piecewise linear regression

$$E[X(t)]=\mu(t): \begin{cases} \mu^D(t): \beta_0 + \beta_1 \cdot t, & t < \tau \\ \mu^{Cr}(t): \beta_0 + \beta_1 \cdot (\tau-1) + \beta_2 \cdot (t-\tau+1), & t \geq \tau \end{cases} \quad (3)$$

where $\beta_1 \geq 0$, $\beta_2 < 0$ when the turning point is a peak. The expected value in (3) holds also when $X(t)$ can be described as a random walk with drift, where the value of the drift parameter changes from β_1 to β_2 at time $t=\tau$.

When $X(t)$ is a differentiated process, then the monitoring can be made by constructing a system for detecting a shift in the level of μ . For example, if the undifferentiated process has the expected value in (3), then the expected value of X is constant conditional on the state, i.e.

$$E[X(t)] = \begin{cases} \mu^D(t): \beta_1, & t < \tau \\ \mu^{Cr}(t): \beta_2, & t \geq \tau \end{cases} \quad (4)$$

where $\beta_1 \geq 0$ and $\beta_2 < 0$. This assumption about the process is made in some HMM approaches, see e.g. Layton (1996), Ivanova et al. (2000) and Layton and Katsuura (2001). When the observations are independent over time, as in (1), the expected values of the differentiated series in (4) imply the linear functions in (3) for the undifferentiated series. If the undifferentiated process is a random walk with drift, or another process with an expected value as in (3), then (4) is valid for the differentiated series. Transformations can also affect the dependency structure. The process X is independent if the undifferentiated process is a random walk with drift. However, if the undifferentiated process is independent as in (1), then the process X will not be independent. Instead, X will be a moving average process of order one with parameter $\theta=1$, so the stochastic part of X is

$$\varepsilon(t) = \omega(t) - \theta \cdot \omega(t-1).$$

The turn may not be abrupt but smooth or gradual. Economic agents may not all act promptly and uniformly at the same moment; their response to news requiring action may contain delays (Teräsvirta (1998)). For literature on surveillance of a gradual change, see Gan (1992), Sveréus (1995) and Chang and Fricker (1999).

2.3 Event to be detected

For each decision time s , a decision is made whether data indicate that a regime shift has occurred or not, i.e. whether τ is in the future or not. Formally in the methodology of statistical surveillance, this is expressed as discriminating between the in- and out-of-control events, expressed as $C(s)$ and $D(s)$, respectively. $C(s)$ and $D(s)$ can be specified in different ways. The most frequently studied case is when $C(s) = \{\bigcup_{t=1}^s C_t\} = \{\tau \leq s\}$, where $C_t = \{\tau = t\}$, i.e. that the change has occurred and the complement, and $D(s) = \{\tau > s\}$. When the change is a peak and μ is given as monotonic functions (see (2)), then $C(s)$ and $D(s)$ are defined as

$$D(s): \mu(1) \leq \dots \leq \mu(s)$$

$$C(s): \mu(1) \leq \dots \leq \mu(\tau-1) \text{ and } \mu(\tau-1) \geq \mu(\tau) \geq \dots \geq \mu(s)$$

where $\tau = \{1, 2, \dots, s\}$ is the time of the change and at least one inequality is strict in the second part. Under the parametric assumption in (3), we specify C and D as

$$D(s): \mu(t) = \beta_0 + \beta_1 \cdot t, t = \{1, 2, \dots, s\}$$

$$C(s) = \{\bigcup_{\tau=1}^s C(\tau)\},$$

where $C(\tau): \mu(t) = \beta_0 + \beta_1 \cdot (\tau-1) + \beta_2 \cdot (t-\tau+1)$, $t = \{1, 2, \dots, s\}$.

One approach to the monitoring is to stop as soon as there is an alarm and then start the monitoring from scratch in order to detect the forthcoming turn. Thus, the monitoring is only made for the next peak, and knowledge of the type of the next turn is assumed. Only data from the current phase (upward or downward trend) is used in the monitoring. This will be contrasted to methods presented in section 4.4, where the inference situation can be described as one of classification of the observations to states. An alarm is called as soon as the system indicates a transition between states. In the classification situation we hence want to detect all possible transitions. The specification of C and D for this approach is

$$D(s): \mu(s-1) \leq \mu(s) \tag{5}$$

$$C(s): \mu(s-1) > \mu(s).$$

When the system indicates a transition between the states, a decision to trade might be taken but thereafter the monitoring is not restarted but continued. This means that less information is used (nothing about the type of turn). Inferential differences and similarities between the two approaches were investigated in Andersson et al. (2002b).

3 A SHORT REVIEW OF THE THEORY OF STATISTICAL SURVEILLANCE

In the theory of statistical surveillance a method for change-point detection is developed from an optimality criterion. This means that the "best" alarm system is developed, conditional on the type of change and the cost of different actions. An optimal alarm set $A(s)$ with the property that, as soon as X_s belongs to $A(s)$ we infer that C has occurred, is constructed. Usually this is done by using an alarm system, consisting of a function $p(x_s)$, and a limit $g(s)$, where the time of an alarm, t_A , is defined as $t_A = \min\{s: p(x_s) > g(s)\}$.

3.1 Optimality and measures of evaluation

It is important that the alarm system utilizes information in an optimal way. Desired properties of the rule are expressed as optimality criteria, which can be formulated in terms of utility functions. For a situation with one decision (hypothesis testing) the probabilities of type I and type II errors are measured by the size and power, respectively, i.e. there is a trade-off between these two probabilities. In a monitoring situation with repeated decisions, the time aspect (the time of the signal, t_A , in relation to the time of the change τ) is important and must be considered in the evaluation. Here we have a trade-off between false alarms and delay of motivated alarms. The specification of utility determines the severity of the penalty for false or delayed alarms. The purpose of this section is to present a utility function and different measures used to evaluate decision rules.

In statistical surveillance the type I error is usually characterized by the average run length conditional of no change, $ARL^0 = E[t_A | \tau = \infty]$, the median run length conditional of no change, $MRL^0 = \text{Median}[t_A | \tau = \infty]$ or the probability of a false alarm (PFA),

$$P(t_A < \tau) = \sum_{i=1}^{\infty} P(\tau = i) \cdot P(t_A < i | \tau = i).$$

The evaluation of a method is made conditional on a limit $g(s)$ that yields a certain ARL^0 (or MRL^0 or PFA).

The ability to detect a change within d time units from τ is reflected by the probability of successful detection

$$PSD(d, t) = P(t_A - \tau \leq d | t_A \geq \tau, \tau = t), \quad (6)$$

which was used by e.g. Frisé (1992) and Frisé and Wessman (1999). $PSD(d, t)$ is calculated conditional on the alarm coming after the turn. However, as a warning, signals right before the peak are also useful. The detection probability around the peak can be evaluated by

$$P(|t_A - \tau| < d),$$

where d is a constant integer. Bojdecki (1979) gives the solution to a maximization of $P(|t_A - \tau| < d)$ with respect to τ . In his solution only the d latest observations are involved (Frisé and de Maré (1991)). When the process under surveillance is independent and normally distributed, the moving average method of statistical surveillance (see section 3.2) with a window width d can be shown to be a special case of the solution of Bojdecki (1979) (Frisé (1994)). Another important aspect in the evaluation is the value of an alarm at a certain time point. The predictive value of an alarm at time t , $PV(t) = P(\tau \leq t | t_A = t)$, suggested by Frisé (1992), reflects the trust you should have in an alarm at a given time point. If the predictive value is low, the alarm system is not reliable.

The “power” of a method of surveillance is measured through measures that reflect the timeliness. A desirable property would be a short delay between the time of a change and the time of an alarm. A widely used optimality criteria in the literature on quality control is the minimal $ARL^1 = E[t_A | \tau = 1]$ for a fixed ARL^0 . This criterion only considers changes that occur at the start of the monitoring ($\tau = 1$), which is not realistic. The evaluation should also take into account $\tau > 1$ because the performance of methods with the same ARL^0 may depend on the value of τ (see Frisé and Sonesson (2002)).

In the specification of utility by Shiryaev (1963) the gain of an alarm is a linear function of the expected delay. The loss associated with a false alarm is a function of the same difference. The utility is

$$u(t_A, \tau) = \begin{cases} h(t_A - \tau) & , t_A < \tau \\ a_1 \cdot (t_A - \tau) + a_2, & t_A \geq \tau \end{cases} \quad (7)$$

where the function $h(t_A - \tau)$ is an arbitrary function. In a situation where the intensity of a change, $P(\tau=t|\tau \geq t)$, is a constant v , the likelihood ratio method (LR, see section 3.2) maximizes the expected value of the utility. This criterion is sometimes referred to as the expected delay (ED) criterion for the following reason. If $h(t_A - \tau)$ is a constant b , then the expected utility is

$$E[u(t_A, \tau)] = b \cdot P(t_A < \tau) + a_1 \cdot ED + a_2. \quad (8)$$

The ED is defined as

$$ED = \sum_{i=1}^{\infty} P(\tau=i) \cdot ED(i)$$

where $ED(i) = E[\max(0, t_A - i) | \tau=i] = E(t_A - \tau | t_A \geq \tau=i) \cdot P(t_A \geq i)$. The expression $E(t_A - \tau | t_A \geq \tau=i)$ is the conditional expected delay of an alarm when $\tau=i$ (CED(i)). When $P(t_A < \tau)$ is fixed, the expected utility is maximized for a minimal ED.

Minimax criteria can be specified in different ways. One specification is the minimum of the maximal value of CED(i) with respect to $\tau=i$. The minimax criterion avoids the requirement of a known distribution of τ . The criterion minimal maximal CED(i) with respect to $\tau=i$ and the worst possible outcome of $Y_{\tau-1}$ for a fixed ARL^0 is used by Moustakides (1986) and it is shown that the CUSUM method (see section 3.2) satisfies the criterion.

Another optimality criterion is the maximal detection probability $P(A(s)|C(s))$ for a fixed false alarm probability $P(A(s)|D(s))$ and a fixed decision time s , see Frisén and de Maré (1991) and Frisén (2003). When $C(s) = \{\tau \leq s\}$ and $D(s) = \{\tau > s\}$, this criterion is satisfied by the LR method. When $C(s) = \{\tau = s - m + 1\}$ and $D(s) = \{\tau > s\}$, the moving average method of statistical surveillance with a window width m is optimal (Frisén (2003)).

