Testing adaptive market efficiency in the presence of non-Gaussian uncertainties

Vykta Wakandigara

A dissertation submitted to the Faculty of Commerce, University of Cape Town, in partial fulfilment of the requirements for the degree of Master of Philosophy.

October 21, 2019

MPhil in Mathematical Finance,
University of Cape Town.

Abstract

One of the central debates in finance concerns the Efficient Market Hypothesis (EMH)—wherein markets are assumed to be efficient in the absolute sense. However, the possibility of time-varying weak-form market efficiency has received increasing attention in recent years. Under the Adaptive Market Hypothesis (AMH) it is postulated that market efficiency is dynamic, which advocates using models with non-constant coefficients. The concept of evolving efficiency has yielded a Test for Evolving Efficiency (TEE) and following that, a Generalised Test for Evolving Efficiency (GTEE) – both with an associated Kalman filtering (KF) technique. Unfortunately, these methods assume that the inherent stochastic processes are Gaussian despite widespread evidence that many real financial time series are non-Gaussian.

Unlike the classical KF, modern filters such as the maximum correntropy Kalman filters (MCC-KF) have been shown to be less sensitive to non-Gaussian uncertainties. These filters utilise a similarity measure known as correntropy–which incorporates higher order information than the mean square criterion that is utilised in the classical KF. As a result, they have been shown to improve filter robustness against outliers or impulsive noises.

In this paper, the South African and American stock markets are tested for adaptive market efficiency using both the standard KF and the MCC-KF. A simulation study shows that the MCC-KF is a more robust estimator of adaptive efficiency but it less accurately estimates unknown system parameters. The South African stock market is found to be inefficient prior to August 2004 but achieves efficiency thereafter. Testing the S&P500 does not provide evidence of inefficiency in the American stock markets. The GTEE, implemented with the MCC-KF, is selected as the best-performing test for the S&P500.

Acknowledgements

I would like to thank Dr. Maria Kulikova, Professor David Taylor and Mr. Obeid Mahomed for their supervision of my Masters' dissertation. Without your buttressing advice and tremendous patience, this paper would have failed to reach completion. I would be remiss if I failed to acknowledge the efforts of Miss Kungela Mzuku who encouraged me to apply myself fully. Finally, this paper is dedicated to my parents, Albert and Berlina Wakandigara, who have fervently financed my education since my childhood.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In an informationally efficient capital market, the Efficient Market Hypothesis (EMH) anticipates that security prices will "fully reflect" all available information (Fama, 1965). Any new knowledge that is relevant to the valuation of a security will be instantaneously incorporated into said security's price as soon as it is available to the market. More formally, if a market is efficient with respect to some information set $\mathcal{F}$, then revealing the information contained in $\mathcal{F}$ to all market participants would have no bearing on the current state of security prices (Malkiel, 1989). Thus, it should be impossible for an investor to make reliable economic profits by trading on the basis of information derived from $\mathcal{F}$, since this information does not determine future prices. Future prices are determined by some other information set $\mathcal{G}$, which according to the EMH, is not available to any participants at the present. Roberts (1967) defines three classes of information sets:

1. Historic prices: A market is *weak-form* efficient if securities prices fully reflect all information contained in historic price sequences.

2. Publicly available information: A market is *semi-strong form* efficient if all prices fully reflect publicly available information in addition to past prices.

3. All available information: A market is *strong-form* efficient if prices fully reflect all information that may be known by any market participants. Consequently, a trader cannot devise a profitable trading strategy that is built on the basis of privately held information.

This taxonomy of informational sets has directed volumes of literature toward the study of weak-form market efficiency and away from the stronger forms (Lim and Brooks, 2011). Although the weak form EMH subjects itself to more accessible testing than its counterparts, the issue of whether certain markets are efficient even in the weak sense has not yet achieved an ultimate conclusion (Ţiţan, 2015).

Currently, the lack of agreement about whether markets are efficient is being sustained by objections from the field of behavioural economics. Behavioural economists argue against the EMH by refuting some of its principal assumptions. For example, an extensive body of research has been published which discredits the notion that humans are perfectly rational; an essential assumption upon which the EMH is premised (Pesaran, 1987). Proponents of the EMH respond by arguing that while occasional human irrationality may exist, its influence is limited by the force of arbitrage. As irrational market players create profit opportunities through mis-pricing, EMH supporters maintain that the (possibly few) market participants that will be behaving rationally will immediately exploit the opportunity to benefit from the mis-priced securities. This rebuttal, however, relies on the ability of arbitrageurs to remove inefficiencies both quickly and completely; faster than any such opportunities may arise. As Lo (2004) points out, the answer to that question favours a more empirical approach rather than a theoretical one. A more detailed discussion on the matter may be found in Malkiel *et al.* (2005), which better illustrates the uncompromising level of disagreement between the advocates of behavioural finance and the academics who support the EMH.

Given the significant evidence of human irrationality, it is not unreasonable to believe that markets may occasionally depart from absolute efficiency; particularly when irrational pricing by certain market participants cannot be exploited by arbitrage capital. This belief, however, need not be intellectually inconsistent with the EMH. Lo (2004) proposes the Adaptive Market Hypothesis (AMH); a theory which aims to reconcile the EMH and the behavioural economists who dispute it. In the AMH, Lo (2004) advocates for a new framework by applying evolutionary principles to financial market interactions. Market participants are viewed as organisms, permitted to succumb to any of the biases that are often observed by behavioural economists. These 'market organisms' are assumed to be bounded in rationality and through trial-and-error, they develop heuristics that assist them in making near-optimal decisions. The amount of profit that they garner offers them feedback on whether their heuristics are currently successful. If these profits are abundant, then the participants need not continue their search for the optimal heuristic; a principle consistent with 'satisficing' (Simon, 1955).

