

Testing the Predictive Power of Technical Analysis' Trading Rules Using Modified Box-Tiao Time Series Models

Alexandros E. Milionis

Department of Statistics and Actuarial-Financial Mathematics, University of the Aegean, Karlovassi, Greece
Email: amilionis@aegean.gr

How to cite this paper: Milionis, A.E. (2025) Testing the Predictive Power of Technical Analysis' Trading Rules Using Modified Box-Tiao Time Series Models. *Journal of Mathematical Finance*, 15, 514-534.
<https://doi.org/10.4236/jmf.2025.153021>

Received: April 1, 2025

Accepted: July 22, 2025

Published: July 25, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In the battery of existing statistical techniques for the testing of the predictive power of trading rules of technical analysis and, in extension, for the testing of the hypothesis of efficiency in financial markets, this work proposes one more for inclusion. The proposed technique is a modification of the Box-Tiao stochastic models for the assessment of the impact of external events on time series. The potential advantages of the proposed methodology over the existing ones (e.g. simple t -tests, bootstrapping techniques) are discussed. Application of the proposed statistical technique to the returns of the FT30 Index of the London Stock Exchange, in conjunction with the moving average technical trading rule, shows good agreement with the empirical findings of other methods.

Keywords

Market Efficiency, Trading Rules of Technical Analysis, Moving Averages, Box-Tiao Models, London Stock Exchange

1. Introduction

The concept of market efficiency is very important in modern financial theory. Although market efficiency is defined differently by different authors (e.g. Beaver [1]; Malkiel [2]; Milionis [3]; Rubinstein [4]), it is the definition due to Fama (1970) that has become the established one, according to which a market is efficient if “prices ‘fully reflect’ all available information.” The classic categorization of available information, introduced by Roberts [5] and adopted by Fama [6], classifies efficiency as weak-form, when the information set includes past prices, semi-strong, when the information set includes all publicly available information, and

strong-form, when the information set includes all publicly or privately available information. In the so-called tests for return predictability (Fama [7]), the information set available, in addition to past prices (which is the information set in the tests for weak-form market efficiency) may also include firm-specific characteristics (e.g., the firm size, the price-earnings ratio, the book-to-market value and the dividend yield), macroeconomic variables (e.g., variables related to the term structure of interest rates and unexpected inflation) or even calendar effects (Fama [7]). In an efficient market, the results from tests of return predictability should not reject the null hypothesis of no predictability (Efficient Markets Hypothesis).¹

Notwithstanding the beliefs of many academics about the validity of the Efficient Markets Hypothesis, at least for its less strict versions, the so-called technical analysis (*i.e.*, the study of market action for the purpose of forecasting future price trends (Murphy [8])) has been a thriving activity of market participants for more than a century. Indeed, relevant surveys revealed that most market analysts in all kinds of speculative markets (stock markets, futures markets, FOREX, crypto) employ technical trading rules (Taylor and Allen [9]; Billingsley and Chance [10]; Cheung and Chinn [11]).

Although until late 1980s most empirical testing was suggestive of non-rejection of the efficiency hypothesis in its weak and semi-strong versions, during the 1990s, there was an abundance of counterevidence, mainly from examining the predictive power of trading rules of technical analysis. Amongst others, this may be attributed to the enhanced statistical testing methodologies introduced in early 1990s. As Park and Irwin [12] in their review about the tests for the predictive power of technical analysis suggest, the relevant statistical testing procedures may be discriminated as “early” (before early 1990s) and “modern” using as criterion the extent of rigorousness and sophistication of the employed statistical methodology. Furthermore, an additional advantage of this approach is that, to an extent, it overcomes the so-called joint hypothesis problem when testing for weak-form market efficiency using more traditional methods (for details see Milionis and Papanagiotou [13]). There is little doubt that among the most influential “modern” studies was the work of Brock *et al.* [14]. Indeed, Brock *et al.* [14], using daily data of the Dow Jones Index of the New York Stock Exchange for nearly a century, showed that many trading rules have predictive power. This conclusion was then generalized for both developed and emerging capital markets (for example, Bessembinder and Chan [15]; Cai *et al.* [16]; Gençay [17]; Hudson *et al.* [18]; Wong *et al.* [19]).

The statistical tests for the predictability of technical trading rules were certainly facilitated by the development of automated computer-guided trading systems and the use of mathematically well-defined, in the sense of Neftci [20], technical trading rules, which can generate precise trading signals.

¹It must be noted that in the above discussion market efficiency is perceived as static. Allowing efficiency to evolve over time a new version of the hypothesis of efficient markets, the so-called Adaptive Markets Hypothesis has been introduced and the framework for its testing has been extended (Lim and Brooks [45]; Patil and Rastogi [42]; Milionis [43], Milionis [44]).

Among such trading rules used by researchers to test market efficiency the one employed most frequently is the so-called moving average (henceforth MA) rule and it is the one which will be used in this work. It is noted that although technical analysis stresses that a buy or a sell decision is best to be a composite one, being based on as many conducive signals as possible (e.g., volume of trade, convergence-divergence indicators, etc., Murphy [8]), quite often the MA rule, due to the precise signals it generates, is used as a stand-alone method, particularly in the automated trend-following trading systems. In that way, it becomes a purely “mechanical”, rather than technical, trading rule. Even in its “mechanical” use, however, the MA trading rule may have several versions (for instance Pring [21]). In one of its simplest versions, two non-centred, moving averages of different lengths are created from the time series of stock prices:

$$MAL_t = \frac{1}{N} \sum_{i=0}^{N-1} \theta_i B^i P_t$$

$$MAS_t = \frac{1}{M} \sum_{i=0}^{M-1} \theta_i B^i P_t \quad \text{with } N > M$$

where MAL_t represents the relatively longer MA with length N , calculated at time t , MAS_t represents the relatively shorter MA with length M , P_t is the stock price at time t , θ_i are non-time varying parameters, and B is the backward shift operator, i.e. $B^i P_{t-i}$. Buy (sell) signals are then generated when the relatively shorter moving average penetrates the longer moving average from below (above).

A usual way of measuring the predictive power of the trading rule is to test for the statistical significance of the difference between: the mean return of individual trading periods characterized as “buy” (trading periods during which according to the MA trading rule the investor’s capital should remain invested in the market) and the mean return of the whole investment period; the mean return of individual trading periods characterized as “sell” (trading periods during which the investment capital should be liquidated or sold short) and the mean return of “buy” trading periods; or the mean return of “sell” periods and either the mean return of the whole investment period or zero (Brock *et al.* [14]).

Although on several occasions simple t -tests have been used to test for the statistical significance of such differences (e.g., Hudson *et al.* [18]; Wong *et al.* [19]), strictly speaking the application of the t -test which assumes normal, stationary, and time-independent distributions is not legitimate since one or more of these assumptions is very often violated in asset returns (see for example Mills [22]). To overcome this problem, bootstrapping techniques have been suggested initially by Brock *et al.* [14] and this approach has been recognized as the established one and is used by most researchers and for many capital, exchange rate, and cryptocurrency markets (e.g., Brock *et al.* [14]; Hudson *et al.* [18]; Ohnishi *et al.* [23]; Olson [24]; Cai *et al.* [16]; Papailias and Thomakos [25]; Luukka *et al.* [26]; Corbeta *et al.* [27]; Miralles-Quiros *et al.* [28]).

Regarding the predictive power of the MA trading rule per se, more recent stud-

ies have shown a gradual attenuation over time, especially for the most developed stock markets (Park and Irwin [12]; Milionis and Papanagiotou [29]). For the New York Stock Exchange for instance, some empirical studies show that the predictive power of the MA trading rule has been lost completely (e.g., Milionis and Papanagiotou [30]).

This is reasonable given the self-destructive nature of the trading rules: once they are revealed publicly, they lose their predictive power.²

In this paper, an alternative statistical procedure for assessing the predictive performance of technical trading rules is suggested and it is argued that it has some advantages over the existing ones. While in this work the MA trading rule will be exclusively employed, the method can be applied using trading signals produced by any trading rule. More specifically, the trading signals generated by the MA trading rule will serve as interventions acting discontinuously in time on the stochastic process assumed to generate asset returns. In principle, the proposed approach allows for the concurrent estimation of measures of the predictive power of the MA trading rule, such as the ones mentioned above, and at the same time taking into consideration linear interdependencies in asset returns.