3.2 Methods of surveillance

The methods reviewed in this section are generally presented as monitoring of the process X , where X can be transformed or untransformed. Since the general aim is to discriminate between event D and event C , these events are specified either in terms of the transformed or the untransformed process of μ (see e.g. $\mu^D(t)$ and $\mu^{C^r}(t)$ in equation (3) and (4)). All methods reviewed are based on likelihood ratios where the difference depends on how the partial likelihood ratios are weighted. The likelihood ratio (LR) method does fulfill several of the optimality criteria described in section 3.1, and has been used as a “benchmark” when evaluating different methods of surveillance, see e.g. Frisén and Sonesson (2002) and Frisén (2003). The alarm rule of the LR method is

$$\frac{f_{X_s}(x_s | C(s))}{f_{X_s}(x_s | D(s))} = \sum_{i=1}^s w(i) \cdot L(s, i) > g_{LR}(s) \quad (9)$$

where

$$L(s, i) = f_{X_s}(x_s | \mu = \mu^{C_i}) / f_{X_s}(x_s | \mu = \mu^D)$$

is the partial likelihood ratio of X_s when $\tau=i$ and $w(i) = P(\tau=i) / P(\tau \leq s)$ is the weight for $L(s, i)$. In $L(s, i)$, μ^{C_i} and μ^D are the expected value of X conditional on $\tau=i$ and $\tau > s$, respectively. It was shown by Frisé and de Maré (1991) that the alarm rule of the LR method can be expressed in terms of the posterior probability conditional on the observed values of X and a constant limit g_{PP} . An alarm is given as soon as $P(C(s) | X_s) > g_{PP}$. This is equivalent to the LR method with the limit $g_{LR}(s) = g_{PP} / (1 - g_{PP}) \cdot P(D(s)) / P(C(s)) = g_{PP} / (1 - g_{PP}) \cdot P(\tau > s) / P(\tau \leq s)$.

The alarm rule of the LR method in (9) requires knowledge of the distribution of τ . Often a geometric distribution with a constant intensity $v = P(\tau=t | \tau \geq t)$ is assumed. When the distribution of τ is unknown, using a non-informative prior for the time of the change point avoids the risk of serious misspecification. When the intensity v tends to zero in the case of a geometric distribution, both $w(t)$ and $g(s)$ tend to constants. Thus, the limiting distribution is the non-informative prior.

The method based on the limiting distribution ($v \rightarrow 0$) is the Shiryayev-Roberts (SR) method (Shiryayev (1963) and Roberts (1966)). The SR method satisfies the expected delay criterion when the process under surveillance has a constant intensity that tends to zero. The SR method gives equal weight to all partial likelihood ratios and the method can be used as an approximation to the LR method. Frisé and Wessman (1999) showed that the approximation works well, even for as large intensities as $v=0.20$. One method that uses the SR approach and is designed for turning point detection by monitoring the undifferentiated process, is the SRLin method (Andersson et al. (2002a) and Andersson et al. (2002b)). This method specifies the trend cycle μ as piecewise linear functions, as in (3) where the turn is symmetric ($\beta_2 = -\beta_1$).

Erroneous or uncertain assumptions regarding the process under surveillance may have a great impact on the performance of the surveillance method. Some methods, for example SRLin, use the assumption of a known parametric function for the trend cycle μ . However, since the parametric function is seldom known, Frisé (1994) suggested a non-parametric approach where the assumptions regarding μ are only the ones of monotonicity and unimodality in (2). Combined with a non-informative prior for τ , this is the SRnp method that has the following alarm rule for detecting a turn in μ :

$$\frac{\sum_{i=1}^s f_{X_s}(x_s | \mu = \hat{\mu}^{C_i})}{f_{X_s}(x_s | \mu = \hat{\mu}^D)} > g_{SRnp}, \quad (10)$$

where g_{SRnp} is a constant. The vector $\hat{\mu}^D$ is the estimator of μ under the monotonicity restriction D (that no turn has occurred, i.e. that $\mu(1) \leq \dots \leq \mu(s)$). The vector $\hat{\mu}^{C_i}$ is the estimator of μ under the restriction C_i (a turn at time i , i.e. that $\mu(1) \leq \dots \leq \mu(i-1)$ and $\mu(i-1) \geq \mu(i) \geq \dots \geq \mu(s)$). The estimators are a maximum likelihood when the disturbances has a Normal distribution as in (1), see Frisé (1986) and Robertson et al. (1988). SRnp was evaluated in Andersson (2001), Andersson (2002), Andersson et al. (2002a) and Andersson et al. (2002b).

Methods for turning point detection can be applied to the undifferentiated process. This is how the LR, SRLin and SRnp methods above are presented. Alternatively, a method for turning point detection can be applied to differentiated data, in which case a turn corresponds to a shift in level (as discussed in section 2.2). The surveillance

methods below are described for the situation when the aim is to detect a shift in level, from positive to negative level.

The CUSUM method of Page (1954) is based on likelihood ratios and signals that a change has occurred in μ as soon as the maximum of the partial likelihood ratios of X_s when $\tau=i$, $L(s, i)$, exceeds a limit g ,

$$\max_{1 \leq i \leq s} \{L(s, i)\} > g,$$

where g is a constant to be determined. For independent and normally distributed variables, the alarm rule can be expressed as

$$C_s - C_{s-i} < g_{\text{CUSUM}} - k \cdot i, \quad (11)$$

for some $i=1, 2, \dots, s$. In (11), $C_s = \sum_{j=1}^s (x(j) - \mu^D)$ is a cumulative sum of the deviations between a reference value, here μ^D , and the observed values, $k = -(\mu^C - \mu^D)/2$ and g_{CUSUM} is a chosen constant. Hence, for this specification, (11) will satisfy the minimax criterion of Moustakides (1986). The size of the shift, for which you optimize the method, is reflected by k . If the size of the shift is small, the value of k will be close to zero. When the size of the shift is large, k is large.

The CUSUM method is different from CUSUM-based tests in the econometric literature (see e.g. Brown et al. (1975)). Econometric CUSUM-based tests are often used for a fixed length series and aimed at a single decision as opposed to the inference situation of surveillance where repeated decisions are made.

The moving average method of statistical surveillance is described in e.g. Wetherhill and Brown (1991) and Frisén (2003). The method can be formulated in terms of the partial likelihood ratio for $\tau=s-m$, $L(s, s-m) = f_{X_s}(x_s | \mu = \mu^{C(s-m)}) / f_{X_s}(x_s | \mu = \mu^D)$, and a constant limit g . An alarm is triggered when $L(s, s-m) > g$. The value of g is not necessarily the same value as for other methods. The method satisfies certain specifications of optimality as described in section 3.1.

For detecting a downward change in the level of an independent and normally distributed process, the alarm rule can be written as

$$\sum_{i=s-m+1}^s (x(i) - \mu^D) < g_{\text{MASUR}}, \quad (12)$$

where m is the window width and g_{MASUR} is a constant. When μ^D is unknown, we use an estimate $\hat{\mu}^D$ instead. One approach is to estimate μ^D from past observation by e.g. a moving average of window width q . At each decision time s , μ^D is estimated by

$$\hat{\mu}^D(s) = \left(\sum_{j=s-m-q+1}^{s-m} x(j) \right) / q. \quad (13)$$

The Shewhart method put all weight to the last partial likelihood ratio $L(s, s)$ and signals an alarm as soon as $L(s, s)$ exceeds a constant g . The method maximize the expected value of the utility $u(t_A, \tau)$ in (7) when $C(s) = \{\tau=s\}$ and $D(s) = \{\tau>s\}$. For an independent and normally distributed variable, the method will signal that a peak has occurred as soon as

$$x(s) - \mu^D < g_{\text{Shewhart}}, \quad (14)$$

where g_{Shewhart} is a constant. The problem of an unknown μ^D can be treated in the same manner as for the moving average method; we can e.g. use the moving average estimator (13). The method has a geometrical false alarm distribution and an conditional expected delay time, $\text{CED}(i)$, that is constant over i . The Shewhart method

was evaluated along with the LR, SR and CUSUM methods in Frisé and Wessman (1999). They showed that when the size of the shift for which the methods are optimized, tends to infinity, then the LR, SR and CUSUM methods all tend to the properties of the Shewhart method.

If the process is very noisy, we might want to monitor the smoothed process. The EWMA method (Roberts (1959), Robinson and Ho (1978), Crowder (1987), Sonesson (2001) and Frisé and Sonesson (2002)) has a built in smoothing mechanism and calls an alarm as soon as

$$Z_s = (1-\lambda) \cdot Z_{s-1} + \lambda \cdot x(s) < g_{EWMA}(s),$$

where $\lambda \in (0,1]$ is the weight parameter and the limit $g_{EWMA}(s)$ can be constant or varying, depending on the variant of EWMA. All past observations are utilized and the most recent ones are given most weight. EWMA is not an optimal method, but approximate optimality can be reached by making modifications of EWMA that behave similar to the LR method, see Frisé and Sonesson (2002) and Frisé (2003).

4 THE RELATION BETWEEN STATISTICAL SURVEILLANCE AND SOME STRATEGIES SUGGESTED FOR FINANCIAL TRADING DECISIONS

As earlier mentioned, there is a wide range of suggested methods for deciding what time is the best to stop the monitoring and trade in the financial market. Sometimes the expression “optimal stopping rules” refer to the approach described by e.g. Shiryaev et al. (1994), Shiryaev (1999), Kukush and Silvestrov (2000) and Jönsson (2001). They treat the case of finding the optimal time to exercise an option where the pricing process of the underlying asset is well-defined with known parameters. In these papers the model for the process is thus completely known, and the optimal stopping domains are obtained by calculating expected outcomes for different payoff-functions of the option. Hence, their approach is different from the one considered here where the process includes unknown statistical parameters and an inferential approach is considered where we want to infer from data whether a regime shift has occurred or not. Their approach will not be further discussed.

We will begin this section by discussing some of the recent studies that relate the theory and methodology of statistical surveillance to decision making problems in finance. Some of the earlier suggested financial decision rules are then investigated more deeply and compared with approaches of statistical surveillance. We will proceed with the work of Lam and Yam (1997) and investigate the relation between the Filter rule, their proposed generalization of the Filter rule and the CUSUM method in section 4.2. How trading rules based on moving averages are related to the optimality criteria used in surveillance has to our knowledge not been investigated. This will be the topic in section 4.3. Another methodology, also considered in finance, is that of a hidden Markov model (HMM), also referred to as a Markov-switching or regime switching model. The suggested decision rule, based on an HMM is hereafter referred to as the Hidden Markov Rule (HMR) and will be discussed in section 4.4.