Where participants do not act optimally, inefficiencies may exist, whereas in a highly competitive system, the market is likely to be efficient and at equilibrium. The EMH characterises the equilibrium state within this ecosystem; where the environmental circumstances allow arbitrageurs to near-instantly eliminate any inefficiencies that may exist in the market. However, as the environment changes, perhaps due to regulation for example, some of the heuristics developed in a previous regime may now be inefficient (Lo, 2004). Thus, the AMH predicts that inefficiencies may persist in the market from time to time.

Even prior to the formalisation of the AMH, several studies had been produced in an effort to test for time-varying efficiency. Many of the early tests of the EMH treated efficiency as a binary characteristic; a market was either found to be efficient or not within a fixed interval of data. Consequently, if it is acceptable that markets become more efficient over time, these tests imply that the change from inefficiency to efficiency is discrete (Lim and Brooks, 2011). Emerson *et al.* (1997) were the first to utilise a time-varying parameter model and an associated Kalman filter, to trace the progression of returns predictability. Kulikova and Taylor (2014) do note, however, that the model specification presented in Emerson *et al.* (1997) was incorrect and was corrected when Zalewska-Mitura and Hall (1999) formalised the Test of Evolving Efficiency (TEE). In the TEE, returns are described by a GARCH-M process. Many of the adaptive tests of efficiency that have followed Zalewska-Mitura and Hall (1999) 's study have aimed to improve the TEE by incorporating more realistic models of financial returns. Kulikova *et al.* (2019) generalise the test by developing the Generalised Test of Evolving Efficiency (GTEE). Rather than assuming that conditional volatility is deterministic, the GTEE substitutes the GARCH-M for a stochastic GARCH-M process. In addition, the GTEE permits for an entirely different volatility feedback function than the TEE. In their investigation, Kulikova *et al.* (2019) find that non-linear feedback functions are preferred for the RTSI and JSE Top 40 series.

Even though many tests of evolving efficiency involve the GARCH in one way or another, modelling returns with a stochastic volatility model is also an alternative. By assimilating the stochastic volatility in mean (SVM) model developed by (Koopman and Hol Uspensky, 2002) into the TEE framework, Holder (2017) constructs the stochastic volatility test of evolving efficiency (SV-TEE). After applying the SV-TEE to various markets, Holder concludes that the TEE and GTEE perform better than the SV-TEE.

Each of the three models (TEE, GTEE, SV-TEE) were implemented assuming that the innovations in the model were conditionally normally distributed. These will be referred to as the normal models i.e. the normal GARCH and the normal stochastic volatility. Since the seminal work by Fama (1965) and Mandelbrot (1972), it has become a stylised fact that the unconditional distribution of returns is leptokurtic. There is overwhelming evidence that the kurtosis implied by the normal GARCH model is often much less than the kurtosis found in financial data (Ghahramani and Thavaneswaran, 2008). Although some of this leptokurtic behaviour can be explained by the conditional heteroskedasticity in the normal GARCH, the rest may need to be directly modelled by a GARCH model that is conditionally leptokurtic. Bollerslev (1987) models a GARCH that is conditionally t-distributed, Nelson (1991) uses a generalised error distribution and Bai *et al.* (2003) uses a Gaussian mixture distribution. These models, though more difficult to estimate, can generate a higher unconditional kurtosis.

As Holder (2017) noted, estimating the TEE with a standard Kalman filter immediately imposes a restriction on the error distributions that the innovations are assumed to follow. Developments in the field of filtering, however, have yielded modern filters that are robust to non-Gaussian noise (Izanloo *et al.* (2016), Kulikova (2017), Chen *et al.* (2017)). The maximum correntropy criterion Kalman filter (MCC-KF) developed by Izanloo *et al.* (2016) maximises correntropy, a cost function that uses higher order signals, to compare similarity in random variables. In systems that are Gaussian, the MCC-KF performs similarly to the Kalman filter. More importantly however, it has been shown to produce better estimation quality than the Kalman filter in environments that are corrupted by non-Gaussian impulses (Izanloo *et al.*, 2016).

This paper tests for weak-form market efficiency in the presence of non-Gaussian returns. To the authors' best knowledge, no other literature has utilised the MCC-KF in state-space estimation of financial data. The order of the paper proceeds as follows. Section 2 presents the theoretical background for the tests of efficiency. Only the TEE and the GTEE are considered in this paper, owing to their superior performance in the literature (Holder, 2017). In Section 3, computable formulae for implementing the KF and MCC-KF are derived, along with their associated quasi-maximum likelihood estimate (QMLE) procedures. Combining these methods with the theoretical background presented in Chapter 2, the nuances of the tests of efficiency are investigated in a simulation study. An important conclusion from the study is that the the MCC-KF achieves lower mean square errors than the KF when measuring measuring predictability where returns are extremely leptokurtic. Nevertheless, the KF realises a lower AIC than the MCC-KF when it is paired with the QMLE technique. This is one of the main contributions of the paper. Regardless, it is shown that either filter estimates time-varying autocorrelation within the estimated confidence bounds. After showing that the tests perform reasonably well in the presence of non-Gaussian returns, the TEE and various forms of the GTEE are applied to the S&P500 and the FTSE/JSE All Share Index (ALSI) in Section 4. The ALSI is selected because it represents 99% of the JSE's total market capital, with each stock in the index assessed for liquidity Ward and Muller (2012). FTSE Russell (2019) constructs the index so that it is both investable and tradable. Where stocks are deemed to be either too illiquid or small, they are excluded from the ALSI. The results show that the South African stock market may have been inefficient prior to 2005 but has been efficient since then. No evidence of inefficiency can be found for the American stock markets. Section 5 concludes the paper.