The rest of the paper is structured as follows: in Section 2, the proposed statistical methodology is explained; in Section 3, the results of the empirical analysis for daily data of the FT-30 Index of the London Stock Exchange are presented, discussed, and compared with those of other studies; Section 4 summarizes and concludes the paper.

2. Methodology and Data

The intended statistical testing procedure is a modification of the so-called impact assessment models originally developed by Box and Tiao [31]. By the term impact assessment is meant a test of the null hypothesis that an event caused a change in a stochastic process measured by a time series. Events (also called “interventions” in the time series literature) may be represented by binary variables. However, the standard parametric or non-parametric statistical tests which are used to test differences in levels (e.g., the t -test, ANOVA, etc.) cannot be used in serially correlated data, as the fundamental assumption of independence among observations is violated. Moreover, a change in level may not take place instantaneously but gradually. For the inadequacy of the t -test for such cases see for instance Abraham [32]. A method of impact assessment which takes into account serial correlation, as well as gradual level shifts is that of Box and Tiao [31] and allows for simultaneous maximum likelihood estimation of the parameters related to the change as well as those related to the serial correlation. Hence, it can be considered as a generalization of the t -test.

The original Box-Tiao methodology, which will be briefly reviewed here, con-

²However, there is empirical evidence that modified versions of the standard MA trading rules may substantially improve their predictive power especially in less developed markets and in lower capitalization (Ohnishi *et al.* [23]; Papailias and Thomakos [25]; Miralles-Quiros *et al.* [28]).

sists of the following steps:

- i) Definition of the event and exact identification of its onset.
- ii) Univariate ARIMA model-building for the pre-intervention section of the time series under consideration.
- iii) Inclusion of (an) intervention component(s), and re-estimation of the full model.

If the time series under consideration is represented by W_t , impact assessment models are of the following general form:

$$Y_t = f(\mathbf{k}, \mathbf{I}_{it}) + N_t \quad (1)$$

where,

$Y_t = W_t$ itself, or some transformation of W_t to ensure variance stationarity (see Milionis (2004) [41]);

$f(\mathbf{k}, \mathbf{I}_{it})$ measures the effect(s) of the external event(s);

$\mathbf{k}$ is the vector of intervention parameters,

$\mathbf{I}_{it}$ $i = 1, 2$ is the vector of intervention variables with I_{1t} binary variables and I_{2t} the first differences of a binary variable (*i.e.* $I_{2t} = \nabla I_{1t}$, where ∇ is the difference operator); in other words I_{1t} represents a step function (*i.e.* equals zero before $t = T$ and one for t greater or equal to T , where T is the time that the intervention occurs) and I_{2t} represents a pulse function (*i.e.* equals one at $t = T$ and zero elsewhere);

N_t is the generally coloured, noise component.

The noise component is described by an ARIMA model (Box and Jenkins, 1976) in the following form:

$$\Phi(B)N_t = c + \Theta(B)\alpha_t$$

where:

α_t is white noise,

c is a constant,

B is the backward shift operator,

$\Phi(B) = 1 - \varphi_1 B - \dots - \varphi_p B^p$ is the autoregressive polynomial of order p in B ,

$\varphi_1, \varphi_2, \dots, \varphi_p$ are constant coefficients,

$\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ is the moving average polynomial of order q in B ,

$\theta_1, \theta_2, \dots, \theta_q$ are constant coefficients.

If the process N_t is integrated of order d , then $\Phi(B)$ can be made stationary by differencing d times so that:

$$\varphi(B) = \nabla^d \Phi(B)$$

with all the roots of $\varphi(B)$ outside the unit circle.

The vector of intervention parameters $\mathbf{k}$ can be expressed in terms of vectors $\boldsymbol{\omega}$, $\boldsymbol{\delta}$, assuming a single external intervention, as follows (Box and Tiao [31]):

$$f(\mathbf{k}, I_{it}) = f(\boldsymbol{\omega}, \boldsymbol{\delta}, I_{it}) = \frac{\Omega(B)}{\delta(B)} I_{it} \quad (2)$$

where:

$$\Omega(B) = \omega_0 + \omega_1 B + \dots + \omega_r B^r,$$

$$\delta(B) = 1 - \delta_1 B - \dots - \delta_s B^s \quad \text{and}$$

$\omega_0, \omega_1, \dots, \omega_r, \delta_1, \delta_2, \dots, \delta_s$ are constant parameters following certain restrictions (Box and Tiao, 1975).

The $\Omega(B)$ polynomial reflects the total change in the level of the series caused by the intervention, while the $\delta(B)$ polynomial expresses the rate at which the series approach the new equilibrium level.

The intervention parameter estimates ω_i , δ_i and the stochastic parameter estimates φ_i , θ_i are asymptotically independent of each other (Box and Tiao [31]; Box *et al.* [33]), hence, the information matrix is block-diagonal. The inverse of the information matrix can be used as an estimator of the asymptotic covariance matrix of the parameter estimates. The estimates of both the intervention and the noise parameters can be obtained simultaneously by maximum likelihood (details in Box and Tiao [31]).

Thus far the method of Box-Tiao has been used to study the impact of events which occur on *continuous* lags. However, the method may be extended to cover the impact of events which occur discontinuously in time, on the condition that the effect of such events is assumed to be *transient*. The assumption about the transient character of an event which acts discontinuously is necessary to justify the use of the same ARIMA model for the whole series.³ Under such circumstances, for the case of discontinuous impacts, the values of the parameter vectors $\boldsymbol{\omega}$, $\boldsymbol{\delta}$ will reflect the *average* effect of the factor which acts discontinuously on a time series.

Another methodological point which should be discussed is that of the specification of the polynomials $\Omega(B)$ and $\delta(B)$. Although in their seminal paper Box and Tiao [31] provide several examples of intervention models, they are not suggestive about the methodology for the identification of the proper specification of $\Omega(B)$ and $\delta(B)$. That makes it difficult to specify the intervention component in those cases for which the form of this component cannot be specified a-priori, on the basis of an existing theory. The only methodological approach for the selection of the proper intervention component that exists in the literature is the three-stage procedure due to McCleary and Hay [34], which will be adopted in this work. In principle this approach is indirectly based on the fact that although in theory the order of $\Omega(B)$ and $\delta(B)$ polynomials can be arbitrarily high, both the general principle of time series analysis for parsimonious models (see Box and Jenkins [35]) and practical experience suggest that for most of the cases

³That means that the proposed method cannot be applied for events for which the character of their impact is either abrupt permanent or gradient permanent in the terminology of Liu [36], and McCleary and Hay [34].

one ω and one δ parameters suffice to fit most intervention effects. Based on this argumentation, in the first stage, McCleary and Hay [34] suggest the following model for the intervention component (see also Liu [36]):

$$\frac{\omega_0}{1-\delta_1 B} I_{2t} = \frac{\omega_0}{1-\delta_1 B} (1-B) I_{1t} \quad \text{with } |\delta_1| < 1 \quad (3)$$

Such a specification represents a “sudden-temporary” effect (see McCleary and Hay [34], as well as Liu [36] for a more detailed discussion). If δ_1 is not statistically significant, or if it is close to unity this specification is ruled out.

In the second stage, the following specification for the intervention component is estimated, representing a “gradual permanent” effect:

$$\frac{\omega_0}{1-\delta_1 B} I_{1t} \quad \text{with } |\delta_1| < 1 \quad (4)$$

If δ_1 is not statistically significant, in the third and final stage a “sudden permanent” effect is assumed:

$$\omega_0 I_{1t} \quad (5)$$

In this work the old Box and Tiao [31] methodology, with the modifications and enhancements described above, will be applied to examine whether or not, the technical trading rule of moving average has predictive power for the FT-30 Index of the London Stock Exchange using daily closing prices for the period 1935-1994. This can be performed in the following way: as explained in the previous section, the role of the MA trading rule is to classify each individual trading period (in this case each individual day) as “buy” or “sell”. The “buy” days will be considered as the discontinuous interventions. Therefore, the intervention variable I_{1t} will be a binary series of “1” and “0” corresponding to the “buy” and “sell” days respectively. It is noted that the time origin for which the intervention variable will be assigned its first value (*i.e.*, a “1” or a “0”) is the period (day) at which the MA trading rule generates its first signal, which is apparently equal to the length of the long moving average considered each time. It is important to note that, in contrast to the existing testing methodologies for which the effect of the trading rule on asset returns is restricted to be constant, the proposed methodology is capable to allow more general patterns for that effect (e.g., the intervention models (3) and (4) discussed previously). This enhanced flexibility that the proposed methodology offers is consistent with the principles of technical analysis (see Pring [21]). Furthermore, it has potentially important implications, as it may result in improved testing procedures and assessment regarding the existence and the extent of the predictive power of trading rules of technical analysis.