The relation between measures of performance often used in statistical surveillance and measures of evaluation of return is discussed in section 4.5. All decision rules considered are presented for the case when we want to detect one turn at a time (the monitoring is restarted after a confirmed turn).

4.1 Recent studies that relate statistical surveillance to decision making problems in finance

Previous to Lam and Yam (1997), Theodossiou (1993), among others, use the methodology of statistical surveillance on financial problems. The application is not one of trading decisions but detection business failures. The CUSUM method is used in order to timely detect business failures. A multivariate process that measures a firm's financial condition is reduced into one process that is being monitored.

Recent studies are those of Severin and Schmid (1999), Schipper and Schmid (2001a) and Schipper and Schmid (2001b). In these papers, the performance of different versions of the CUSUM, EWMA and Shewhart methods are compared with respect to detecting changes in GARCH processes. GARCH processes are sometimes used to describe volatility. In Severin and Schmid (1999), these three methods are compared with respect to a change in the mean of an ARCH(1) process. A case study is made on stock market data. Whereas Schipper and Schmid (2001b) aim to detect a change in the variance of a GARCH process, the aim in Schipper and Schmid (2001a) is to simultaneously detect an additive outlier, modelled as a one time change in the mean, and a changed variance. Schipper and Schmid (2001b) use the CUSUM and EWMA methods on the following indicators: the squared observations, the logarithm of the squared observations, the conditional variance and the residuals of a GARCH model estimated conditional that no change has occurred. A case study is made on stock market data. Schipper and Schmid (2001a) use EWMA to monitor the level of the process simultaneously with the indicators above. In all the three studies, ARL^1 is evaluated for a fixed ARL^0 .

Steland (2002) addresses the need in finance for detecting change-points online. The use of a Shewhart-type method for detecting a change in the drift of a finance related process is discussed. The event to be detected is a temporary change in the drift μ (a jump away from μ^D and then back to μ^D). Both a GARCH and an independent normally distributed process are considered. The indicator under surveillance is a non-parametric kernel estimator of μ . Different types of estimators are discussed, e.g. EWMA. Also here is ARL^0 fixed and the evaluation is made with respect to ARL^1 .

4.2 Two Filter rules

Lam and Yam (1997) propose a generalized Filter rule (GFR) for monitoring the untransformed process Y . At decision time s , an alarm signal that a peak has occurred is given if

$$(y(s)-y(s-i))/y(s-i) < g_{GFR} - k_{GFR} \cdot i \quad (15)$$

for some $i=1, 2, \dots, s$ and where g_{GFR} and k_{GFR} are chosen constants. Detecting a trough (time to buy) is solved in an analogous way. GFR is derived with the CUSUM method as the starting-point. To see the relation between (15) and the CUSUM method, first consider the monitoring of an independent and normally distributed process X in order to detect a downward shift in the level from μ^D to μ^C . The alarm rule of the CUSUM method can then be expressed as (11), that is

$$C_s - C_{s-i} < g_{CUSUM} - k \cdot i,$$

for some $i=1, 2, \dots, s$, where $C_s = \sum_{j=1}^s (x(j)-\mu^D)$. This means that the alarm rule can be written as

$$\sum_{j=s-i+1}^s (x(j)-\mu^D) < g_{CUSUM} - k \cdot i,$$

which can be written as

$$\sum_{j=s-i+1}^s x(j) < g_{CUSUM} - k \cdot i + \mu^D \cdot i.$$

If we let $X(t)=\ln Y(t)-\ln Y(t-1)$, then we have the alarm rule

$$\sum_{j=s-i+1}^s (\ln y(j)-\ln y(j-1)) < g_{CUSUM} - k \cdot i + \mu^D \cdot i,$$

that is

$$\ln y(s)-\ln y(s-i) < g_{CUSUM} - (k-\mu^D) \cdot i$$

for the undifferentiated process $\ln Y$ or

$$y(s)/y(s-i) < \exp\{g_{CUSUM} - (k-\mu^D) \cdot i\}$$

for the process Y . This is equivalent to

$$(y(s)-y(s-i))/y(s-i) < \exp\{g_{CUSUM}\} \cdot \exp\{-(k-\mu^D) \cdot i\} - 1. \quad (16)$$

Let $\exp\{g_{CUSUM}\} = (1+g_{GFR})$ and $\exp\{-(k-\mu^D) \cdot i\} = (1+k_{GFR})^{-i}$. We approximate $(1+k_{GFR})^{-i}$ by $(1-k_{GFR} \cdot i)$ and make this substitution in the limit in (16). The rule (16) then becomes

$$(y(s)-y(s-i))/y(s-i) < (1+g_{GFR}) \cdot (1-k_{GFR} \cdot i) - 1.$$

If $g_{GFR} \cdot k_{GFR} \cdot i \approx 0$, this approximately equals (15), i.e. the alarm rule of GFR. Thus, applying GFR on the variable Y where $g_{GFR} = \exp\{g_{CUSUM}\} - 1$ and $k_{GFR} = \exp\{k-\mu^D\} - 1$ is approximately the same as applying the CUSUM method on the variable $\ln Y(t) - \ln Y(t-1)$. It should be pointed out however, that only when the process X is independent and normally distributed, GFR will be approximately minimax optimal.

A special case of GFR is obtained when $k_{GFR} = 0$. This is the widely used Filter rule (FR) proposed by Alexander (1961), which has been discussed by e.g. Alexander (1964), Fama and Blume (1966) and Taylor (1986) and Lam and Yam (1997). The rule is also referred to as the Trading range break (Neely (1997)). This alarm rule, as given by Lam and Yam (1997), is to alarm that a peak has occurred as soon as

$$\left(\max_{t \leq s} \{y(t)\} - y(s) \right) / \max_{t \leq s} \{y(t)\} > g_{FR} \quad (17)$$

where g_{FR} is a constant. The constant g_{FR} is generally between 0.005 percent and 0.03 (Neely (1997)). The statistical properties of this method seem not to have been examined. Sometimes the maximum in a moving window is considered instead of the overall max, see Neely (1997).

Lam and Yam showed that the FR used on $Y(t)$, is equivalent to a special case of the CUSUM rule (11) used on $\ln Y(t) - \ln Y(t-1)$. The special case is when $k = \mu^D$. It is easily seen that (17) is equivalent to $\ln y(s) - \max_{t \leq s} \{\ln y(t)\} < \ln(1-g_{FR})$. Then we have that

$$\ln y(s) - \max_{t \leq s} \{\ln y(t)\} \Leftrightarrow C_s - C_{s-i} + \mu^D \cdot i$$

for some $i \leq s$ and where $C_s = \sum_{j=1}^s (x(j) - \mu^D)$. Thus, FR in (17) is equivalent to

$$C_s - C_{s-i} < \ln(1 - g_{FR}) - \mu^D \cdot i.$$

Hence, when X is independent and normally distributed and $-(\mu^C - \mu^D)/2 = \mu^D$, FR has the same properties as the CUSUM method and FR will be minimax optimal. The special case when $-(\mu^C - \mu^D)/2 = \mu^D$ implies $\mu^C = -\mu^D$. When μ is on the form (4) for the differentiated process X , it means that $\beta_2 = -\beta_1$, i.e. the location of β_1 and β_2 are symmetric with respect to the zero line. This would be the case when undifferentiated process $\ln Y$ is a random walk with drift where the value of the drift parameter changes from β_1 to $-\beta_1$. The effect on the performance of the CUSUM method when the size of the shift for which the method is optimized, here measured by $-(\mu^C - \mu^D)$, is seriously misspecified is investigated in Frisén and Wessman (1999).

As was mentioned in section 2.2, if the undifferentiated process is independent, then the first difference X will be an MA(1) process. The partial likelihood ratio $L(s, i)$ for a change in the mean of a MA process where the disturbance term has a Normal distribution is derived in Petzold et al. (2003). The problem of dependent observations is further discussed in section 5.3.3.

The performance of GFR and FR were evaluated in Lam and Yam using one set of real data, namely the Hang Seng Index for the period 24 November 1969 to 6 January 1993, for different combinations of g_{CUSUM} and k . It was found that GFR offered an improvement over the FR with respect to measures of the return for some combinations. However, no discussion was made regarding the relation between k and the size of the shift.

4.3 Moving average rules

There are many rules suggested for financial trading decisions based on moving averages. The moving average rule as given by e.g. Neftci (1991), Brock et al. (1992), Levich and Thomas (1993), Mills (1997) and Neely (1997) will call an alarm for a peak (sell signal) as soon as the difference between two non-centred overlapping moving averages, one with a narrow window width m and one with a wide window width n , is below a limit, such that

$$\frac{1}{m} \cdot \sum_{i=s-m+1}^s x(i) - \frac{1}{n} \cdot \sum_{i=s-n+1}^s x(i) < g(s)_{MAR}. \quad (18)$$

The limit is often a constant, usually set to zero. Alarm rule (18) can be written as

$$\sum_{i=s-m+1}^s (x(i) - \hat{\mu}^D(s)) < g'(s)_{MAR}$$

where $\hat{\mu}^D(s) = \left(\sum_{i=s-n+1}^{s-m} x(i) \right) / (n-m)$ and $g'(s)_{MAR} = m \cdot n \cdot g(s)_{MAR} / (n-m)$. This is the usual moving average rule of surveillance expressed in (12) where μ^D is estimated by the moving average estimator $\hat{\mu}^D(s)$ in (13) with $q = n-m$ and $g_{MASur} = (m \cdot (m+q)) \cdot g(s)_{MAR} / q$.

A special case of (18) often considered is $m=1$ that is

$$x(s) - \frac{1}{n} \cdot \sum_{i=s-n+1}^s x(i) < g(s)_{MAR}. \quad (19)$$

Alarm rule (19) can be seen as comparing the last observation to the estimated trend and can be expressed as

$$x(s) - \hat{\mu}^D(s) < g'(s)_{MAR},$$

where $\hat{\mu}^D(s) = \left(\sum_{i=s-n+1}^{s-1} x(i) \right) / (n-1)$ and $g'(s)_{MAR} = (n/(n-1)) \cdot g(s)_{MAR}$. This is the situation when only the last observation is used in the surveillance, i.e. the Shewhart situation that was presented in (14). Here μ^D is estimated by the moving average estimator $\hat{\mu}^D(s)$ in (13).