Chapter 2

Methodology: Testing Evolving Efficiency

2.1 The time-varying model approaches for examining the AMH

Two approaches can be used to examine the AMH. The first approach involves a model with time-varying parameters which should track changing levels of efficiency within a fixed time window (Ito *et al.* (2014), Ito *et al.* (2016)). The second approach applies classical tests of static efficiency in rolling estimation windows (Kim *et al.* (2011), Lim *et al.* (2013)). The rationale behind the second approach is that varying test results among the windows suggest the presence of evolving efficiency, which therefore supports the AMH. A notable weakness in this methodology, however, is the arbitrary selection of an optimal window width. This paper intends to circumvent the aforementioned weakness, and as a result, it applies two tests with time-varying parameters; specifically the Test of Evolving Efficiency (TEE) and the Generalised Test of Evolving Efficiency (GTEE). Although both tests measure the forecastability of returns by estimating auto-regressive parameters, the differences among them are due to each test's unique specification of the process generating returns. The TEE is the forerunning model; initially introduced by Zalewska-Mitura and Hall (1999) and Emerson *et al.* (1997). It estimates the evolution of efficiency by following the path of time-dependant parameters within the GARCH-M framework. The GTEE, developed by Kulikova *et al.* (2019) generalises the TEE by assuming that the returns are influenced by a stochastic GARCH-M process rather than a standard GARCH-M process. The GTEE also allows for a wide variety of volatility feedback functions. Each of the original models will be discussed in this chapter. Thereafter, the importance of allowing for non-Gaussian errors in these models will be considered.

2.2 The Test for Evolving Efficiency

The weak form of the EMH implies that no profitable trading opportunities may be constructed from historical prices. If a trader finds that the historical record of returns predicts future returns, they can form a profitable trading strategy from this information; barring limitations in execution such as transaction costs. Thus, the aim of TEE is to estimate efficiency by finding predictable patterns in returns; assuming such patterns are exploitable.

A preliminary step in testing for efficiency is to establish a plausible model for returns. It has long since been recognised that volatility in financial time series tends to cluster together (Fama, 1965), (Mandelbrot, 1972), implying that uncertainty changes over time and that it is autocorrelated. The advent of the ARCH model (Engle, 1982) led to new research that was focused on applying and extending the conventional ARCH formulation. An especially important extension is the generalised ARCH (GARCH) framework introduced by Bollerslev (1986) ,which permits current estimates of conditional variance to depend on previous conditional variances. A further generalisation to the GARCH is the GARCH-in-mean (Bollerslev *et al.*, 1988) which extended the original ARCH-M model by Engle *et al.* (1987). More specifically, if ϵ_t denotes a real-time, discrete, stochastic process with variance h_t conditioned on information set $\mathcal{F}_{t-1}$ and x_t is a vector of weakly-exogenous variables, then the GARCH(p, q)-M process for returns has the following dynamics:

$$r_t = \gamma^\top x_t + \delta h_t + \epsilon_t, \quad \epsilon_t | \mathcal{F}_{t-1} \sim N(0, h_t) \tag{2.1}$$

$$h_t = \omega + \sum_{k=1}^{q} \alpha_k {\epsilon^2}_{t-k} + \sum_{k=1}^{p} \beta_k h_{t-k} \tag{2.2}$$

The benefit that the GARCH-in-mean (GARCH-M) model confers is that it allows the conditional variance of returns to influence their conditional expection (Engle *et al.*, 1987). This modification may not only increase the predictive power of a standard GARCH but may also introduce an ancillary interpretation. A positive relationship between the conditional variance and conditional expected returns would provide evidence for the risk aversion of investors while a negative relationship would suggest the presence of the leverage effect as discussed by Black (1976). The TEE, which is based on the GARCH-M(1,1) model, takes on the following form:

$$r_t = \beta_{0,t} + \beta_{1,t} r_{t-1} + \delta h_t + \epsilon_t, \quad \epsilon_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, h_t), \quad t \in T = 1, 2, ...N, \tag{2.3}$$

$$h_t = a_0 + a_1 \epsilon_{t-1}^2 + b_1 h_{t-1}, \tag{2.4}$$

$$\beta_{i,t} = \beta_{i,t-1} + w_{i,t}, \quad i = 0, 1, \quad w_{i,t} \sim \mathcal{N}(0, \sigma_i^2) \tag{2.5}$$

where $(r_t)_{t \in T}$ is the returns process and $(h_t)_{t \in T}$ is the conditional volatility of returns given the information $\mathcal{F}_{t-1}$ that is adapted to the history of prices up to time $t-1$. As evidenced by equation (2.5), the $(\beta_{i,t})_{t \in T}$ with $i \in [0, 1]$ are taken to be stochastic processes; assuming different values in time. To guarantee that h_t is positive, it is assumed that $a_0 > 0$ and $a_1, b_1 \geq 0$. In (Zalewska-Mitura and Hall (1999), Emerson *et al.* (1997)), the TEE is estimated using a standard Kalman filter. Estimation of the parameter vector $\theta^1 = (\delta, a_0, a_1, b_1, \sigma_0, \sigma_1)$ is performed via the Quasi-Maximum Likelihood Estimation (QMLE) procedure. Weak form efficiency implies that $\beta_{1,t} = 0$ ($\forall t \in T$). By explicitly modelling the returns' heteroskedastic behaviour and dependance on a changing variance, the TEE avoids detecting spurious correlation (Zalewska-Mitura and Hall, 1999).

2.3 The Generalised Test for Evolving Efficiency

The GARCH-M equation assumes that the relationship in equation (2.2) is exact. Hall (1990) is unconvinced by this specification, asserting that equation (2.2) should at least include some measurement error. Hall (1990) then alters the conditional volatility equation (2.2) by adding a stochastic term to it:

$$r_t = \gamma^\top x_t + \delta h_t + \epsilon_t, \quad \epsilon_t | \mathcal{F}_{t-1} \sim N(0, h_t) \tag{2.6}$$

$$h_t = \omega + \sum_{k=1}^{q} \alpha_k {\epsilon^2}_{t-k} + \sum_{k=1}^{p} \beta_k h_{t-k} + \omega_t \quad \omega_t \sim N(0, Q_t) \tag{2.7}$$

Equations (2.6) and (2.7) completely specify the stochastic GARCH-in-Mean (SGARCH-M) model.