The total time period that the data cover in this work was purposely selected, so as to approximately coincide with that of previous studies of other scholars, who examined the predictive performance of the MA trading rule for the same stock exchange index but using different techniques (Markellos [37]; Mills [22]). This will make it possible to directly compare the results derived following the statistical approach suggested above with those of the previous aforementioned

studies. Moreover, the total time period stops at 1994, *i.e.* roughly at the end of the so-called “early” period, as defined by Park and Irwin [12], as from the second half of the 90ies the predictive power of the MA trading rule almost disappeared. The total time will be divided into three 20-year sub-periods: 1935–1954, 1955–1974 and 1975–1994 and each sub-period will also be examined. The particular combinations of long and short moving averages for the application of the moving average trading rules which will be employed, are deliberately taken to coincide with those of the previously mentioned studies and are the following: 1–50, 1–150, 5–150, 1–200 and 2–200, where the first number indicates the length of the short moving average and the second number, the length of the long moving average.^{3,4}

3. Empirical Analysis

First, it is useful to present and comment on some summary descriptive statistics relating to the dataset. Such statistics are shown in **Table 1**, which refer to the daily returns of the FT-30 Index expressed as logarithmic differences of successive closing prices of the Index. From the results of **Table 1**, it is evident that over the whole period, as well as in any sub-period, index returns are non-normal, asymmetric (negatively skewed) and leptokurtic, as is also found in Mills [22]. Further, the first order autocorrelation coefficient for index returns ($\rho(1)$) is significant at 5% level for the whole period, as well as for all sub-periods (it is noted, however, that significance testing for autocorrelations should be taken as indicative but not exact, for index returns are not normal, as documented above. In addition, the Ljung-Box statistic (denoted as LBQ), which is a portmanteau statistic for the test of significance of more than one autocorrelation coefficient jointly (Ljung and Box [38]) was used on the residuals of an AR(1) model of index returns. Again higher-than-first-order autocorrelation significant at the 5% level exists for all time periods. Moreover, there is strong evidence of substantial interdependence in the squares of index returns, as the last two columns of **Table 1** reveal. The existence of such interdependence in higher moments of index returns is also conducive to the existence of trading rule abnormal returns (Neftci [20]).

³It must be noted that the whole approach of using the same data and the same selective trading rules as in Mills [22] inevitably makes our results susceptible to data snooping. This entails overstated significance levels for hypothesis testing hence, overoptimistic conclusions about the predictive power of trading rules. The problem may be tackled by using the so-called “data snooping adjusted p-values” provided by White’s bootstrap reality check methodology (White [46]). Indeed, results in similar studies in which only popular trading rules were used selectively (e.g. Brock *et al.* [14]) have been shown to suffer by data snooping. However, the whole matter is still controversial as under certain conditions even the “data snooping adjusted p-values” may also be biased (Hansen [47] [48]). As one of our main goals in this work was the direct comparison of the results from the proposed methodology with those of the existing traditional methods, the issue of data snooping was of lesser importance, and we have not dealt with it any further.

⁴Caution is needed on how trading rules’ results are linked to market efficiency testing. As Milionis and Papanagiotou [49] note the existence of predictive power in an MA trading rule with specific combinations of long and short moving averages does not necessarily mean a rejection of the weak-form market efficiency hypothesis.

Table 1. Summary statistics for daily index returns.

Time period	mean*10 ⁻⁴	std	skewness	kurtosis	KS-test	autocorrelation			
						Returns		Squared Returns	
						$\rho(1)$	LBQ(10) (in res. of AR(1) model)	$\rho^2(1)$	LBQ(10) (in res. of AR(1) model)
1935-1994	2.1912 (**) (2.67)	0.0104	-0.1438	11.5320	0.4816 (**)	0.0983 (**)	72.44 (**)	0.4202 (**)	431.06 (**)
1935-1954	1.3056 (1.52)	0.0060	-0.4332	17.1266	0.4879 (**)	0.3441 (**)	47.24 (**)	0.2921 (**)	437.39 (**)
1955-1974	0.1103 (0.07)	0.0102	-0.1844	10.3541	0.4826 (**)	0.0953 (**)	30.13 (**)	0.3886 (**)	88.44 (**)
1975-1995	5.2711 (**) (2.94)	0.0126	-0.0956	7.4298	0.4794 (**)	0.0464 (**)	45.12 (**)	0.4324 (**)	169.20 (**)

The Kolmogorov-Smirnov (KS-test) critical value is 0.0110 (5%). X_9^2 critical values (for LBQ-statistic): 14.68 (10%), 16.91 (5%). One asterisk (*) indicates significance at 10% level. Two asterisks (**) indicate significance at 5% level.

As will become evident below, it is to the methodological benefit of the empirical analysis if any linear dependencies in index returns are initially ignored, *i.e.*, if it is assumed that the logarithmic differences of index prices are white noise. It will be further assumed that in Equation (2) above $r = s = 0$, so that $\Omega(B) = \omega_0$ and $\delta(B) = 1$.

Combining Equation (1) and Equation (2) and taking the above assumption into consideration it is easily seen that index returns may be described by the following model:

$$R_t = c + \omega_0 I_{1t} + a_t \quad (6)$$

In Equation (6), which will be called the “benchmark model” hereafter, a_t is temporarily assumed to be white noise, the parameter ω_0 represents the relative average increase (if any) on returns during the “buy” periods, as compared to the “sell” periods, while the parameter c represents the average return of “sell” days. Therefore, it makes sense to compare the results on the statistical significance of the parameter ω_0 with the results from an ordinary t -test for the mean difference of “buy-sell” for index returns and results on the statistical significance of parameter c to those of the ordinary t -test for the statistical significance of the mean return of “sell” days. This comparison aims to serve one more purpose: to show the advantage in terms of statistical efficiency of the proposed methodology (“benchmark” model) as compared to the t -test for some cases. Indeed, with the benchmark model more efficient use of the available data is made, as the *whole* data set is used for the concurrent estimation of both c and ω_0 using equation (6), in contrast to the use of two distinct t -tests, in which only a subset of the data is used for the “sell” days. It is noted that the “benchmark” model is simply a linear regression model with a dummy explanatory variable and it is stressed that for the present case this model will only be used for the sake of comparisons with the t -test, as from the results of autocorrelation in **Table 1** it is evident that the residuals of this model are expected to be autocorrelated and, hence, strictly speaking, any

significance testing of model parameters is not valid under such circumstances.

The effect of linear dependencies on index returns, which does exist as evidenced by the results of **Table 1**, as well as non-constant effects of the trading rule, may be taken into account by considering the model:

$$R_t = c + f(\omega_0, \delta_1, I_{it}) + \frac{\Theta(B)}{\Phi(B)} a_t \quad (7)$$

In model (7), the specification (3) will be initially used for $f(\omega_0, \delta_1, I_{it})$ and, if necessary, the specifications (4) and (5) will be used subsequently. Model (7) will be called the “full” model henceforth and the estimates of c , ω_0 and δ_1 will be purified from the effect of linear interdependencies in index returns. It is noted that for the intervention components specifications (3) and (4) both intervention parameters should be significant.

Another building block of our analysis is the univariate ARMA models that have to be created for index returns and for each case. This task is useful because ARMA models can be used as a first step in the model building of an impact assessment model. Moreover, in that way, it will be possible to assess the improvement in the explanatory power of the “full” models, which will include the intervention parameters, in comparison to that of the univariate models. The model building procedure followed for the creation of the univariate models for each time period is the classical one suggested by Box and Jenkins [35]. Both the pattern of the autocorrelation function, as well as that of the partial autocorrelation function, in all cases were suggestive of clearly stationary index returns. Moreover, pure moving average processes of order greater than one were preferable to purely autoregressive, or mixed processes (further details are available from the authors on request). This is in accordance with the empirical findings of other studies (e.g., Milionis *et al.* [39]; Milionis and Moschos [40]). It differs, however, from the empirical finding of Mills [22] who fits AR(2) or even AR(1) models to almost identical data. In the course of the empirical analysis of this study, AR(2) models did not render white noise residuals. It is also noted that autoregressive conditional heteroscedasticity (ARCH), which according to the results of **Table 1** may be present in the data, has not been taken into consideration in this initial stage of our research.