As previously mentioned, the moving average rule of surveillance is optimal for discriminating between $C(s) = \{\tau = s - m + 1\}$ and $D(s) = \{\tau > s\}$ and the Shewhart method is optimal for $C(s) = \{\tau = s\}$ and $D(s) = \{\tau > s\}$. However, the optimality is not so clear-cut in (18) and (19), since μ^D is estimated and the results depend on the properties of the estimator.

Combinations often used for the window widths m and n in the alarm rule (18) are 1-50, 1-150, 2-200 and 5-200 (Gencay and Stengos (1997)). In the applied work by Brock et al. (1992), Levich and Thomas (1993), Mills (1997) and Neely (1997), no explicit motivation behind the choice of combinations is made and to our knowledge, no explicit rule of thumb for the choice of window widths has been suggested. Neely (1997) points out that the width is often determined by trial and error by the practitioner. Different aspects on the choice of the window widths is discussed below.

4.3.1 *The consequences of the choice of the window width in the moving average estimator of μ^D*

If we are to use the moving average method or the Shewhart method of surveillance for peak detection and μ^D is unknown, we must estimate it. If the parametric form of the process is known, the parameters can be estimated. If no such knowledge is at hand, one approach is to use the moving average estimator $\hat{\mu}^D$ in (13). Then however, we must determine the window width of the moving average, that is we must determine q in (13). Below we will discuss different aspects on the choice of q . We will consider three aspects that should be taken into account. These are the cycle length, the variance of the process under surveillance and the steepness of the trend. Focus will be given on alarm rule (19) that is the Shewhart approach with alarm rule (14) where μ^D is estimated by the moving average estimator (13) with window width q .

Say that the cycle length (peak to peak) is constant and equal to c . If a window width is used in the estimation that is much wider than c , then it is merely a question of comparing $x(s)$ with an estimator $\hat{\mu}$ that is unconditioned on the states. This will cause a considerable delay of motivated alarms since $\hat{\mu}$ will severely underestimate μ^D in a peak detection situation. If the monitoring is started right after a trough, the false alarm rate will be high for early time points. The reason for this is that when a wide window is used, observations before the trough are used in the moving average estimator at the start of the surveillance. A consequence is that the level will be overestimated at the start.

If the variance is large, a large window might be motivated because it will reduce the false alarm rate. However, the delay of a motivated alarm will increase with the window width. If there is considerable heteroscedasticity, it should be taken into account by using a weighted moving average instead of a simple one.

It was shown by Andersson and Bock (2001) that a trend cycle estimated by a moving average does not always preserve the true time of the turning point. For a certain window width, the location of the estimated turning point time resulting from the moving average depends on the steepness of the post-peak trend μ^C . If μ^C is steep,

the turning point time is preserved by the moving average and thus there is no systematic delay in the turning point detection.

Despite the wide spread use, window-based methods are sub-optimal because information about the process is lost since all observations are not used in the alarm statistic. When a narrow window width is used considerable information about the process is lost. If a wide window is used, the regime change will be smoothed and hard to detect. If two consecutive windows of fixed length are used, as in (18), the ability to detect a gradual change is low (Sveréus (1995)). Thus, instead of using window-based methods, we should use methods that utilize all observations, for example the EWMA method or the LR method, since they cause no loss of information. An advantage of window-based methods however, is that the in-control process, here represented by μ^D , does not need to be known.

As we have seen, the use of a moving average estimator of μ on non-stationary data is problematic. Thus, information from historical data or other prior knowledge on the form of the curve μ is very valuable.

4.4 HMR

The Hidden Markov Rule (HMR), as given by Marsh (2000) and Dewachter (2001), is proposed for trading in the foreign exchange market. It is assumed that trends in currency prices depend on an unobservable process. The X process is here the difference of the logarithmic transformed process, and X is described as a HMM where the expected value of the differentiated process X is constant depending on the state (see (4)). The switching between the states in (4) is governed by a first-order time-homogenous two-state hidden Markov chain $J(t)=\{1, 2\}$. The states 1 and 2 denote the expansion and recession phase, respectively, and are such that $E[X(t)|J(t)=1]=\beta_1$ and $E[X(t)|J(t)=2]=\beta_2$, respectively. In Dewachter, also a first-order hidden Markov chain governs the variance.

Marsh and Dewachter proposed an alarm statistic based on the one-step ahead predicted expected value of the differentiated process X conditional on past values. At decision time s the alarm rule is to signal that a peak has occurred if

$$E[X(s+1)|x_s] < 0 \quad (20)$$

where $E[X(s+1)|x_s]=P(J(s+1)=1|x_s)\cdot\beta_1+P(J(s+1)=2|x_s)\cdot\beta_2$ is the minimum mean squared prediction error (MSPE) forecast. The alarm limit zero used in (20) is not chosen to satisfy a certain level of the type I error. Kwan et al. (2000) suggest the following rule for selling

$$E[X(s+1)|x_s] < -2\cdot c,$$

where c is a transaction cost. Kwan et al. (2000) does not model X as a HMM. Instead, a price-trend model (Taylor (1980)) is used. The simplest form of the price-trend model assumes that X has a stochastic trend and the change in the trend is governed by a Bernoulli process. It was however shown by Dewachter (2001) that the HMM above belongs to the class of price trend models. From this point of view, the alarm limit zero, used above in (20), might be a natural choice when we have no transaction costs.

When the unobservable process is assumed to be governed by a HMM, a common approach is to use the posterior probability. The inference situation can be characterized as one of classification of the observations to states, as described in section 2.3, where

an alarm is called as soon as the system indicates a transition between the states. The rule for detecting a peak is to call an alarm as soon as

$$P(J(s)=2|x_s) > g_{PP}. \quad (21)$$

It was shown by Frisé and de Maré (1991) that when $D(s)=C(s)^c$, the complement of the critical event, then the posterior probability is inferentially equivalent to the LR method in (9). Since this method is optimal in various ways, it follows that HMR is also optimal in this case. However, the results of HMR and the LR method (for example in the sense of CED(i)) also depend on the assumptions and estimation procedures that are used for the parameters. The impact of these aspects was evaluated by Andersson et al. (2002b). Another important aspect is how the alarm limit g_{PP} is determined. This is discussed below.

4.4.1 How to choose the alarm limit (the effect of different specifications)

The alarm limit determines the sensitivity of the alarm statistic. Provided that $p_{21} \neq p_{22}$, the alarm rule for HMR in (20) is equivalent to using the rule (21) for the posterior probability with the alarm limit

$$g_{PP} = \frac{p_{11} \cdot \beta_1 + p_{12} \cdot \beta_2}{(\beta_1 - \beta_2) \cdot (p_{22} - p_{21})}, \quad (22)$$

where $p_{ij} = P(J(t)=j|J(t-1)=i)$, $i, j=1, 2$, are transition probabilities. The limit g_{PP} is sometimes set to 0.5, see e.g. Koskinen and Öller (1998) and Ivanova et al. (2000). This limit will, according to Zellner et al. (1991), give minimal expected loss if the loss of a false alarm equals that of a missed or delayed alarm (i.e. no signal of change at τ). For the situation when the location of the expected values β_1 and β_2 in (4) are symmetric with respect to the zero line (i.e. $\beta_2 = -\beta_1$), and the expected length of the expansion phase is the same as the expected length of the recession phase, that is $p_{12} = p_{21}$, then the limit g_{PP} in (22) equals 0.5. However, when the expected length of the expansion is shorter than for the recession ($p_{21} < p_{12} < p_{11}$), then g_{PP} is smaller than 0.5. This makes sense since we should then allow the alarm rule to be more sensitive for detecting peaks.

An alarm system that uses the information about the type of the next turning point is more efficient. Then the system is constructed for detecting one turn at a time, instead of classifying the whole sequence of turning points. If we know that the next turn is a peak, then the transition probability p_{21} equals to zero. Then p_{12} can be interpreted as the intensity of a peak, $v = P(\tau=t|\tau \geq t)$, given that τ has a geometric distribution. For a symmetric turn, the alarm limit g_{PP} in (22) then can be written as

$$g_{PP} = (1 - 2 \cdot p_{12}) / (1 - p_{12}) = (1 - 2 \cdot v) / (1 - v),$$

which tends to one as v tends to zero and minus infinity when v tends to one. This is reasonable since the limit 1 implies that an alarm signal should never be called (since $v \rightarrow 0$ means that the expected time of the turning point tends to infinity). The limit 0 implies that an alarm signal should be triggered immediately after the monitoring was started (since $v \rightarrow 1$ means that the expected time of the turning point tends to one).

Thus the symmetry of the turn and the intensity of the turns greatly influence the alarm limit. In surveillance methodology, the alarm limit is often determined indirectly, in order to control the false alarms, by setting the alarm limit so that a specific value of ARL^0 (or MRL^0 or PFA) is achieved. The value 0.5 is reasonable if the loss and gain of

a false and delayed alarm are the same and if the process is symmetric (both with respect to β and the transition probabilities), but otherwise quite arbitrary, since it does not give any control of the false alarm rate. A limit based on reasonable values of p_{11} , p_{22} , β_1 and β_2 , as in (22), is less arbitrary than the choice 0.5. However, if the values that are being used are far from the true values (i.e. not optimal), the alarm rate will be affected. Thus, HMR is sensitive to the values of the parameters.

Marsh and Dewachter use constant transition probabilities, p_{12} and p_{21} . A constant transition probability implies a constant intensity, i.e. a geometric distribution of the time of the change-point. This is a rather restrictive assumption that may not be realistic. When the intensity changes with time that is being spent in the regime, it is referred to as duration dependence. Whether regimes in the financial markets exhibit duration dependence have been investigated by e.g. McQueen and Thorley (1994), Cochran and Defina (1995) and Maheu and McCurdy (2000). When we are not sure about the form of the intensity, a non-informative prior for the change point time avoids the risk of serious misspecification. When a non-informative prior for τ is used as in the SR method, no transition probability needs to be estimated.

Another problem that Marsh (2000) addresses is the problem of instability of the estimates of β_1 and β_2 over different time periods. When no assumption is made regarding the parametric form of the curve as in the SRnp method, the surveillance system is robust against changes in the characteristics of the curve over time.