The GARCH-M is a special case of the SGARCH-M, which occurs when $Q_t = 0$. The matrix Q_t is considered to be deterministic. In addition, if both $Q_t = 0$ and $\delta = 0$ the SGARCH-M becomes the GARCH. While the SGARCH-M is more general than the GARCH-M; the increased plausibility of the model comes at the expense of a more complex estimation procedure. For example, since ω_t is normally distributed, Q_t must be sufficiently small to impede h_t from taking on negative values.

Hall (1990) advocates for the Kalman filter as an estimator of the SGARCH-M so that estimates at time t are updated by the information in $\mathcal{F}_{t-1}$. This is unlike the standard estimation procedure.

In Kulikova *et al.* (2019), the authors incorporate the SGARCH-M into the design of a new Generalised Test of Evolving Efficiency (GTEE). Kulikova *et al.* (2019) hypothesize further that alternative volatility feedback functions than those specified in the TEE might better explain the influence that volatility exerts on returns. The GTEE is given by:

$$r_t = \beta_{0,t} + \beta_{1,t} r_{t-1} + \delta f(h_t) + \epsilon_t, \quad \epsilon_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, h_t), \quad t = 1, 2, ...N, \tag{2.8}$$

$$h_t = a_0 + a_1 \epsilon_{t-1}^2 + b_1 h_{t-1} + \omega_t, \quad \omega_t \sim N(0, Q_t) \tag{2.9}$$

$$\beta_{i,t} = \beta_{i,t-1} + w_{i,t}, \quad i = 0, 1, \quad w_{i,t} \sim \mathcal{N}(0, \sigma_i^2) \tag{2.10}$$

with each variable defined as it were in the TEE case. The GTEE reduces to the TEE when $Q_t = 0$ and the volatility feedback is the identity function i.e. $f(h_t) = h_t$. As in the TEE, $a_0 > 0$ and $a_1, b_1 \geq 0$ to ensure the positivity of h_t. The parameter vector associated with the GTEE is given by $\theta^2 = (\delta, a_0, a_1, b_1, \sigma_0, \sigma_1, Q_t)$. The empirical study in (Holder, 2017) concludes that the GTEE with non-linear feedback functions performs better than the TEE and the GTEE with linear volatility feedbacks in periods of high volatility. This paper will implement the same non-linear feedback functions as in the previously mentioned study. These two functions are given by $f(h_t) = \sqrt{h_t}$ and $f(h_t) = \log(h_t)$. These functions are chosen because they have performed well in particular markets as alternatives to the linear feedback function (Kulikova *et al.* (2019), Engle *et al.* (1987)).

2.4 Importance of non-Gaussian extensions

Gaussian GARCH models have been successful in capturing certain aspects of financial time series such as time-dependant volatility and volatility clustering. However, normal GARCH models do not accommodate the heavy tails, extreme events and excess kurtosis that are often found in real financial data (Urquhart and McGroarty, 2014).

In many papers, models with conditionally leptokurtic error distributions have provided a better fit to financial data than their Gaussian counterparts. Bollerslev (1987) finds that more of the observed kurtosis in the monthly S&P500 Composite Index could be accounted for by GARCH models with conditionally t-distributed errors. Jorion (1988) and Nelson (1991) also find that more leptokurtic conditional error distributions are better suited to their data, with each paper advocating for a different density. These findings suggest that equations (2.3) and (2.8) may not best represent the evolution of actual returns. The results from the TEE and GTEE may not be optimal if a Kalman filter is used to estimate $(\beta_{1,t})_{t\in T}$ where the returns are generated by a conditionally leptokurtic distribution. In addition, the assumption that the $(\beta_{i,t})_{t\in T, i\in[0,1]}$ are Gaussian processes is plausible since this process is hidden and thus their true distribution cannot be strictly assumed to be normal.

This paper aims to reduce the impact of non-Gaussian errors on the estimation quality of each of the TEE and GTEE. Some papers have found that Kalman filters degenerate when estimating non-Gaussian dynamic systems; a particular example being tracking a moving object in the presence of outlying observations (Bilik and Tabrikian, 2010). Where simulation studies have been conducted, the MCC-KF has proven to be more robust than the Kalman filter when estimating processes corrupted by non-Gaussian noise (Izanloo *et al.* (2016), Kulikova (2017), Chen *et al.* (2017)). This new filter has found its applications in physical and engineering-based systems. In this paper, however, the TEE and GTEE will be recast in state-space form thus making them suitable for estimation by adaptive filtering techniques. The derivation of the Kalman filter and MCC-KF is discussed in the next chapter whereafter their implementation with respect to the TEE and GTEE is presented. The performance of the filters in each test is compared in a simulation study.

Chapter 3

State-space approach for the TEE model structures

3.1 The Kalman Filter approach

3.1.1 The linear Kalman filter

Consider the classical state-space model representation for linear Gaussian Hidden Markov Models:

$$x_k = F_{k-1}x_{k-1} + B_{k-1}u_{k-1} + G_{k-1}w_{k-1}, \qquad w_k \sim \mathcal{N}(0, Q), \tag{3.1}$$
$$y_k = H_k x_k + v_k \qquad v_k \sim \mathcal{N}(0, R) \tag{3.2}$$

where the subscript k, $k \geq 0$, refers to the discrete time, i.e. y_k means $y(t_k)$ and so on. Assume for the rest of the chapter that the vectors $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^d$ and $y_k \in \mathbb{R}^m$ are, respectively, the unknown (hidden) dynamic state, the control input and the available measurements. Equations (3.1) and (3.2) are called the *state* and *measurement* equations, respectively. Going forward, the state and the measurement uncertainty processes will be denoted by $\{w_k\}$ and $\{v_k\}$ respectively. Both error processes are independent zero-mean white-noise with covariance matrices $Q \geq 0$ and $R > 0$, respectively. They are also uncorrelated with initial state $x_0 \sim \mathcal{N}(\bar{x}_0, \Pi_0)$, $\Pi_0 \geq 0$.