The results of the empirical analysis, based on the parameter estimates using the “benchmark” model as well as the “full” models, for the total time period and each sub-period, for all trading rules are presented in **Table 2** to **Table 5**. Together with the values and the associated t -statistics for the parameters c , ω_0 and δ_1 , the results from ordinary t -tests are also quoted in these tables.⁵ Additional pieces of information also included in each of the **Table 2** to **Table 5** are the values of the coefficient of determination (R^2) for the univariate models as well as for the “full” models selected after following the three stage procedure proposed by McCleary and Hay [34]. The value of this coefficient is directly related to the de-

⁵Computationally wise for the estimation of model parameters an old, specialized routine of maximum likelihood estimation for multivariate time series models, included in sub-programme BMDP2T of the BMDP statistical software, was used.

gree of linear interdependencies in the index returns series; hence it may be taken as a measure of the departure from randomness; the latter is among the fundamental assumptions for the application of the t -test. Moreover, the value of R^2 will also be useful in assessing the improvement in the explanatory power of the models describing index returns due to the separation of returns into “buy” and “sell” returns, as a result of the application of the MA trading rule.

The final (*i.e.* after the inclusion of the intervention component) model specification for the noise component for each case, the estimates of model parameters, as well as the corresponding t -statistics, are presented in **Table 6**.

Table 2. Results for the 1935-1994 time period (t -statistics in parentheses; ### denotes absolute values of t -statistics > 100).

Trading Rule	“Benchmark” model		Ordinary <i>t</i> -test		Univariate ARIMA Model	<i>R</i> ²	“Full” model							
							Intervention Component							
							$\frac{\omega_0}{1-\delta_1 B}(1-B)I_{1t}$		$\frac{\omega_0}{1-\delta_1 B}I_{1t}$		$\omega_0 I_{1t}$			
	<i>c</i> · 10 ^{−3}	ω_0 · 10 ^{−3}	Mean “sell”. 10 ^{−3}	Mean “buy-sell” difference. 10 ^{−3}	<i>R</i> ²	<i>c</i> · 10 ^{−3}	ω_0 · 10 ^{−3}	δ_1	<i>c</i> · 10 ^{−3}	ω_0 · 10 ^{−3}	δ_1	<i>c</i> · 10 ^{−3}	ω_0 · 10 ^{−3}	
1-50	−0.437 (**) (−3.40)	1.111 (**) (6.64)	−0.437 (**) (−3.06)	1.111 (**) (6.64)	0.0161	0.0170	0.217 (**) (2.32)	1.393 (**) (4.85)	0.993 (**) (92.66)	−0.234 (−1.63)	0.763 (**) (2.47)	−0.440 (−0.80)	−0.241 (−1.64)	0.778 (**) (4.00)
1-150	−0.164 (−1.23)	0.611 (**) (3.60)	−0.164 (−1.08)	0.612 (**) (3.60)	0.0161	0.0169	0.214 (**) (2.19)	0.960 (**) (2.63)	0.976 (**) (71.43)	0.030 (0.19)	0.453 (1.32)	−0.393 (−0.39)	−0.038 (−0.02)	0.346 (**) (2.81)
5-150	−0.054 (−0.41)	0.434 (**) (2.55)	−0.054 (−0.36)	0.435 (**) (2.55)	0.0161	0.0169	0.214 (**) (2.17)	0.930 (**) (2.14)	0.988 (**) (82.02)	0.059 (0.37)	0.237 (0.74)	0.266 (0.30)	0.054 (0.34)	0.309 (**) (2.47)
1-200	−0.170 (−1.25)	0.611 (**) (3.57)	−0.170 (−1.11)	0.611 (**) (3.56)	0.0165	0.0174	0.213 (**) (2.17)	1.375 (**) (2.99)	0.989 (**) (85.13)	−0.023 (−0.17)	0.668 (**) (3.21)	−0.970 (**) (−70.60)	−0.026 (−0.17)	0.362 (**) (2.96)
2-200	−0.123 (−0.91)	0.538 (**) (3.14)	−0.123 (−0.81)	0.538 (**) (3.14)	0.0165	0.0174	0.213 (**) (2.17)	1.533 (**) (3.25)	0.969 (**) (67.78)	0.016 (0.10)	0.316 (0.87)	0.099 (0.10)	−0.019 (−0.12)	0.373 (**) (2.83)
	1	2	3	4	5	6	7	8	9	10	11	12	13	14

One asterisk (*) indicates significance at 10% level. Two asterisks (**) indicate significance at 5% level.

Table 3. Results for the 1935-1954 time period (t -statistics in parentheses; ### denotes absolute values of t -statistics > 100).

Trading Rule	“Benchmark” model		Ordinary t -test	Univariate ARIMA Model	“Full” model									
					Intervention Component									
					$\frac{\omega_0}{1-\delta_1 B}(1-B)I_{1t}$			$\frac{\omega_0}{1-\delta_1 B}I_{1t}$			$\omega_0 I_{1t}$			
	$c \cdot 10^{-3}$	$\omega_0 \cdot 10^{-3}$	Mean “sell”. 10^{-3}	Mean “buy-sell” difference. 10^{-3}	R^2	R^2	$c \cdot 10^{-3}$	$\omega_0 \cdot 10^{-3}$	δ_1	$c \cdot 10^{-3}$	$\omega_0 \cdot 10^{-3}$	δ_1	$c \cdot 10^{-3}$	$\omega_0 \cdot 10^{-3}$
1-50	-0.619 (**) (-4.49)	1.219 (**) (6.93)	-0.619 (**) (-3.40)	1.219 (**) (6.89)	0.1337	0.1353	-0.045 (-0.31)	-0.844 (**) (-2.13)	-0.182 (-0.43)	-0.353 (-1.59)	0.000 (0.40)	0.993 (**) (###)	-0.208 (-1.07)	0.331 (**) (2.19)

Continued

1-150	-0.329 (**) (-2.33)	0.708 (**) (3.96)	-0.329 (*) (-1.75)	0.708(**) (3.96)	0.1356	0.1367	0.112 (1.03)	-1.065 (**) (-1.97)	-0.217 (-0.48)	-0.017 (-0.11)	0.168 (0.55)	0.484 (0.53)	-0.180 (-1.00)	0.278 (**) (2.00)
5-150	-0.172 (-1.22)	0.456 (**) (2.55)	-0.171 (-0.92)	0.455(**) (2.54)	0.1356	0.1370	0.113 (1.04)	-0.372 (-0.76)	-0.396 (-0.56)	-0.254 (-0.96)	0.000 (0.20)	0.992 (**) (###)	-0.258 (-1.43)	0.322 (**) (2.32)
1-200	-0.322 (**) (-2.25)	0.688 (**) (3.82)	-0.322 (*) (-1.67)	0.688 (**) (3.81)	0.1371	0.1383	0.111 (1.01)	-0.768 (-1.30)	0.142 (0.18)	-0.370 (-1.51)	0.000 (0.75)	0.996 (**) (###)	-0.201 (1.11)	0.287 (**) (2.08)
2-200	-0.303 (*) (-1.85)	0.660 (**) (3.67)	-0.304 (-1.57)	0.660 (**) (3.67)	0.1371	0.1388	-0.0014 (-0.01)	0.580 (**) (1.98)	0.991 (**) (###)	-0.272 (-1.52)	0.620 (**) (2.74)	-0.952 (**) (###)	-0.280 (*) (-1.66)	0.313 (**) (2.06)
	1	2	3	4	5	6	7	8	9	10	11	12	13	14

One asterisk (*) indicates significance at 10% level. Two asterisks (**) indicate significance at 5% level.

Table 4. Results for the 1955-1974 time period (*t*-statistics in parentheses; ### denotes absolute values of *t*-statistics > 100).