4.5 *The relation between measures used in statistical surveillance with measures used in evaluation of return*

In papers regarding financial decision rules, the expected return from investments is often used as a measure of performance. One measure of return that is often used is

$$r(t_A) = x(t_A) - x(0). \quad (23)$$

where X is the (logarithm of the) price. The above expression measures the return of buying the asset at $t=0$ and selling it at the alarm time t_A . The return in (23) does not take the timeliness of an alarm into account. However, as we will see below, under certain assumptions, return and timeliness are related.

The expected return $E[r(t_A)]$ is maximized when $E[X]$, is maximized. In the current setting, this occurs when we call an alarm (sell) at the time of the peak that is at $t_A = \tau - 1$, i.e. $\max\{E[r(t_A)]\} = E[r(\tau - 1)]$. For $t_A \neq \tau - 1$, $E[r(t_A)]$ decreases as $|t_A - \tau|$ increases and maximizing $E[r(t_A)]$ is equivalent to minimizing $|t_A - \tau|$. Consider the expected difference

$$\delta(t_A, t) = E[r(t_A) - r(\tau) | \tau = t] = E[X(t_A) - X(\tau) | \tau = t]$$

and the timeliness measure $\text{PSD}(d, t)$ in (6). The return in (23) is measured along the log-price scale whereas the timeliness measure $\text{PSD}(d, t)$ is measured along the time scale. When $X(t)$ is a known function of time, the timeliness measures can easily be expressed as a measure of the return. If $E[X(t)]$ is on the piecewise linear form in (3), $P(\delta(t_A, \tau) \geq d' | t_A \geq \tau, \tau = t)$, where $d' < 0$, is equivalent to $\text{PSD}(d, t) = P(t_A - \tau \leq d | t_A \geq \tau, \tau = t)$ where $d = d' / \beta_2 > 0$. A similar relation can be expressed for $t_A < \tau$.

In a financial setting it is important that transaction costs are taken into account. If transaction costs are ignored, it would be optimal to have a method that frequently generates trading signals. When we consider transaction costs, we have a penalty for each alarm which favors infrequent trading.

In section 3.1, we exemplified the expected utility by (8) where $h(t_A-\tau)=b$ in (7). This corresponds to a constant loss of a false alarm. However, as pointed out, the LR method with a constant intensity of a change will maximize the expected utility for an arbitrary function $h(t_A-\tau)$ as long as the gain of an alarm is a linear function of the expected delay. Hence, regardless of any specification of $h(t_A-\tau)$ in accordance with a realistic model of the expected return, the LR method will be optimal. If we have an expected return that increases as t_A approaches τ , we can for example use the function

$$u(t_A, \tau) = \begin{cases} b_1 \cdot (t_A - \tau) + b_2 & , t_A < \tau \\ a_1 \cdot (t_A - \tau) + a_2 & , t_A \geq \tau \end{cases}$$

where $b_1 > 0$, $a_1 < 0$ and where b_2 and a_2 would depend on e.g. the transaction cost. The function $u(t_A, \tau)$ above can be interpreted as the difference $E[r(t_A)] - E[r(\tau)]$, where the expectation is taken with respect to the disturbance ε . If the relationship between the gain of an alarm and the expected delay is nonlinear, the solution will no longer hold. But it might work well if the nonlinearity is not too pronounced.

5 EVALUATION BY DATA ON THE HANG SENG INDEX

The most common way to evaluate decision rules in financial literature is by using one or several case studies. Here we will illustrate some of the methods by applying them to two different periods of a real time series. The performance will be measured by the timeliness of the alarm and the return (23) of resulting transactions. The relative merits of case studies as an instrument of evaluation is also illustrated and discussed.

We will use the same time series that Lam and Yam (1997) used namely closing Hang Seng Index (HSI). HSI is a market-value weighted index of the stock prices of the 33 largest companies on the Hong Kong stock market. The weight each stock is assigned in the index is related to the price of the stock. HSI can thus be seen as the price of a portfolio of stock. A period of logarithmized values of HSI (4 January 1999 to 26 July 2002) is shown in Fig. 1. The two evaluation periods are marked with solid vertical lines. The methods are illustrated on two peaks, but troughs could equally well be used.

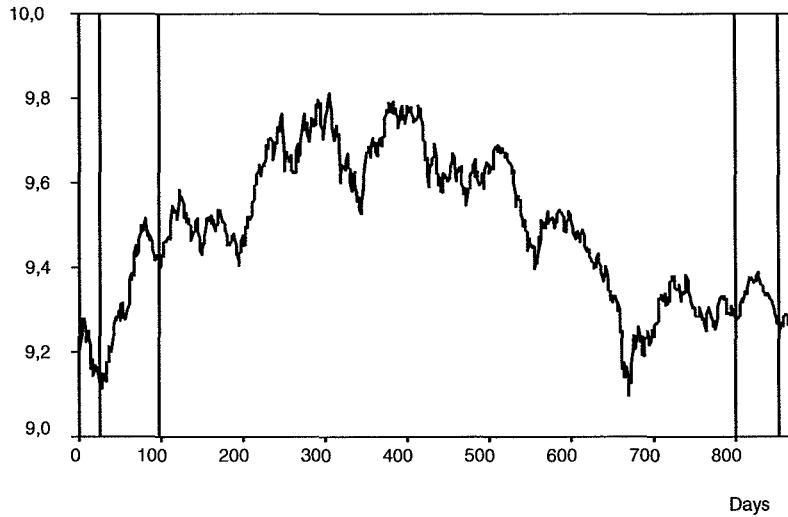


Fig. 1. Logarithm of values of daily observations of closing HSI from 4 January 1999 to 26 July 2002. The two periods considered for evaluation, 10 February 1999 to 28 May 1999 and 10 April 2002 to 26 June 2002, are marked with solid vertical lines.

The statistical surveillance methods that will be evaluated are SRnp (alarm rule (10)) and CUSUM (alarm rule (11)). The strategies suggested for financial trading decisions that will be evaluated, are GFR and FR. The alarm rules are expressed in (15) and (17), respectively. All these methods are used on the evaluation periods and the time of alarm is noted.

5.1 Characterizing the evaluation periods

The aim is to detect a turn in the trend cycle. To our knowledge, no official dating of turning points in HSI exists. We must hence determine the time of change τ , i.e. date the peak, for the two periods. First, for each period, we estimate the trend cycle μ under the assumption that the period includes a single peak. The estimation is made by maximum likelihood under the restriction that the regression is inversely U-shaped and under the assumption that the disturbances are normally distributed as in model (1) (see Frisé (1986)). Then τ is determined such that

$$\mu(1) \leq \dots \leq \mu(\tau-1) \text{ and } \mu(\tau-1) \geq \mu(\tau) \dots \geq \mu(t).$$

The values of τ for the two evaluation periods are marked in Fig. 2 and Fig. 3, respectively.

For descriptive purpose we first fit a model to the logarithmized values of the data in the evaluation periods. We use the following model:

$$\ln Y(t) = \mu(t) + \varepsilon(t),$$

where $\varepsilon(t) = \theta \cdot \varepsilon(t-1) + \delta(t)$, $\delta(t) \sim \text{iid } N[0, \sigma_\delta^2]$, $t=1, 2, \dots$, and where Y is the price of a unit of HSI. The expected value of $\ln Y$ is modeled using the piecewise linear function (3), that is

$$\mu(t) : \begin{cases} \mu^D(t) : \beta_0 + \beta_1 \cdot t, & t < \tau \\ \mu^{C\tau}(t) : \beta_0 + \beta_1 \cdot (\tau-1) + \beta_2 \cdot (t-\tau+1), & t \geq \tau \end{cases}$$

The first evaluation period under consideration, 10 February 1999 to 28 May 1999, is displayed in Fig. 2 below. The time of change is estimated to 6 May 1999 ($\tau=55$). The monitoring is started at 11 February ($t=1$).

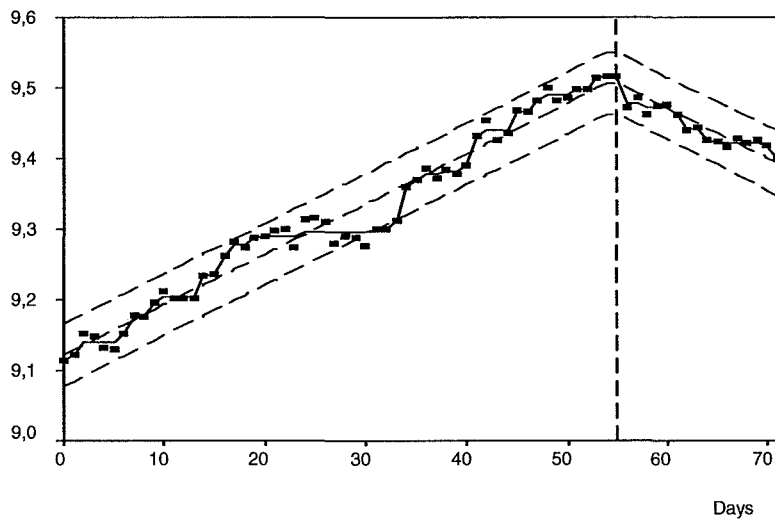


Fig. 2. First evaluation period. Logarithm of values of daily observations of the HSI from 10 February 1999 to 28 May 1999 is marked with dots. The time of change, 6 May 1999 ($\tau=55$) is marked with a dotted vertical line. The μ -curve estimated by inversely U-shaped regression is marked with a solid curve. The μ -curve for the piecewise linear model (3), estimated by ordinary least squares, with $\tau=55$ and the 2.5th and 97.5th percentiles of values of the observations according to the model are marked as dotted lines.

The second evaluation period, 10 April 2002 to 26 June 2002, is shown in Fig. 3. The time of the change is estimated to 21 May 2002 ($\tau=27$). The monitoring is started at 11 April ($t=1$).