When the classical KF estimates a hidden dynamic state process $\{x_k\}_{k=1}^N$ from an observed sequence $\{y_k\}_{k=1}^N$, the result is a sequence of minimum mean-square estimates, $\{\hat{x}_{k|k}\}_{k=1}^N$, for *linear Gaussian* state-space models. The quantity $\hat{x}_{k|k}$ represents a state estimate at time t_k, given the available measurements $\{y_1, \ldots, y_k\}$. Below, the classical KF recursion is given (Kailath *et al.*, 2000, Theorem 9.2.1):

Algorithm 1. KF (*Conventional KF implementation*)

INITIALIZATION: $(k=0)$ $\hat{x}_{0|0}=\bar{x}_0$ and $P_{0|0}=\Pi_0$.

TIME UPDATE: $(k=\overline{1,N})$ ▷ PRIORI ESTIMATION

$\hat{x}_{k|k-1}=F_{k-1}\hat{x}_{k-1|k-1}+Bu_{k-1}$,

$P_{k|k-1}=F_{k-1}P_{k-1|k-1}F_{k-1}^T+G_{k-1}QG_{k-1}^T$.

MEASUREMENT UPDATE: $(k=\overline{1,N})$ ▷ POSTERIORI

▷ ESTIMATION

$R_{e,k}=H_kP_{k|k-1}H_k^T+R$,

$K_k=P_{k|k-1}H_k^TR_{e,k}^{-1}$,

$\hat{x}_{k|k}=\hat{x}_{k|k-1}+K_ke_k$ where $e_k=y_k-H_k\hat{x}_{k|k-1}$,

$P_{k|k}=(I-K_kH_k)P_{k|k-1}$.

The step where the one-step ahead predicted (*a priori*) estimate, $\hat{x}_{k|k-1}$, is computed together with the corresponding error covariance matrix

$$P_{k|k-1}=\mathbf{E}\left\{(x_k-\hat{x}_{k|k-1})(x_k-\hat{x}_{k|k-1})^T\right\}$$

is called the time update. As soon as y_k becomes available at time t_k, the KF updates the state estimate through the feedback gain matrix K_k. This updating step is called the measurement update, where both the *a posteriori* estimate $\hat{x}_{k|k}$ and its respective error covariance matrix $P_{k|k}$ are computed. An important consequence of applying the KF to Gaussian state-space models (3.1), (3.2) is that $e_k\sim\mathcal{N}(0,R_{e,k})$ where $\{e_k\}$ are the discrete-time KF innovations.

3.1.2 The Extended Kalman Filter

Now consider the following non-linear dynamic system:

$$x_k=\mathcal{F}(x_{k-1},u_{k-1},w_{k-1}),\qquad w_k\sim\mathcal{N}(0,Q),\tag{3.3}$$

$$y_k=\mathcal{H}(x_k,v_k),\qquad v_k\sim\mathcal{N}(0,R)\tag{3.4}$$

The Extended Kalman filter (EKF) is derived by performing Taylor series expansions of the state and measurement equations (3.3) and (3.4). When linearizing the state equation, the following Jacobian matrices are derived:

$$F_{k-1}=\frac{\partial\mathcal{F}}{\partial x}\bigg|_{\hat{x}_{k-1|k-1}},\qquad G_{k-1}=\frac{\partial\mathcal{F}}{\partial w}\bigg|_{\hat{w}_{k-1|k-1}}.$$

These matrices are used in the EKF's predicted covariance computation. Similarly the Jacobian matrices for the measurement equation are given by:

$$H_k=\frac{\partial\mathcal{H}}{\partial x}\bigg|_{\hat{x}_{k|k-1}},\qquad M_k=\frac{\partial\mathcal{H}}{\partial v}\bigg|_{\hat{v}_{k|k-1}}.$$

The matrices H_k and M_k form part of the EKF's residual covariance calculation. With these partial derivatives at hand, the EKF algorithm is readily implemented:

Algorithm 2. EKF (*Conventional EKF implementation*)

INITIALIZATION: $(k = 0)$ $\hat{x}_{0|0} = \bar{x}_0$ and $P_{0|0} = \Pi_0$.

TIME UPDATE: $(k = \overline{1, N})$ ▷ PRIORI ESTIMATION

$\hat{x}_{k|k-1} = \mathcal{F}(\hat{x}_{k-1|k-1}, u_{k-1}, 0)$

$P_{k|k-1} = F_{k-1} P_{k-1|k-1} F_{k-1}^T + Q$

MEASUREMENT UPDATE: $(k = \overline{1, N})$ ▷ POSTERIORI ▷ ESTIMATION

$R_{e,k} = H_k P_{k|k-1} H_k^T + R,$

$K_k = P_{k|k-1} H_k^T R_{e,k}^{-1},$

$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k e_k$ where $e_k = y_k - \mathcal{H}(\hat{x}_{k|k-1}, 0),$

$P_{k|k} = (I - K_k H_k) P_{k|k-1}.$

3.2 The maximum correntropy criterion Kalman Filter approach

3.2.1 The Standard MCC-KF

Suppose that the examined state-space model is non-Gaussian, i.e. the processes $\{w_k\}$ and $\{v_k\}$ are zero-mean, white, uncorrelated, and have known covariance matrices Q_k and R_k, respectively.