Trading Rule	“Benchmark” model		Ordinary <i>t</i> -test		Univariate ARIMA Model	“Full” model									
						Intervention Component									
						$\frac{\omega_0}{1-\delta_1 B}(1-B)I_{1t}$					$\frac{\omega_0}{1-\delta_1 B}I_{1t}$			$\omega_0 I_{1t}$	
	$c \cdot 10^{-3}$	$\omega_0 \cdot 10^{-3}$	Mean “sell”. 10^{-3}	Mean “buy-sell” difference. 10^{-3}	R^2	R^2	$c \cdot 10^{-3}$	$\omega_0 \cdot 10^{-3}$	δ_1	$c \cdot 10^{-3}$	$\omega_0 \cdot 10^{-3}$	δ_1	$c \cdot 10^{-3}$	$\omega_0 \cdot 10^{-3}$	
1-50	-0.519 (**) (-2.62)	1.058 (**) (3.62)	-0.517 (**) (-2.28)	1.057 (**) (3.61)	0.0219	0.0230	0.022 (0.13)	1.059 (1.57)	5.838 (0.10)	-0.422 (*) (-1.73)	1.203(**) (1.97)	-0.493 (-0.73)	-0.445 (*) (-1.84)	0.834 (**) (2.51)	
1-150	-0.503 (**) (-2.30)	0.925 (**) (3.14)	-0.503 (**) (-2.03)	0.925 (**) (3.14)	0.0211	0.0224	0.006 (0.04)	0.722 (1.08)	-0.700 (*) (-1.90)	-0.415 (*) (-1.78)	0.514 (*) (1.76)	-0.912 (**) (-38.04)	-0.441 (**) (-1.96)	0.786 (**) (2.38)	
5-150	-0.458 (**) (-2.10)	0.843 (**) (2.86)	-0.458 (*) (-1.86)	0.843 (**) (2.86)	0.0211	0.0224	-0.019 (-0.12)	1.858 (1.54)	0.153 (0.24)	-0.450 (*) (-1.86)	0.455 (0.72)	-0.208 (-0.13)	-0.445 (*) (-1.87)	0.783 (**) (2.36)	
1-200	-0.527 (**) (-2.41)	0.985 (**) (3.33)	-0.528 (**) (-2.15)	0.985 (**) (3.33)	0.0210	0.0220	0.003 (0.02)	0.816 (0.80)	0.074 (0.06)	-0.453 (**) (-1.96)	1.449 (**) (2.89)	-0.900 (**) (-32.19)	-0.476 (**) (-2.00)	0.892 (**) (2.77)	
2-200	-0.489 (**) (-2.23)	0.913 (**) (3.08)	-0.489 (**) (-1.99)	0.913 (**) (3.08)	0.0210	0.0220	0.006 (0.04)	1.039 (0.87)	0.223 (0.19)	-0.462 (*) (-1.94)	1.156 (1.21)	-0.334 (-0.31)	-0.462 (**) (-1.96)	0.868 (**) (2.68)	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	

One asterisk (*) indicates significance at 10% level. Two asterisks (**) indicate significance at 5% level.

Table 5. Results for the 1975-1994 time period (t -statistics in parentheses; ### denotes absolute values of t -statistics > 100).

Trading Rule	“Benchmark” model		Ordinary <i>t</i> -test		Univariate ARIMA Model	<i>R</i> ²	“Full” model							
							Intervention Component							
							$\frac{\omega_0}{1-\delta_1 B}(1-B)I_{1t}$			$\frac{\omega_0}{1-\delta_1 B}I_{1t}$		$\omega_0 I_{1t}$		
	<i>c</i> · 10 ^{−3}	<i>ω</i> ₀ · 10 ^{−3}	Mean “sell”. 10 ^{−3}	Mean “buy-sell” difference. 10 ^{−3}	<i>R</i> ²	<i>R</i> ²	<i>c</i> · 10 ^{−3}	<i>ω</i> ₀ · 10 ^{−3}	<i>δ</i> ₁	<i>c</i> · 10 ^{−3}	<i>ω</i> ₀ · 10 ^{−3}	<i>δ</i> ₁	<i>c</i> · 10 ^{−3}	<i>ω</i> ₀ · 10 ^{−3}
1-50	−0.011 (−0.04)	0.958 (**) (2.62)	−0.011 (−0.04)	0.958 (**) (2.61)	0.0106	0.0117	0.565 (**) (2.67)	1.617 (**) (2.57)	0.996 (**) (74.60)	0.268 (0.90)	1.195(**) (2.90)	−0.715 (**) (−2.77)	0.341 (1.02)	0.705 (**) (2.61)
1-150	0.388 (1.21)	0.054 (0.15)	0.389 (1.17)	0.053 (0.15)	0.0047	0.0050	0.428 (**) (2.30)	1.175 (1.52)	0.969 (**) (70.45)	0.483 (**) (2.45)	−0.167 (−0.85)	−0.992 (**) (###)	0.602 (1.45)	−0.259 (−0.64)
5-150	0.517 (1.54)	−0.135 (−0.38)	0.517 (1.59)	−0.135 (−0.37)	0.0047	0.0051	0.429 (**) (2.31)	0.9938 (1.25)	0.987 (**) (73.61)	1.430 (**) (2.27)	−0.032 (*) (−1.78)	0.978 (**) (###)	0.682 (1.62)	−0.376 (−0.94)
1-200	0.481 (1.49)	−0.030 (−0.08)	0.481 (1.38)	−0.030 (−0.08)	0.0054	0.0055	0.461 (**) (2.64)	1.882 (**) (2.65)	0.986 (**) (72.11)	0.280 (1.10)	0.909 (**) (2.59)	−0.970 (**) (###)	0.571 (1.56)	−0.158 (−0.41)
2-200	0.552 (*) (1.73)	−0.132 (−0.37)	0.552 (*) (1.66)	−0.132 (−0.37)	0.0054	0.0055	0.460 (**) (2.64)	1.678 (**) (2.23)	0.988 (**) (73.71)	1.556 (**) (2.77)	−0.034 (**) (−2.01)	0.978 (**) (###)	0.641 (*) (1.88)	−0.259 (−0.67)
	1	2	3	4	5	6	7	8	9	10	11	12	13	14

One asterisk (*) indicates significance at 10% level. Two asterisks (**) indicate significance at 5% level.

Table 6. Specification of the noise component of the “full” model and parameter estimates (t -statistics in parentheses, e_t = white noise).

Time Period	Trading Rule	N_t
1935-1994	1-50	$N_t = 0.088e_{t-1} - 0.017e_{t-6} - 0.020e_{t-7} + 0.024e_{t-14} + e_t$ (10.77) (-2.09) (2.44) (2.97)
	1-150	$N_t = 0.091e_{t-1} - 0.016e_{t-7} + 0.031e_{t-9} + 0.027e_{t-14} + e_t$ (11.22) (-1.97) (3.84) (3.36)
	5-150	$N_t = 0.093e_{t-1} - 0.015e_{t-7} + 0.031e_{t-9} + 0.026e_{t-14} + e_t$ (11.39) (-1.95) (3.86) (3.32)
	1-200	$N_t = 0.093e_{t-1} - 0.015e_{t-7} + 0.031e_{t-9} + 0.026e_{t-14} + e_t$ (11.39) (-1.95) (3.86) (3.32)
	2-200	$N_t = 0.092e_{t-1} - 0.015e_{t-7} + 0.032e_{t-9} + 0.026e_{t-14} + e_t$ (11.27) (-1.95) (3.89) (3.24)
1935-1954	1-50	$N_t = 0.317e_{t-1} + 0.195e_{t-2} + 0.061e_{t-3} + 0.041e_{t-14} + e_t$ (22.29) (13.25) (4.36) (2.93)
	1-150	$N_t = 0.321e_{t-1} + 0.203e_{t-2} + 0.072e_{t-3} - 0.038e_{t-13} + e_t$ (22.46) (13.81) (5.08) (2.79)
	5-150	$N_t = 0.322e_{t-1} + 0.204e_{t-2} + 0.073e_{t-3} - 0.038e_{t-13} + e_t$ (22.57) (13.87) (5.13) (2.81)
	1-200	$N_t = 0.321e_{t-1} + 0.204e_{t-2} + 0.067e_{t-3} - 0.043e_{t-13} + e_t$ (22.27) (13.77) (4.69) (2.99)
	2-200	$N_t = 0.323e_{t-1} + 0.207e_{t-2} + 0.072e_{t-3} - 0.037e_{t-13} + e_t$ (22.45) (14.00) (5.05) (2.74)