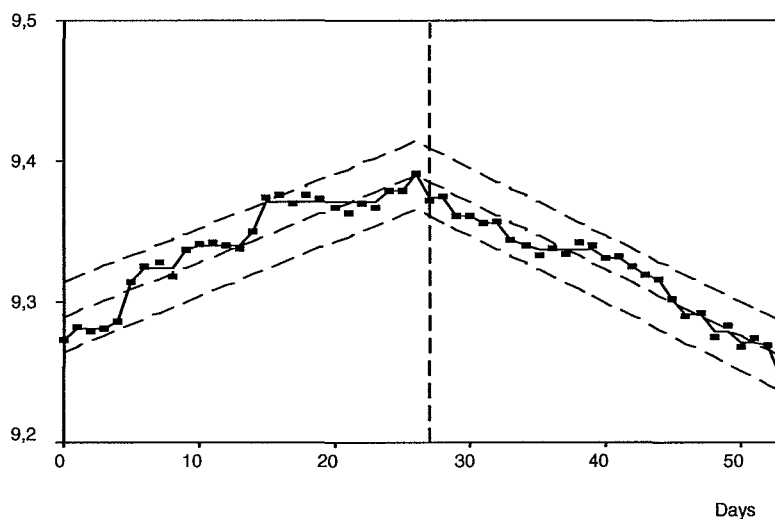


Fig. 3. Second evaluation period. Logarithm of values of daily observations of the HSI from 10 April 2002 to 26 June 2002 is marked with dots. The time of change, 21 May 2002 ($\tau=27$) is marked with a dotted vertical line. The μ -curve estimated by inversely U-shaped regression is marked with a solid curve. The μ -curve for the piecewise linear model (3), estimated by ordinary least squares, with $\tau=27$ and the 2.5th and 97.5th percentiles of values of the observations according to the model are marked as dotted lines.

5.2 Specification of the surveillance systems

The alarm limit in a surveillance system should be determined so as to get control over the false alarms. Generally, the alarm limit is set so that ARL^0 , MRL^0 or PFA equal a certain acceptable value. When we want to compare several methods, their respectively alarm limits are adjusted to yield the same ARL^0 (or MRL^0 or PFA). The specific value of an alarm limit depends on several factors, such as the signal-noise ratio and the dependency structure of the process under surveillance. The signal noise is a measure of how well the regime shift can be distinguished from the noise. We measure the signal-noise ratio as β_1/σ_δ .

The observations on HSI are used in the surveillance. However, the methods differ concerning the transformations. SRnp is applied to $\ln Y(t)$, whereas CUSUM is applied to the logarithmized and differentiated process, $\ln Y(t) - \ln Y(t-1)$ and GFR and FR are applied to $Y(t)$.

Some of the alarm statistics depend on known or estimated values of the parameters β_1 , β_2 and σ_δ . We estimate the parameters by observations from the first evaluation period. The resulting estimated signal-noise ratio is 0.46. The alarm limits for SRnp and CUSUM are set to yield an MRL^0 of approximately 60 days, conditional on the estimated model for $\ln Y$ above. For SRnp it suffices to estimate the variance of the process. Then the alarm limit, g_{SRnp} , is determined by simulations so that $MRL^0=60$. But for CUSUM, also the optimal value of k in the limit of the CUSUM method, $-(\mu^C - \mu^D)/2$, which here corresponds to $-(\beta_2 - \beta_1)/2$, must be estimated. The value of k obtained by using the estimated parameters is

$$k = -(\hat{\beta}_2 - \hat{\beta}_1)/2 = 0.00689.$$

Then the other limit parameter, g_{CUSUM} , is also determined by simulations so that $MRL^0=60$.

In a practical online monitoring situation, where the process is continually observed, the parameters must be estimated by observations that are available at the current time point, e.g. past observations of the process. We will take this aspect into account in the following way. When we evaluate the methods in the second period, we do not re-estimate the parameters. Instead, we use the parameters estimated by observations from the first period again.

In Lam and Yam (1997) no discussion is made regarding the false alarm rate. Instead the parameter values g_{CUSUM} and k are determined to yield maximum return. As a comparison, their best values are also used. The best limit for FR is $\{-0.09\}$ and for GFR, the best parameter combination of g_{CUSUM} and k is $\{-0.06, -0.003\}$. Due to the similarity between GFR and CUSUM, only the CUSUM method will be investigated. The two alarm systems are referred to as FR(L&Y best) and CUSUM(L&Y best).

5.3 Results

When evaluating on a real time series including several turns, different approaches can be used, sometimes depending on the measure of evaluation. In active surveillance it is often the case that only the first alarm and the time to the first alarm are noted. Thus the expected delay only refers to the delay time for the first alarm. But when measures such as rate of return is used, then it is natural to restart the surveillance after an alarm, since it is likely that an asset would be sold, then bought, then sold again. Below both approaches will be tried.

5.3.1 Results when the surveillance systems are not restarted after an alarm is given

The results of the two evaluations when the monitoring is not restarted after the first alarm time, i.e. a single trading decision is made during each period, is reported in the tables below. The measure of return, $r(t_A)$ in (23) is reported. We assume that it is possible to sell for the price at the alarm time t_A . In Table 1, the return of buying a unit of the index at 9 February 1999 ($t=0$) for the price 9244.49 and selling it at t_A is reported. Transaction costs are ignored. For CUSUM the parameter values $\hat{\beta}_1$ and $\hat{\beta}_2$ are estimated from the same period as is used in the evaluation (in-sample performance).

Table 1. Results of the evaluation on the first period (10 February 1999 to 28 May 1999). Estimated time of change $\tau=55$.

	t_A	$t_A-\tau$	$ t_A-\tau $	$r(t_A)$
SRnp(MRL ⁰ =60)	29	-26	26	0.158
CUSUM(MRL ⁰ =60)	56	1	1	0.341
FR(L&Y best)	64	9	9	0.294
CUSUM(L&Y best)	64	9	9	0.294

SRnp gives an alarm at the plateau (see Fig. 2). As opposed to CUSUM and FR that impose parametric restrictions on the turning point curve through the value of k , SRnp places no such restrictions. All information about the curve comes from data. Hence, SRnp is in this sense more sensitive to the values of the observations compared to the other methods. The success of CUSUM(MRL⁰=60) might be due to the large drop in the level right after τ (see Fig. 2). The magnitude of the drop is almost twice the estimated standard deviation of the process.

In addition to the evaluation of the FR that use the limit of Lam and Yam (1997), referred to as FR(L&Y best) in the table above, we also apply the FR method with another limit. We use the limit g_{CUSUM} that yield MRL⁰=60 for the CUSUM method. We do this in order to see whether there will a difference in performance between using the optimal value of k (estimated by $-(\hat{\beta}_2-\hat{\beta}_1)/2$) and using the value $k=\hat{\beta}_1$ implied by FR. The result is that FR gives an alarm at the same time point as CUSUM, i.e. at $t=56$. This is because the turn is, according to our estimate, almost symmetric and hence the estimated optimal value of k is close to $\hat{\beta}_1$.

The results of the evaluation on the second period when the parameter estimates obtained in the previous evaluation period (out-of-sample performance) are used again are reported in Table 2. It is assumed that a unit of the index is bought at 10 April 2002 ($t=0$) for the price 10643.14.

Table 2. Results of the evaluation on the second period (10 April 2002 to 26 June 2002). Estimated time of change $\tau=27$.

	t_A	$t_A-\tau$	$ t_A-\tau $	$r(t_A)$
SRnp(MRL ⁰ =60)	29	2	2	0.088
CUSUM(MRL ⁰ =60)	33	6	6	0.071
FR(L&Y best)	46	19	19	0.017
CUSUM(L&Y best)	53	26	26	-0.003

As seen in the first two rows in the table, SRnp outperforms CUSUM when $k=0.00689$, the value estimated from the first evaluation period, is used in the alarm

rule of CUSUM. If $k=0.00431$, which is the estimate from the current period, is used instead, CUSUM gives an alarm at $t=27$, i.e. at τ . The alarm yields a return of 0.099 instead of 0.071. Hence, for a fixed value of g_{CUSUM} , the choice of k has an effect here. The signal-noise ratio is very similar for both periods. When we again apply the FR with the limit g_{CUSUM} that yield $\text{MRL}^0=60$ for CUSUM, an alarm is given at $t=33$, i.e. at the same time as CUSUM. For both evaluation periods, the alarm limits suggested by Lam and Yam (1997) give late alarms. Hence, the surveillance systems based on these limits seem to be inappropriate for these periods.

5.3.2 Results when the surveillance systems are restarted after an alarm is given

We will now consider the situation when we are not restricted to stop the monitoring after the first alarm was given. Instead, we are allowed to restart the monitoring immediately, i.e. we can make a sequence of buy and sell decisions during the period.

With this approach, we are given the following results for the first period. After the first alarm at $t=29$ of SR_{np} , a signal that a trough has occurred (buy signal) is called at $t=35$. An additional alarm that a peak has occurred (sell) is given at $t=57$. The cumulated return is 0.270. The other methods give no alarms later than the times reported in Table 1. For the second period, no alarms are called later than the times reported in Table 2.

5.3.3 The effect of autocorrelation

For the evaluation periods that we considered, there was a strong autocorrelation in the data. Many methods are constructed under the assumption of independent observations and might be misleading in their results if they are used directly on data with a dependency process.

In statistical surveillance there are several ways of dealing with dependent data (see Pettersson (1998) and Friséen (2003)). One approach is to use the correct likelihood ratio statistic. The statistic for a change in the mean of a stationary AR(1) process with normally distributed disturbances is derived in Pettersson (1998). Another approach is to calculate the residuals of an estimated model of the dependency structure or to differentiate the observations in order to get rid of the dependency. A third approach is to just adjust the alarm limit so that the MRL^0 (or ARL^0 or PFA) is correct for the dependent process. Sometimes the dependence problem is ignored; no consideration is given whether the data has a dependency structure or not.

When Lam and Yam calculate the ARL^0 analytically, a possible dependency structure is not discussed. The data in our two periods has an AR(1) structure with a rather high estimated AR parameter ($\hat{\theta}=0.7$). Whether the dependence is considered in the model or not has a large impact on the results here. Thus further improvements can be expected by the use of methods of surveillance that takes the dependency structure into account.

5.3.4 Case studies versus Monte Carlo methods

The use of a single case study for evaluation can be questioned. How well does the actual data set represent the process of interest? How extreme is the outcome of the process? If it is very extreme, then the resulting alarm time is very rare. However, it is impossible to know whether an outcome is extreme or not if not several examples are available or if the process is replicated. If we e.g. replicate the process by Monte Carlo

methods, then we can make statements about the properties of a method under the assumptions in that study.

6 DISCUSSION AND CONCLUDING REMARKS

Inferential differences and similarities between statistical surveillance and some strategies suggested for financial trading decisions have been investigated. The latter were represented by two variants of the filter rule, a rule based on moving averages and a decision rule based on a hidden Markov model. All three methods aim at prospectively detecting turning points in a time series.