As stated before, applying the KF to a state-space system produces a series of linear minimum mean square estimates, $\{\hat{x}_{k|k}\}_{k=1}^N$. Where the system is disturbed by non-Gaussian impulses, robust estimates may be constructed by maximising the correntropy criterion rather than minimising mean square errors (MSE) Cinar and Príncipe (2012); **?**. Instead of minimising the MSE between the hidden state x_k and its estimate $\hat{x}_k$, the aim is to maximise the following Gaussian kernel:

$$G_\sigma(x_k, \hat{x}_k) = \exp\{-\|x_k - \hat{x}_k\|^2/(2\sigma^2)\}$$

It is possible to use other kernels in the MCC-KF however, the Gaussian kernel is considered in this paper since it is the most commonly used in the literature (Wang *et al.*, 2018). Clearly, $G_\sigma(x_k, \hat{x}_k)$ is maximised when $x_k = \hat{x}_k$. Applying a Taylor expansion to $G_\sigma(x_k, \hat{x}_k)$ shows how the Gaussian kernel captures more than second order information in measuring similarity between random variables. Bearing this in mind, **?** derives the MCC-KF from maximising a weighted cost function that is a linear combination of $G_\sigma(y, Hx_k)$ and $G_\sigma(x_k, Hx_{k-1})$. The improved version in Kulikova (2016) is derived for the case when $B_k = 0$ (equation (3.1)). It is not difficult to extend the proposed MCC-KF technique for the case when $B_k \neq 0$. Following Kulikova (2016), we have

Algorithm 3. MCC-KF (*Improved conventional version*)

1 INITIALIZATION: $(k = 0)$ $\hat{x}_{0|0} = \bar{x}_0$ and $P_{0|0} = \Pi_0$.
TIME UPDATE: $(k = \overline{1, N})$ ▷ PRIORI ESTIMATION
2 $\hat{x}_{k|k-1} = F_{k-1}\hat{x}_{k-1|k-1} + Bu_{k-1}$,
3 $P_{k|k-1} = F_{k-1}P_{k-1|k-1}F_{k-1}^T + G_{k-1}QG_{k-1}^T$.
MEASUREMENT UPDATE: $(k = \overline{1, N})$ ▷ POSTERIORI
▷ ESTIMATION
4 $L_k = \frac{G_\sigma\left(\|y_k - H_k\hat{x}_{k|k-1}\|_{R^{-1}}\right)}{G_\sigma\left(\|\hat{x}_{k|k-1} - F_{k-1}\hat{x}_{k-1|k-1} - Bu_{k-1}\|_{P_{k|k-1}^{-1}}\right)}$,
5 $R_{e,k} = H_k P_{k|k-1} L_k H_k^T + R$,
6 $K_k^L = P_{k|k-1} L_k H_k^T R_{e,k}^{-1}$,
7 $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k^L e_k$, $\quad e_k = y_k - H_k\hat{x}_{k|k-1}$,
8 $P_{k|k} = (I - K_k^L H_k)P_{k|k-1}$.

This paper will apply the MCC-KF filter instead of the classical KF for estimating the TEE model in the presence of non-Gaussian uncertainties. The filter is also used for calculating the likelihood function for QML estimation procedure (it helps to estimate the parameters of the TEE model).

3.2.2 The Extended MCC-KF

The MCC-KF may be extended to estimate the non-linear system described by equations (3.3) and (3.4). By computing the Jacobian matrices derived from equations (3.3) and (3.4), the MCC-KF may be linearised with a first-order Taylor expansion; an adaptation that is analogous to the EKF:

Algorithm 4. EMCC-KF (*Improved conventional version*)

INITIALIZATION: $(k=0)$ $\hat{x}_{0|0}=\bar{x}_0$ and $P_{0|0}=\Pi_0$.

TIME UPDATE: $(k=\overline{1,N})$ ▷ PRIORI ESTIMATION

$\quad \hat{x}_{k|k-1}=\mathcal{F}(\hat{x}_{k-1|k-1},u_{k-1},0)$,

$\quad P_{k|k-1}=F_{k-1}P_{k-1|k-1}F_{k-1}^T+G_{k-1}QG_{k-1}^T$.

MEASUREMENT UPDATE: $(k=\overline{1,N})$ ▷ POSTERIORI ▷ ESTIMATION

$$L_k=\frac{G_\sigma\left(\|y_k-H_k\hat{x}_{k|k-1}\|_{R^{-1}}\right)}{G_\sigma\left(\|\hat{x}_{k|k-1}-F_{k-1}\hat{x}_{k-1|k-1}-Bu_{k-1}\|_{P_{k|k-1}^{-1}}\right)},$$

$\quad R_{e,k}=H_kP_{k|k-1}L_kH_k^T+R$,

$\quad K_k^L=P_{k|k-1}L_kH_k^TR_{e,k}^{-1}$,

$\quad \hat{x}_{k|k}=\hat{x}_{k|k-1}+K_k^Le_k, \quad e_k=y_k-\mathcal{H}(\hat{x}_{k|k-1},0)$,

$\quad P_{k|k}=(I-K_k^LH_k)P_{k|k-1}$.

Currently, there are many ways to extend the MCC-KF, with varying forms being proposed in the literature (Yang and Huang (2017), Liu *et al.* (2016)). This paper uses the algorithm described as the MCC-EKF, in the paper by Yang and Huang (2017). To arrive at the EMCC-KF used in this paper, the improvements suggested by Kulikova (2016) may be applied to the MCC-EKF. The MCC-KF differs from the EMCC-KF two ways. Firstly, F_k, G_k and H_k are Jacobians in the latter case. Lastly, e_k is the difference between the observation and a non-linear function in the EMCC-KF.

3.3 State-space representation of the Tests of Efficiency

3.3.1 The TEE

The representation of econometric models into state-space form has received increasing attention in the past 20 to 25 years; see the comprehensive survey published recently in Wilcox and Hamano (2017). The Markovian nature of state space models and efficient recursive computations they require make such models attractive to econometricians. Given that the econometrician is sufficiently informed about the model's system dynamics, state space models allow for optimal estimation of unobserved components via the adaptive filters, optimal prediction and estimation of unknown model parameters through the maximisation of a likelihood function (Harvey, 1987).