Continued

1955-1974	1-50	$N_t = 0.091e_{t-1} - 0.031e_{t-4} - 0.030e_{t-7} + 0.033e_{t-9} + 0.044e_{t-13} + e_t$ (6.38) (2.17) (2.17) (2.34) (3.06)
	1-150	$N_t = 0.089e_{t-1} - 0.028e_{t-7} + 0.033e_{t-9} + 0.044e_{t-13} + e_t$ (6.28) (1.96) (2.32) (3.08)
	5-150	$N_t = 0.089e_{t-1} - 0.028e_{t-7} + 0.033e_{t-9} + 0.044e_{t-13} + e_t$ (6.28) (1.96) (2.33) (3.08)
	1-200	$N_t = 0.087e_{t-1} - 0.028e_{t-4} - 0.028e_{t-7} + 0.031e_{t-9} + 0.043e_{t-13} + e_t$ (6.11) (1.98) (1.97) (2.15) (2.98)
	2-200	$N_t = 0.089e_{t-1} - 0.028e_{t-4} - 0.028e_{t-7} + 0.031e_{t-9} + 0.043e_{t-13} + e_t$ (6.19) (1.97) (1.99) (2.17) (3.01)
1975-1994	1-50	$N_t = 0.035e_{t-1} + 0.038e_{t-4} + 0.029e_{t-9} + 0.066e_{t-10} + 0.029e_{t-14} + e_t$ (2.48) (2.67) (2.06) (4.68) (2.08)
	1-150	$N_t = 0.034e_{t-4} + 0.036e_{t-9} + 0.045e_{t-10} + e_t$ (2.39) (2.56) (3.18)
	5-150	$N_t = 0.034e_{t-4} + 0.037e_{t-9} + 0.045e_{t-10} + e_t$ (2.40) (2.51) (3.20)
	1-200	$N_t = 0.034e_{t-4} - 0.034e_{t-7} + 0.037e_{t-9} + 0.040e_{t-10} + e_t$ (2.37) (2.42) (2.61) (2.83)
	2-200	$N_t = 0.034e_{t-4} - 0.034e_{t-7} + 0.037e_{t-9} + 0.040e_{t-10} + e_t$ (2.37) (2.38) (2.61) (2.82)

Based on the results presented in **Table 2** to **Table 6** the following interesting remarks can be made:

a) For all time periods the values of the parameters c and ω_0 estimated by using the “benchmark” model (columns (1), (2)) are identical to the corresponding values for the mean return of “sell” days and the mean “buy-sell” difference respectively, as is evident from the results for the ordinary t -test (columns (3) and (4) respectively).

b) Referring to the “benchmark” model, though the t -statistics associated with the ω_0 parameter (column (2)) are identical to those reported in the results of the t -test for the significance of the “buy-sell” difference for all time periods (column (4)), the t -statistics associated with the parameter c (column (1)) are *constantly higher* in absolute value as compared to those of the mean return of “sell” days with the ordinary t -test (column (3)), implying more efficient estimation (lower standard errors for c , as compared to the standard errors associated with the ordinary t -test for the significance of the mean return of “sell” days). This enhanced estimation efficiency is due to the increased degrees of freedom (use of the whole data set) and the concurrent estimation of both parameters obtained by using the “benchmark” model.

c) In all cases the intervention component for the first stage (specification (3)) is ruled out as the value of δ_1 is either not statistically significant or very close to unity.

d) In all cases the intervention component for the second stage is ruled out as the value of δ_1 is either not statistically significant or very close to the bounds of

model stability (± 1). Therefore, it is necessary to proceed with the estimation of the third stage (specification (5) for the intervention component).

e) It is of importance to compare the estimates as well as the associated t -statistics of the parameters c and ω_0 derived from the “benchmark” model (columns (1) and (2)) and the 3rd stage “full” model (columns (13), (14)). In all cases, except for the last sub-period, the estimates of c are less negative with the “full” model than with the “benchmark” model. On the other hand, the estimates of ω_0 are always less positive with the “full” model than with the “benchmark” model. This is due to the effect of autocorrelation which is incorporated in the “full” model, but not in the “benchmark” model. The effect is most pronounced in the first sub-period. This can be easily confirmed by the comparison of the estimates of c and ω_0 derived by each model for the total time period (Table 2) and the first sub-period (Table 3). Indeed, while the estimates of c obtained using the “benchmark” model for all trading rules are less negative for the total time period, with the “full” model it is the other way round. Similarly, the ω_0 estimates with the “benchmark” model for all cases are more positive for the first sub-period, but with the “full” model it is the other way round. This can be easily explained by looking at the R^2 values of the univariate model in each time period (column 5). The first sub-period has by far the highest R^2 value; hence, the effect of autocorrelation is expected to be the highest for this time period.

f) Despite the differences in the results of the “benchmark” and the “full” model, it is apparent that in all time periods qualitatively the conclusions regarding the predictive power of the trading rules are, by and large, the same. Of particular importance is the fact that although for the first two sub-periods (1935-1954 and 1955-1974) the hypothesis of weak-form efficiency is rejected by most trading rules, for the last sub-period (1975-1994) the hypothesis of weak-form efficiency is not rejected in all, but one tests.

g) The model specification for the noise component (Table 6) in all cases is a pure moving average model, as a result of smaller values of the Akaike criterion. This is not unusual for stock index returns (see for instance Milionis *et al.* [39]; Milionis and Moschos [40]. Further, in several model specifications the inclusion of scattered MA components at relatively high lags (e.g., lags 9, 10, 13, 14) may be a reflection of the attempt of a linear stochastic model, as are the Box-Jenkins models, to encompass non-linearities such as autoregressive conditional heteroscedasticity (see Milionis, 2004 for a detailed discussion on that matter).

It is also of great interest to compare the results of this study with those of other studies, which used data for the same time period but followed different methodological approaches. More specifically, Mills [22] followed the established bootstrapping approach, while Markellos (1999) suggested and used the co-integration cumulative profit (CCP) test. To facilitate comparisons, the results for each time period of the three different methodologies have been gathered together and are presented in Table 7 to Table 10.

Table 7. Comparative results for the 1935-1994 time period.

Trading Rule	This study		Mills [22]		Markellos [37]
	c	ω	c	ω	CCP test
1-50	NS	**	$\emptyset$	$\emptyset$	**
1-150	ns	**	$\emptyset$	$\emptyset$	**
5-150	Ns	**	$\emptyset$	$\emptyset$	**
1-200	Ns	**	$\emptyset$	$\emptyset$	**
2-200	ns	**	$\emptyset$	$\emptyset$	**

One asterisk (*) indicates significance at 10% level. Two asterisks (**) indicate significance at 5% level. NS denotes no statistical significance at 10% level. $\emptyset$ denotes no available results. CCP stands for co-integration cumulative profit test.

Table 8. Comparative results for the 1935-1954 time period.

Trading Rule	This study		Mills [22]		Markellos [37]
	c	ω	c	ω	CCP test
1-50	Ns	**	**	**	Ns
1-150	Ns	**	**	**	**
5-150	Ns	**	$\emptyset$	$\emptyset$	**
1-200	Ns	**	*	**	**
2-200	*	**	$\emptyset$	$\emptyset$	**

One asterisk (*) indicates significance at 10% level. Two asterisks (**) indicate significance at 5% level. NS denotes no statistical significance at 10% level. $\emptyset$ denotes no available results. CCP stands for co-integration cumulative profit test.

Table 9. Comparative results for the 1955-1974 time period.

Trading Rule	This study		Mills [22]		Markellos [37]
	c	ω	c	ω	CCP test
1-50	*	**	**	**	**
1-150	**	**	$\emptyset$	$\emptyset$	**
5-150	*	**	**	**	**
1-200	**	**	**	**	**
2-200	**	**	$\emptyset$	$\emptyset$	**

One asterisk (*) indicates significance at 10% level. Two asterisks (**) indicate significance at 5% level. NS denotes no statistical significance at 10% level. $\emptyset$ denotes no available results. CCP stands for co-integration cumulative profit test.