A desirable property of any prospective monitoring method is that a change can be detected quickly (timeliness) without having too many unmotivated signals of change (safety). In the general theory of statistical surveillance, the aim is to optimize the method with respect to these properties. Examples of measures that reflect timeliness are the expected delay and the absolute probability of successful detection. The latter is of particular interest in a financial setting since trading right before a turn is valuable. The predictive value is another useful measure that reflects the trust you should have in an alarm.

When financial decision rules are used in practice, measures of the return from transactions made are used to evaluate the performance. Since the aim of a financial decision rule generally is to maximize the expected return (adjusted for the risk exposure and transaction costs), return-measures are natural to use. Most of the results in the theory of statistical surveillance has been developed on the sole use of the timeliness scale. However, timeliness and rate of return are, as we have seen, often related. Thus, evaluation that both reflects timeliness and return should be used.

It is well known that the CUSUM method is minimax optimal if the observations under surveillance are independent and normally distributed. The FR has been shown to be a special case of the CUSUM method. The special case is when the location of in- and out-of-control means are symmetric with respect to the zero line, which is for example the case when there is a change in the drift of a random walk where the in- and out-of-control drift parameters are equal in absolute values. Thus, when this is the case, and the differentiated observations are independent and normally distributed, FR is also minimax optimal. However, this type of symmetry is a rather restrictive assumption and the CUSUM method is more general. GFR is an approximation of CUSUM that does take the aspect of the relation between the slopes of the upward and downward trend into account. Thus, GFR is approximately minimax optimal.

Methods based on moving averages are common and relatively simple to construct. One method, widely used for trading, consists of comparing two overlapping moving averages of different length. In this paper it is demonstrated that this approach is equivalent to a moving average method described in the literature of surveillance, where the in-control mean is estimated by a moving average. A special case often considered is to compare the current observation with a simple moving average of observations. It was shown that this method is similar to the Shewhart method of surveillance where, again, the in-control mean is estimated by a moving average. The optimality properties of these two surveillance methods are known. This does not however necessarily imply that the moving average trading rule is optimal since the result depends on e.g. the properties of the moving average estimator. Another objection to the moving average methods is that it is often of interest to detect a change

that occurred in an unknown past and therefore methods that use all available observations, e.g. the full LR method should be used instead of methods that use only the last m observations.

In both business cycles and financial trading rules there are several methods that use a Hidden Markov model approach. The inferential structure of the HMR method is equivalent to that of the LR method when the specifications of the events that are to be discriminated between are the same. This requires however that we have knowledge of the type of the next turn. Thus the method is also optimal in the sense that the expected utility is maximized. However, the results from the monitoring also depend on the knowledge (or method for estimating) of parameters and the distribution of τ .

Improper assumptions regarding the process under surveillance may have a great impact on the performance. When the distribution of the time of the change τ is unknown, a non-informative prior for τ can be used in order to avoid the risk of serious misspecification. This is what the SRLin and SRnp methods use. A way of avoiding the risk of a serious misspecification of the regression is to use the non-parametric approach of the SRnp method. The case studies illustrate the advantage of the SRnp method when the current turn has a different shape than the previous.

Since there are many problems to deal with in the implementation of surveillance in finance, further research is needed. Much effort has been made on the modeling aspect of processes related to finance, e.g. volatility, the arrival time of transactions and smooth transitions between regimes. However, not much work has been made in the implementation of these models in a surveillance system.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Marianne Frisé, for her encouraging guidance and support during the work of this report. I would also like to thank my assistant supervisor Dr Eva Andersson, Christian Sonesson and assistant Professor Helgi Tomasson, Faculty of Economics and Business Administration, University of Iceland, for valuable comments. This paper is a part of a research project on statistical surveillance at Göteborg University, supported by the Swedish Council for Research in the Humanities and Social Sciences. The work on this paper has also been supported by Kungliga and Hvitfeldtska Stiftelsen and Wilhelm and Martina Lundgrens Vetenskapsfond 1.

REFERENCES

- Alexander, S. S. (1961) Price Movements in Speculative Markets: Trends or Random Walks. *Industrial Management Review*, **2**, 7-26.
- Alexander, S. S. (1964) Price Movements in Speculative Markets: Trends or Random Walks, No. 2. *Industrial Management Review*, 25-46.
- Andersson, E. (2001) Turning Point Detection Using Non-parametric Statistical Surveillance. Evaluation of Some Influential Factors. Research Report, Department of Statistics, Göteborg University, 2001:3.
- Andersson, E. (2002) Monitoring Cyclical Processes. A Non-parametric Approach. *Journal of Applied Statistics*, **29**, 973-990.
- Andersson, E. and Bock, D. (2001) On Seasonal Filters and Monotonicity. Research Report, Department of Statistics, Göteborgs university, 2001:4.
- Andersson, E., Bock, D. and Frisé, M. (2002a) Some statistical aspects on methods for detection of turning points in business cycles. *Submitted*.
- Andersson, E., Bock, D. and Frisé, M. (2002b) Statistical surveillance of cyclical processes with application to turns in business cycles. *Submitted*.
- Black, F. (1976) The pricing of commodity contracts. *Journal of Financial Economics*, **3**, 167-179.
- Boehm, E. A. and Moore, G. H. (1991) Financial Market Forecasts and Rates of Return Based on Leading Index Signals. *International Journal of Forecasting*, **7**, 357-374.
- Bojdecki, T. (1979) Probability maximizing approach to optimal stopping and its application to a disorder problem. *Stochastics*, **3**, 61-71.
- Brock, W., Lakonishok, J. and LeBaron, B. (1992) Simple Technical Trading Rules and the Stochastic Properties of Stock Returns. *Journal of Finance*, **47**, 1731-1764.
- Brooks, C., Rew, A. G. and Ritson, S. (2001) A trading strategy based on the lead-lag relationship between the spot index and futures contract for the FTSE 100. *International Journal of Forecasting*, **17**, 31-44.
- Brown, R. L., Durbin, J. and Evans, J. M. (1975) Techniques for Testing the Constancy of Regression Relationships over Time. *Journal of the Royal Statistical Society B*, **37**, 149-192.
- Chang, J. T. and Fricker, R. D. (1999) Detecting when a monotonically increasing mean has crossed a threshold. *Journal of Quality Technology*, **31**, 217-234.
- Chang, T. C. and Gan, F. F. (1995) A Cumulative Sum Control Chart For Monitoring Process Variance. *Journal of Quality Technology*, **27**, 109-119.
- Chen, J. and Gupta, A. K. (1997) Testing and locating variance changepoints with application to stock prices. *Journal of the American Statistical Association*, **92**, 739-747.
- Cochran, J. S. and Defina, R. H. (1995) Duration dependence in the US stock market cycle: a parametric approach. *Applied Financial Economics*, **5**, 309-318.
- Crowder, S. (1987) A simple method for studying run-length distributions of exponentially weighted moving average charts. *Technometrics*, **29**, 401-407.
- Dewachter, H. (1997) Sign predictions of exchange rate changes: Charts as proxies for Bayesian inferences. *Weltwirtschaftliches Archiv-Review of World Economics*, **133**, 39-55.
- Dewachter, H. (2001) Can Markov switching models replicate chartist profits in the foreign exchange market? *Journal of International Money and Finance*, **20**, 25-41.
- Dufour, A. and Engle, R. F. (2000) Time and the price impact of a trade. *Journal of Finance*, **55**, 2467-2498.
- Easley, D. and O'Hara, M. (1992) Time and the Process of Security Price Adjustment. *Journal of Finance*, **47**, 577-605.
- Edwards, R. D. and Magee, J. (1992) *Technical Analysis of Stock Trends*, John Magee Inc., Boston.
- Engle, R. F. and Russell, J. R. (1998) Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, **66**, 1127-1162.
- Fama, E. F. and Blume, M. E. (1966) Filter Rules and Stock-Market Trading. *Journal of Business*, **39**, 226-241.
- Franses, P. H. and van Dijk, D. (2000) *Non-linear Time Series Models in Empirical Finance*, Cambridge University Press, Cambridge.
- Frisé, M. (1986) Unimodal regression. *The Statistician*, **35**, 479-485.
- Frisé, M. (1992) Evaluations of Methods for Statistical Surveillance. *Statistics in Medicine*, **11**, 1489-1502.
- Frisé, M. (1994) Statistical Surveillance of Business Cycles. Research Report, Department of Statistics, Göteborg University, 1994:1 (Revised 2000).