The TEE model, which is specified in equations (2.3) - (2.5) may be represented in the following way:

$$\begin{bmatrix} h_k \\ \beta_{0,k} \\ \beta_{1,k} \end{bmatrix} = \underbrace{\begin{bmatrix} b_1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{F(\theta)} \begin{bmatrix} h_{k-1} \\ \beta_{0,k-1} \\ \beta_{1,k-1} \end{bmatrix} + \underbrace{\begin{bmatrix} a_0 & a_1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{B(\theta)} \underbrace{\begin{bmatrix} 1 \\ e_{k-1}^2 \end{bmatrix}}_{u_{k-1}}$$

$$+ \underbrace{\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}}_{G(\theta)} \begin{bmatrix} w_{0,k-1} \\ w_{1,k-1} \end{bmatrix},$$

$$y_k = \underbrace{\begin{bmatrix} \delta & 1 & y_{k-1} \end{bmatrix}}_{H(\theta)} \begin{bmatrix} \hat{h}_{k|k-1} \\ \beta_{0,k} \\ \beta_{1,k} \end{bmatrix} + e_k, \quad e_k \sim \mathcal{N}(0, \underbrace{\hat{h}_{k|k-1}}_{R(\theta)})$$

where the process noise is

$$\begin{bmatrix} w_{0,k} \\ w_{1,k} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \underbrace{\begin{bmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_1^2 \end{bmatrix}}_{Q(\theta)} \right),$$

The process $\{h_k\}_{k=0}^N$ is the hidden volatility process, given by the GARCH(1,1) equation (2.4), which needs to be estimated from the given returns data $\{y_k\}_{k=0}^N$. The core result of the TEE methodology is the time-varying coefficients $\beta_{0,k}$ and $\beta_{1,k}$, which need to be estimated together with h_k. More precisely, the *evolution* of the slope coefficient $\beta_{1,k}$ reflects the time-varying change in weak-form market efficiency. If $\beta_{1,k}$ equals to zero (within its confidence intervals), then the TEE yields the conclusion that this market is weak form efficient.

It is stressed that e_k are return residuals, i.e. from equation (2.3) we have $e_k = y_k - \beta_{0,k} - \beta_{1,k} y_{k-1} - \delta h_t$. As mentioned in Hall (1991), the initial values for the state vector and e_0 are assumed to be known and, hence, the new residual is $\hat{e}_1 = r_1 - \hat{\beta}_{0,1|0} - \hat{\beta}_{1,1|0} y_0 - \delta \hat{h}_{1|0}$; see equation (9) in Hall (1991). The computed value $\hat{e}_1$ is, then, used in equation (2.4) to obtain the new predicted estimate $\hat{h}_{2|1}$, i.e. $\hat{e}_{k-1}$ is considered to be a known input at time instance k.

3.3.2 The GTEE

$$\begin{bmatrix} h_k \\ \beta_{0,k} \\ \beta_{1,k} \end{bmatrix} = \underbrace{\begin{bmatrix} b_1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{F(\theta)} \begin{bmatrix} h_{k-1} \\ \beta_{0,k-1} \\ \beta_{1,k-1} \end{bmatrix} + \underbrace{\begin{bmatrix} a_0 + a_1 e_{k-1}^2 \\ 0 \\ 0 \end{bmatrix}}_{u_{k-1}}$$

$$+ \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{G(\theta)} \begin{bmatrix} w_k \\ w_{0,k-1} \\ w_{1,k-1} \end{bmatrix},$$

$$y_k = \underbrace{\begin{bmatrix} \delta & 1 & y_{k-1} \end{bmatrix}}_{H(\theta)} \begin{bmatrix} f(\hat{h}_{k|k-1}) \\ \beta_{0,k} \\ \beta_{1,k} \end{bmatrix} + e_k, \quad e_k \sim \mathcal{N}(0, \underbrace{\hat{h}_{k|k-1}}_{R(\theta)})$$

where the process noise is

$$\begin{bmatrix} w_k \\ w_{0,k} \\ w_{1,k} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \underbrace{\begin{bmatrix} \sigma_\omega^2 & 0 & 0 \\ 0 & \sigma_0^2 & 0 \\ 0 & 0 & \sigma_1^2 \end{bmatrix}}_{Q(\theta)} \right),$$

3.4 Quasi-maximum likelihood estimation of unknown system parameters

Econometric state-space models are often parameterized. This means that the entries of system matrices $\{F, H, R, Q, B\}$ may depend on the unknown system parameter vector $\theta \in \mathbb{R}^p$, which needs to be estimated together with the dynamic state x_k from only the available signal $Y_N = \{y_k\}_{k=0}^N = \{y_0, \ldots, y_N\}$. In practice, the system parameters θ_i, $i = 1, \ldots, p$ are usually estimated by the method of maximum likelihood. If an econometric time series model is represented in the state-space form, then the computation of likelihood function can be done through the KF recursion, because it enables the likelihood function to be broken down in terms of one-step-ahead prediction errors Harvey (1989):

$$\ln L(\theta|Y_N) = \sum_{k=1}^{N} \ln p(y_k|Y_{k-1})$$

$$= -\frac{mN}{2} \ln 2\pi - \frac{1}{2} \sum_{k=1}^{N} \left\{ \ln(\det R_{e,k}) + e_k^T R_{e,k}^{-1} e_k \right\} \tag{3.5}$$

where $Y_N = \{y_0, \ldots, y_N\}$ is N-step measurement history and $\{e_k\}$ are the innovations, generated by the underlying filtering algorithm.