Table 10. Comparative results for the 1975-1994 time period.

Trading Rule	This study		Mills [22]		Markellos [37]
	c	ω	c	ω	CCP test
1-50	NS	**	Ns	**	ns
1-150	Ns	Ns	Ns	Ns	**
5-150	Ns	Ns	$\emptyset$	$\emptyset$	Ns
1-200	Ns	Ns	ns	Ns	Ns
2-200	Ns	ns	$\emptyset$	$\emptyset$	ns

One asterisk (*) indicates significance at 10% level. Two asterisks (**) indicate significance at 5% level. NS denotes no statistical significance at 10% level. $\emptyset$ denotes no available results. CCP stands for co-integration cumulative profit test.

Examination of **Table 7** to **Table 10** shows that all methodologies agree that there is a pronounced weakening of the predictive power of the trading rule in the last sub-period (1975-1994). This justifies our decision not to proceed with newer data. Moreover, the results for the significance testing of the ω_0 parameter (*i.e.*, the “buy-sell” difference) of this study agree completely with those of Mills [22] for all time periods and combinations of MA lengths. In addition, the results for the significance testing of the c parameter (mean return for “sell” days) of this study are quite similar, yet not identical, with those of Mills [22], as in this study the null hypothesis of no statistical significance of c is rejected at a lower confidence level. Direct comparison regarding these parameters with the results of Markellos [37] is not possible as the approach followed by the latter is based on a different reasoning. Qualitatively however, it is evident that all methodologies, by and large, agree in their conclusions regarding the predictive performance of the MA trading rule and that the results of this study are somewhat closer to those of Mills [22] as compared to those of Markellos [37].

4. Summary and Conclusions

In this work, a statistical technique based on modified Box-Tiao impact assessment stochastic models is proposed and applied for the assessment of the predictive power of trading rules of technical analysis. Although in this work, the proposed statistical technique was used in stock market data in conjunction specifically with the moving average trading rule, it may be used more generally for the testing of the predictive power of any technical trading rule which generates precise trading signals and for any speculative market. As the predictive power of the MA trading rule almost diminishes from the second half of the 90’s and onwards, it would not be so interesting to just confirm this fact. Therefore, a dataset prior to that period was chosen. Furthermore, for the chosen period, there exist two independent studies with different testing procedures with which the results of our approach were compared.

It is important to clarify that the proposed technique is not a trading rule, but a statistical testing procedure which can be utilized to test the predictive power of any trading rule with some considerable advantages over the two prevailing procedures, namely the simple t -test and the bootstrap. As it was documented, the proposed methodology measures of predictive power of a trading rule, such as the mean value of the “sell” periods and the mean “buy-sell” difference, can be concurrently estimated and tested for their statistical significance, but at the same time, controlling for linear dependencies in asset returns. Application of this approach to the daily closing prices of the FT-30 Index of the London Stock Exchange gave results very similar to those of the established methodology (bootstrapping) and similar to other alternative methods (co-integration cumulative profit test). In particular, the results of the proposed methodology provide evidence of a pronounced weakening of the predictive power of the MA trading rule for the last sub-period (1975-1994), which implies non-rejection of the weak-form

market efficiency hypothesis, confirming the findings of other methods. In general, however, with the statistical methodology proposed in this work, abnormal returns attributed to the predictive power of the trading rule are found to be clearly reduced as compared to those estimated by ordinary t -tests and this reduction seems to be a little higher than the one found by bootstrapping.

To summarize: in comparison to the t -test, the proposed technique is more accurate, as not only does it use more degrees of freedom, but also it describes the possible existence of predictive power in a generally dynamic way as compared to the exclusively static way. The second advantage also applies to the comparison of the proposed modified Box-Tiao impact assessment models as compared to bootstrapping. It is also noted that the proposed technique could potentially be combined with the more recent artificial intelligence approaches to market prediction.

The proposed methodology may be enhanced further, as it is susceptible to technical improvements. As a first step, the problem with non-normal distributions in asset returns may be alleviated by considering the more general class of GED distributions for the testing procedure. Further, second order dependencies in asset returns, such as autoregressive conditional heteroscedasticity, may also be taken into account by considering a general ARMA-GARCH model for the noise component of the impact assessment model.

Hence, overall, the proposed methodology seems promising and potentially has some advantages as compared to the established ones. Therefore, it would be of much interest to apply the methodology proposed in this work employing different trading rules, financial markets, and time spans.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Beaver, W. (1981) Market Efficiency. *The Accounting Review*, **56**, 23-37.
- [2] Malkiel, B. (1992) Efficient Market Hypothesis. In: Newman, P., Milgate, M. and Eatwell, J., Eds., *New Palgrave Dictionary of Money and Finance*, MacMillan, 139.
- [3] Milionis, A.E. (2007) Efficient Capital Markets: A Statistical Definition and Comments. *Statistics & Probability Letters*, **77**, 607-613.
<https://doi.org/10.1016/j.spl.2006.09.007>
- [4] Rubinstein, M. (1975) Securities Market Efficiency in an Arrow-Debreu Economy. *American Economic Review*, **65**, 812-824.
- [5] Roberts, H.V. (1959) Stock-Market "Patterns" and Financial Analysis: Methodological Suggestions. *The Journal of Finance*, **14**, 1-10.
<https://doi.org/10.1111/j.1540-6261.1959.tb00481.x>
- [6] Fama, E.F. (1970) Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, **25**, 383-417. <https://doi.org/10.2307/2325486>
- [7] Fama, E.F. (1991) Efficient Capital Markets: II. *The Journal of Finance*, **46**, 1575-1617. <https://doi.org/10.1111/j.1540-6261.1991.tb04636.x>
- [8] Murphy, J.J. (1999) Technical Analysis of the Financial Markets: A Comprehensive

- Guide to Trading Methods and Applications. New York Institute of Finance, Prentice Hall.
- [9] Taylor, M.P. and Allen, H. (1992) The Use of Technical Analysis in the Foreign Exchange Market. *Journal of International Money and Finance*, **11**, 304-314.
[https://doi.org/10.1016/0261-5606\(92\)90048-3](https://doi.org/10.1016/0261-5606(92)90048-3)
 - [10] Billingsley, R.S. and Chance, D.M. (1996) Benefits and Limitations of Diversification among Commodity Trading Advisors. *Journal of Portfolio Management*, **27**, 5-17.
 - [11] Cheung, Y. and Chinn, M.D. (2001) Currency Traders and Exchange Rate Dynamics: A Survey of the US Market. *Journal of International Money and Finance*, **20**, 439-471. [https://doi.org/10.1016/s0261-5606\(01\)00002-x](https://doi.org/10.1016/s0261-5606(01)00002-x)
 - [12] Park, C. and Irwin, S.H. (2007) What Do We Know about the Profitability of Technical Analysis? *Journal of Economic Surveys*, **21**, 786-826.
<https://doi.org/10.1111/j.1467-6419.2007.00519.x>
 - [13] Milionis, A.E. and Papanagiotou, E. (2008) On the Use of the Moving Average Trading Rule to Test for Weak Form Efficiency in Capital Markets. *Economic Notes*, **37**, 181-201. <https://doi.org/10.1111/j.1468-0300.2008.00198.x>
 - [14] Brock, W., Lakonishok, J. and Lebaron, B. (1992) Simple Technical Trading Rules and the Stochastic Properties of Stock Returns. *The Journal of Finance*, **47**, 1731-1764. <https://doi.org/10.1111/j.1540-6261.1992.tb04681.x>
 - [15] Bessembinder, H. and Chan, K. (1995) The Profitability of Technical Trading Rules in the Asian Stock Markets. *Pacific-Basin Finance Journal*, **3**, 257-284.
[https://doi.org/10.1016/0927-538x\(95\)00002-3](https://doi.org/10.1016/0927-538x(95)00002-3)
 - [16] Cai, B.M., Cai, C.X. and Keasey, K. (2005) Market Efficiency and Returns to Simple Technical Trading Rules: Further Evidence from U.S., U.K., Asian and Chinese Stock Markets. *Asia-Pacific Financial Markets*, **12**, 45-60.
<https://doi.org/10.1007/s10690-006-9012-y>
 - [17] Gençay, R. (1996) Non-Linear Prediction of Security Returns with Moving Average Rules. *Journal of Forecasting*, **15**, 165-174.
[https://doi.org/10.1002/\(sici\)1099-131x\(199604\)15:3<165::aid-for617>3.0.co;2-v](https://doi.org/10.1002/(sici)1099-131x(199604)15:3<165::aid-for617>3.0.co;2-v)
 - [18] Hudson, R., Dempsey, M. and Keasey, K. (1996) A Note on the Weak Form Efficiency of Capital Markets: The Application of Simple Technical Trading Rules to UK Stock Prices 1935 to 1994. *Journal of Banking & Finance*, **20**, 1121-1132.
[https://doi.org/10.1016/0378-4266\(95\)00043-7](https://doi.org/10.1016/0378-4266(95)00043-7)
 - [19] Wong, W., Manzur, M. and Chew, B. (2003) How Rewarding Is Technical Analysis? Evidence from Singapore Stock Market. *Applied Financial Economics*, **13**, 543-551.
<https://doi.org/10.1080/0960310022000020906>
 - [20] Neftci, S.N. (1991) Naive Trading Rules in Financial Markets and Wiener-Kolmogorov Prediction Theory: A Study of "Technical Analysis". *The Journal of Business*, **64**, 549-571. <https://doi.org/10.1086/296551>
 - [21] Pring, M.J. (2002) Technical Analysis Explained. 4th Edition, McGraw-Hill.
 - [22] 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.
[https://doi.org/10.1002/\(sici\)1099-1158\(199710\)2:4<319::aid-jfe53>3.0.co;2-6](https://doi.org/10.1002/(sici)1099-1158(199710)2:4<319::aid-jfe53>3.0.co;2-6)
 - [23] Ohnishi, T., Mizuno, T., Aihara, K., Takayasu, M. and Takayasu, H. (2004) Statistical Properties of the Moving Average Price in Dollar-Yen Exchange Rates. *Physica A: Statistical Mechanics and Its Applications*, **344**, 207-210.
<https://doi.org/10.1016/j.physa.2004.06.118>