- Frisén, M. (2003) Statistical Surveillance. Optimality and Methods. *to appear in International Statistical Review*.
- Frisén, M. and de Maré, J. (1991) Optimal Surveillance. *Biometrika*, **78**, 271-80.
- Frisén, M. and Sonesson, C. (2002) Optimal Surveillance Based on Exponentially Weighted Moving Averages. Research Report, Department of Statistics, Göteborg University, 2002:1.
- Frisén, M. and Wessman, P. (1999) Evaluations of likelihood ratio methods for surveillance. Differences and robustness. *Communications in Statistics. Simulations and Computations*, **28**, 597-622.
- Gan, F. (1998) Designs of one- and two-sided exponential EWMA charts. *Journal of Quality Technology*, 55-69.
- Gan, F. F. (1992) Cusum Control Charts Under Linear Drift. *Statistician*, **41**, 71-84.
- Gencay, R. and Stengos, T. (1997) Technical Trading Rules and the Size of the Risk Premium in Security Returns. *Journal of Nonlinear Dynamics and Econometrics*, **2**, 23-34.
- Hsu, D. A. (1982) A Bayesian Robust Detection of Shift in the Risk Structure of Stock Market Returns. *Journal of the American Statistical Association*, **77**, 29-39.
- Inclan, C. and Tiao, G. C. (1994) Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance. *Journal of the American Statistical Association*, **89**, 913-923.
- Ivanova, D., Lahiri, K. and Seitz, F. (2000) Interest rate spreads as predictors of German inflation and business cycles. *International Journal of Forecasting*, **16**, 39-58.
- Jönsson, H. (2001) Monte Carlo studies of American type call options with discrete time. *Theory of Stochastic Processes*, **7**, 163-188.
- Koskinen, L. and Öller, L.-E. (1998) A Classifying Procedure for Signalling Turning Points. Research Report, National institute of economic research, Stockholm, Sweden, 59.
- Kukush, A. G. and Silvestrov, D. S. (2000) Optimal pricing for American type options with discrete time. Research Report, Department of Mathematics and Physics, Mälardalen University, 1.
- Kwan, J. W. C., Lam, K., So, M. K. P. and Yu, P. L. H. (2000) Forecasting and Trading Strategies Based on a Price Trend Model. *Journal of Forecasting*, **19**, 485-498.
- Lai, T. L. (1995) Sequential change-point detection in quality control and dynamic systems. *Journal of the Royal Statistical Society B*, **57**, 613-658.
- Lam, K. and Yam, H. C. (1997) CUSUM Techniques for Technical Trading in Financial Markets. *Financial Engineering and the Japanese Markets*, **4**, 257-74.
- Layton, A. P. (1996) Dating and predicting phase changes in the U.S. business cycle. *International Journal of Forecasting*, **12**, 417-428.
- Layton, A. P. and Katsuura, M. (2001) A new turning point signalling system using the Markov switching model with application to Japan, the USA and Australia. *Applied Economics*, **33**, 59-70.
- Leitch, G. and Tanner, E. (1991) Economic Forecast Evaluation: Profits Versus the Conventional Error Measures. *The American Economic Review*, **81**, 580-590.
- Levich, R. M. and Thomas, L. R. (1993) The Significance of Technical Trading-Rule Profits in the Foreign Exchange Market: A Bootstrap Approach. *Journal of International Money and Finance*, **12**, 451-474.
- Lo, A. W. (2000) Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation. *Journal of Finance*, **55**, 1705-1770.
- Maheu, J. M. and McCurdy, T. H. (2000) Identifying Bull and Bear Markets in Stock Returns. *Journal of Business and Economic Statistics*, **18**, 100-112.
- Marsh, I. W. (2000) High-frequency Markov Switching Models in the Foreign Exchange Market. *Journal of Forecasting*, **19**, 123-134.
- McQueen, G. and Thorley, S. (1994) Bubbles, Stock Returns and Duration Dependence. *Journal of Financial and Quantitative Analysis*, **29**, 379-401.
- Mills, T. C. (1997) Technical analysis and the London Stock Exchange: testing trading rules using the FT30. *International Journal of Finance & Economics*, **2**, 319-331.
- Moore, G. H., Boehm, E. A. and Banerji, A. (1994) Using Economic Indicators to Reduce Risk in Stock-Market Investments. *International Journal of Forecasting*, **10**, 405-417.
- Moustakides, G. V. (1986) Optimal stopping times for detecting changes in distributions. *Annals of Statistics*, **14**, 1379-87.
- Neely, C. J. (1997) Technical Analysis in the Foreign Exchange Market: A Layman's Guide. *The Federal Reserve Bank of St. Louis Review*, **79**, 23-38.
- Neely, C. J., Weller, P. and Dittmar, R. (1997) Is technical analysis in the foreign exchange market profitable? A genetic programming approach. *Journal of Financial and Quantitative Analysis*, **32**, 405-426.

- Neftci, S. N. (1991) Naive Trading Rules in Financial-Markets and Wiener-Kolmogorov Prediction-Theory - a Study of Technical Analysis. *Journal of Business*, **64**, 549-571.
- O'Hara, M. (1995) *Market Microstructure Theory*, Blackwell Publisher, Cambridge.
- Page, E. S. (1954) Continuous inspection schemes. *Biometrika*, **41**, 100-114.
- Pettersson, M. (1998) Evaluation of some methods for statistical surveillance of an autoregressive process. Research Report, Department of Statistics, Göteborg University, 1998:4.
- Petzold, M., Sonesson, C., Bergman, E. and Kieler, H. (2003) Detection of intrauterine growth restriction. *Submitted*.
- Pring, M. J. (1985) *Technical Analysis Explained*, McGraw-Hill, New York.
- Roberts, S. W. (1959) Control Chart Tests Based on Geometric Moving Averages. *Technometrics*, **1**, 239-250.
- Roberts, S. W. (1966) A Comparison of some Control Chart Procedures. *Technometrics*, **8**, 411-430.
- Robertson, T., Wright, F. T. and Dykstra, R. L. (1988) *Order Restricted Statistical Inference*, John Wiley & Sons Ltd.
- Robinson, P. B. and Ho, T. Y. (1978) Average Run Lengths of Geometric Moving Average Charts by Numerical Methods. *Technometrics*, **20**, 85-93.
- Rydberg, T. H. and Shepard, N. (1999) BIN models for trade-by-trade data. Modelling the number of trades in fixed interval of time. *Manuscript*.
- Schipper, S. and Schmid, W. (2001a) Control charts for GARCH processes. *Nonlinear Analysis*, **47**, 2049-2060.
- Schipper, S. and Schmid, W. (2001b) Sequential Methods for Detecting Changes in the Variance of Economic Time Series. *Sequential Analysis*, **20**, 235-262.
- Severin, T. and Schmid, W. (1999) Monitoring changes in GARCH processes. *Allgemeines Statistisches Archiv*, **83**, 281-307.
- Shiryayev, A. N. (1963) On optimum methods in quickest detection problems. *Theory of Probability and its Applications*, **8**, 22-46.
- Shiryayev, A. N. (1999) *Essentials of Stochastic Finance*, World Scientific Pub. Co, Singapore.
- Shiryayev, A. N., Kabanov, Y. M., Kramkov, O. D. and Melnikov, A. V. (1994) Toward the Theory of Pricing of Options of Both European and American Types. I. Discrete Time. *Theory of Probability and its Applications*, **39**, 14-60.
- Sonesson, C. (2001) Evaluations of Some Exponentially Weighted Moving Average Methods. Research Report, Department of Statistics, Göteborg University, 2001:6.
- Sonesson, C. and Bock, D. (2003) A review and discussion of prospective statistical surveillance in public health. *Journal of the Royal Statistical Society A*, **166**, 5-21.
- Srivastava, M. S. and Wu, Y. (1993) Comparison of EWMA, CUSUM and Shiryaev-Roberts Procedures for Detecting a Shift in the Mean. *Annals of Statistics*, **21**, 645-670.
- Steland, A. (2002) Nonparametric Monitoring of Financial Time Series by Jump-preserving Control Charts. *Statistical Papers, Berlin*, **43**, 401-422.
- Sweeney, R. J. (1986) Beating the Foreign Exchange Market. *The Journal of Finance*, **41**, 163-182.
- Sweeney, R. J. (1988) Some New Filter Rule Tests: Methods and Results. *Journal of Financial and Quantitative Analysis*, **23**, 285-300.
- Sver us, A. (1995) Detection of successive changes. Statistical methods in postmarketing surveillance. Research Report, Department of Statistics, Göteborg university, 1995:2.
- Szakmary, A. and Davidson III, W. N. (1999) Filter Tests in Nasdaq Stocks. *The Financial Review*, **34**, 45-70.
- Taylor, S. (1986) *Modelling Financial Time Series*, John Wiley & Sons, Chichester.
- Taylor, S. J. (1980) Conjectured Models for Trends in Financial Prices, Tests and Forecasts. *Journal of the Royal Statistical Society A*, **143**, 338-362.
- Ter svirta, T. (1994) Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models. *Journal of the American Statistical Association*, **89**, 208-218.
- Ter svirta, T. (1998) Modelling Economic Relationships with Smooth Transition Regressions. In *Handbook of Applied Economic Statistics* (Eds, Ullah, A. and Giles, D. E. A.) Marcel Dekker, New York, pp. 507-552.
- Theodossiou, P. T. (1993) Predicting Shifts in the Mean of a Multivariate Time Series Process: An Application in Predicting Business Failures. *Journal of the American Statistical Association*, **88**, 441-449.
- Van Dijk, D., Ter svirta, T. and Franses, P. H. (2002) Smooth Transition Autoregressive Models - A Survey of Recent Developments. *Econometric Reviews*, **21**, 1-47.
- Vardeman, S. and Ray, D. (1985) Average Run Lengths for CUSUM Schemes When Observations Are Exponentially Distributed. *Technometrics*, **27**, 145-150.

- Wessman, P. (1998) Some Principles for surveillance adopted for multivariate processes with a common change point. *Communications in Statistics. Theory and Methods*, **27**, 1143-1161.
- Wessman, P. (1999) Studies on the surveillance of univariate and multivariate processes. Doctoral thesis, Department of Statistics, Göteborg University, Sweden, Göteborg.
- Wetherhill, G. B. and Brown, D. W. (1991) *Statistical Process Control: Theory and Practice*, Chapman and Hill, London.
- Wichern, D. W., Miller, R. B. and Hsu, D.-A. (1976) Changes of Variance in First-Order Autoregressive Time Series Models-With an Application. *Applied Statistics*, **25**, 248-256.
- von Collani, E. and Sheil, J. (1989) An Approach to Controlling Process Variability. *Journal of Quality Technology*, **21**, 87-96.
- Zellner, A., Hong, C. and Min, C.-K. (1991) Forecasting Turning Points in International Output Growth Rates Using Bayesian Exponentially Weighted Autoregression, Time-Varying Parameter, and Pooling Techniques. *Journal of Econometrics*, **49**, 275-304.
- Zhang, M. Y., Russell, J. R. and Tsay, R. S. (2001) A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics*, **104**, 179-207.

Research Report

- | | | |
|--------|--------------------------------------|---|
| 2002:1 | Frisén, M. & Sonesson, C.: | Optimal surveillance. Based on exponentially weighted moving averages. |
| 2002:2 | Frisén, M.: | Statistical surveillance. Optimality and methods. |
| 2002:3 | Jonsson, R. & Persson, A.: | Bayes prediction of binary outcomes based on correlated discrete predictors. |
| 2002:4 | Persson, A.: | Prediction of work resumption among men and women with lower back- and neck pain in a Swedish population. |
| 2002:5 | Persson, A.: | Prediction of work resumption in theory and practice. |
| 2002:6 | Andersson, E., Bock, D. & Frisé, M.: | Detection of turning points in business cycles. |
| 2002:7 | Andersson, E., Bock, D. & Frisé, M.: | Some statistical aspects on methods for detection of turning points in business cycles. |
| 2002:8 | Andersson, E., Bock, D. & Frisé, M.: | Statistical surveillance of cyclical processes with application to turns in business cycles. |
| 2002:9 | Holgersson, T.: | Testing for non-normality in multivariate regression with nonspherical disturbances. |
| 2003:1 | Holgersson, T. & Shukur, G.: | Testing for multivariate heteroscedasticity. |
| 2003:2 | Holgersson, T.: | Testing for multivariate autocorrelation. |
| 2003:3 | Petzold, M. & Sonesson, C.: | Detection of intrauterine growth restriction. |