- [24] Olson, D. (2004) Have Trading Rule Profits in the Currency Markets Declined over Time? *Journal of Banking & Finance*, **28**, 85-105. [https://doi.org/10.1016/s0378-4266\(02\)00399-0](https://doi.org/10.1016/s0378-4266(02)00399-0)
- [25] Papailias, F. and Thomakos, D.D. (2015) An Improved Moving Average Technical Trading Rule. *Physica A: Statistical Mechanics and Its Applications*, **428**, 458-469. <https://doi.org/10.1016/j.physa.2015.01.088>
- [26] Luukka, P., Pätäri, E., Fedorova, E. and Garanina, T. (2015) Performance of Moving Average Trading Rules in a Volatile Stock Market: The Russian Evidence. *Emerging Markets Finance and Trade*, **52**, 2434-2450. <https://doi.org/10.1080/1540496x.2015.1087785>
- [27] Corbet, S., Eraslan, V., Lucey, B. and Sensoy, A. (2019) The Effectiveness of Technical Trading Rules in Cryptocurrency Markets. *Finance Research Letters*, **31**, 32-37. <https://doi.org/10.1016/j.frl.2019.04.027>
- [28] Miralles-Quirós, J.L., Miralles-Quirós, M.d.M. and Valente Gonçalves, L.M. (2018) The Profitability of Moving Average Rules: Smaller Is Better in the Brazilian Stock Market. *Emerging Markets Finance and Trade*, **55**, 150-167. <https://doi.org/10.1080/1540496x.2017.1422428>
- [29] Milionis, A.E. and Papanagiotou, E. (2011) A Test of Significance of the Predictive Power of the Moving Average Trading Rule of Technical Analysis Based on Sensitivity Analysis: Application to the NYSE, the Athens Stock Exchange and the Vienna Stock Exchange. Implications for Weak-Form Market Efficiency Testing. *Applied Financial Economics*, **21**, 421-436. <https://doi.org/10.1080/09603107.2010.532105>
- [30] Milionis, A.E. and Papanagiotou, E. (2013) Decomposing the Predictive Performance of the Moving Average Trading Rule of Technical Analysis: The Contribution of Linear and Non-Linear Dependencies in Stock Returns. *Journal of Applied Statistics*, **40**, 2480-2494. <https://doi.org/10.1080/02664763.2013.818624>
- [31] Box, G.E.P. and Tiao, G.C. (1975) Intervention Analysis with Applications to Economic and Environmental Problems. *Journal of the American Statistical Association*, **70**, 70-79. <https://doi.org/10.1080/01621459.1975.10480264>
- [32] Abraham, B. (1980) Intervention Analysis and Multiple Time Series. *Biometrika*, **67**, 73-80. <https://doi.org/10.2307/2335318>
- [33] Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994) Time Series Analysis Forecasting and Control. 3rd Edition, Prentice-Hall.
- [34] McCleary, R. and Hay, R.A. (1980) Applied Time Series Analysis for the Social Sciences. Sage.
- [35] Box, G.E.P. and Jenkins, G. (1976) Time Series Analysis: Forecasting and Control. Rev. Edition, Holden Day.
- [36] Liu, L.M. (1988) Box-Jenkins Time Series Analysis. In: *BMDP Statistical Software Manual*, University of California Press, 429.
- [37] Markellos, R.N. (1999) Investment Strategy Evaluation with Cointegration. *Applied Economics Letters*, **6**, 177-179. <https://doi.org/10.1080/135048599353582>
- [38] Ljung, G.M. and Box, G.E.P. (1978) On a Measure of Lack of Fit in Time Series Models. *Biometrika*, **65**, 297-303. <https://doi.org/10.1093/biomet/65.2.297>
- [39] Milionis, A.E., Moschos, D. and Xanthakis, M. (1998) The Influence of Foreign Markets on the Athens Stock Exchange. *Spoudai*, **48**, 140-156.
- [40] Milionis, A.E. and Moschos, D. (2000) On the Validity of the Weak-Form Efficient Markets Hypothesis Applied to the London Stock Exchange: Comment. *Applied Economics Letters*, **7**, 419-421. <https://doi.org/10.1080/135048500351087>

- [41] Milionis, A.E. (2004) The Importance of Variance Stationarity in Economic Time Series Modelling. A Practical Approach. *Applied Financial Economics*, **14**, 265-278. <https://doi.org/10.1080/0960310042000220040>
- [42] Patil, A. and Rastogi, S. (2019) Time-Varying Price-Volume Relationship and Adaptive Market Efficiency: A Survey of the Empirical Literature. *Journal of Risk and Financial Management*, **12**, Article No. 105. <https://doi.org/10.3390/jrfm12020105>
- [43] Milionis, A.E. (2019) A Simple Return Generating Model in Discrete Time: Implications and Potential Applications for Market Efficiency Testing and Technical Analysis. *Theoretical Economics Letters*, **9**, 973-985. <https://doi.org/10.4236/tel.2019.94063>
- [44] Milionis, A.E. (2023) On the Assimilation of New Information on Asset Prices. In: *An Overview on Business, Management and Economics Research*, Vol. 5, B P International, 11-25. <https://doi.org/10.9734/bpi/aobmer/v5/6218b>
- [45] Lim, K. and Brooks, R. (2011) The Evolution of Stock Market Efficiency over Time: A Survey of the Empirical Literature. *Journal of Economic Surveys*, **25**, 69-108. <https://doi.org/10.1111/j.1467-6419.2009.00611.x>
- [46] White, H. (2000) A Reality Check for Data Snooping. *Econometrica*, **68**, 1097-1126. <https://doi.org/10.1111/1468-0262.00152>
- [47] Hansen, P.R. (2005) A Test for Superior Predictive Ability. *Journal of Business & Economic Statistics*, **23**, 365-380. <https://doi.org/10.1198/073500105000000063>
- [48] Hansen, P.R., Lunde, A. and Nason, J.M. (2011) The Model Confidence Set. *Econometrica*, **79**, 453-497.
- [49] Milionis, A.E. and Papanagiotou, E. (2009) A Study of the Predictive Performance of the Moving Average Trading Rule as Applied to NYSE, the Athens Stock Exchange and the Vienna Stock Exchange: Sensitivity Analysis and Implications for Weak-Form Market Efficiency Testing. *Applied Financial Economics*, **19**, 1171-1186. <https://doi.org/10.1080/09603100802375519>