Derivation of formula (3.5) is based on the assumption of optimality of the KF estimator for Gaussian state-space models with the property $e_k \sim \mathcal{N}(0, R_{e,k})$; see Schweppe (1965). However, in non-Gaussian settings, the KF exhibits only sub-optimal behaviour, i.e. the KF solution yields minimum mean *linear* square estimators rather than minimum mean square estimators, and the above mentioned property is violated. Hence, the use of formula (3.5) in non-Gaussian settings yields the quasi-maximum likelihood estimation (Ruiz (1994),Harvey *et al.* (1994)).

A significant amount of research effort has been committed to uncovering necessary and sufficient conditions for consistent estimation with the quasi-maximum likelihood estimator (QMLE). Although the Methods of Moments Estimator may produce more asymptotically efficient estimator than the QMLE under non-normality, the QMLE is often simpler to implement and requires less non-parametric estimation (Bollerslev and Wooldridge, 1992). For a pure GARCH process, Francq *et al.* (2004) shows that under some fairly standard assumptions, the QMLE is consistent. Extending this proof to the ARMA-GARCH process only requires the additionally assumption that $\mathbb{E}(e_k) = 0$ and then consistency becomes a consequence. Finally, the QMLE is asymptotically normal when the sixth order moments of e_k exist and stationarity restrictions are imposed over the entire parameter space (Francq and Zakoian, 2011).

3.5 Simulation Study

In order to assess the efficacy of the three tests along with their associated filters, a simulation study is proposed. The purpose of the study is three-fold. Firstly, the study aims to evaluate the filters' respective abilities to estimate the time-varying sequence

$$\{\beta_{1,k}\}_{k=1}^{n}$$

in a non-Gaussian system, where n represents the number of time steps in a single path. The study, secondly, is interested in how well each filter recovers GARCH-M parameters. Aside from inferring a market's path towards efficiency, the tests of efficiency also provide information about the relationship between risk and return for instance; through the GARCH-M formulation.

In the study, a single $\{\beta_{1,k}\}_{k=1}^{n}$ path is simulated $m = 100$ times. Only one $\beta_{1,k}$ function is used throughout the investigation. To test the robustness of the filters, a $\beta_{1,k}$ function which does not follow the random walk specifications of equations

(2.5) and (2.10) is selected. This is done since the aim of the filters is to discover a true $\beta_{1,k}$ function for a given returns series, and therefore this function may take on any form. The $\beta_{1,k}$ function in this study is non-linear so that it may be thus be determined whether the tests can capture non-standard paths:

$$\beta_{1,k} = \begin{cases} 0, & 0 \leq k \leq \frac{n}{2} \\ \frac{8k^2}{n^2} - \frac{8k}{5} + \frac{2}{5}, & \frac{n}{2} \leq k \leq n \end{cases} \tag{3.6}$$

For a single iteration m, the $\{\beta_{1,k}\}_{k=1}^{n}$ path generated by (3.6) is incorporated into the returns process. Where returns are generated according to the specification of a particular test of efficiency, that same test is used to recover the $\{\beta_{1,k}\}_{k=1}^{n}$ path from said returns. For example, where the GTEE is being assessed, the returns will be generated by the equations (2.8) and (2.9), which are implied by the test. Thus, it is assumed that the test correctly specifies the returns generating process.

Following a similar procedure to the Monte Carlo Test of (Zalewska-Mitura and Hall, 1999), the TEE, the GTEE with $f(h_t) = h_t$, the GTEE with $f(h_t) = \log(h_t)$ and the the GTEE with $f(h_t) = \sqrt{h_t}$ are used to estimate a pre-specified $\beta_{1,k}$ function, in this case equation (3.6). This procedure is repeated $m = 100$ times and each time an estimate of the path of $\beta_{1,k}$ and a 99% confidence interval are produced. Hence, a sequence $\{\hat{\beta}_{1,k}\}_{k=1}^{n}$ of $\beta_{1,k}$ estimates is calculated. For each iteration $j \in \{1, \ldots, m\}$, both the MCC-KF and the KF are applied in the estimation of an entire $\beta_{1,k}$ path. The parameters are not pre-specified in the filters and are therefore estimated using the QMLE approach. The log-likelihood value and associated parameter estimates for each filter are recorded.

The most crucial part of the investigation is that the innovations are non-Gaussian. To test the robustness of the (E)MCC-KF, the distribution of the innovations ϵ_k may conditionally follow any non-Gaussian law. In most of the MCC-KF literature, Gaussian mixture errors are usually chosen to highlight the robustness of the MCC-KF in engineering-based systems. However, in a finance context, it is desirable that the distribution be as leptokurtic as possible while still conforming to the framework implied by the GARCH models. The exposition that follows solves this problem by showing how one might generate mixture Gaussian GARCH returns. Firstly, it is noted that the mean equation for the Gaussian GARCH-M (2.1) can be formulated alternatively as follows:

$$r_t = \gamma^\top x_t + \delta h_t + \sqrt{h_t} z_t, \quad z_t \sim N(0,1) \tag{3.7}$$

where z_t is an i.i.d. random variable generated from the standard normal distribution and h_t has the usual definition (2.2). It can be easily seen that ϵ_t in (2.1) is equivalent to $\sqrt{h_t}z_t$ in (3.7). Ausín and Galeano (2007) state that in order to generate innovations ϵ_t from the Gaussian mixture distribution – where the ϵ_t have h_t as their conditional variance, then one may generate i.i.d. z_t as a mixture of two Gaussian distributions with the following dynamics:

$$z_t \sim \begin{cases} N(0, \sigma^2), & \text{with probability } \rho \\ N(0, \frac{1}{\lambda}\sigma^2), & \text{with probability } 1-\rho \end{cases} \tag{3.8}$$

where $0 < \lambda < 1; 0 < \rho < 1$ and,

$$\sigma^2 = \frac{1}{\rho + \frac{1-\rho}{\lambda}} \tag{3.9}$$

so that $\text{Var}(z_t) = 1$.

Thus, substituting these z_t back into (3.7) would generate highly leptokurtic returns. This paper chooses to generate perturbations $z_t \sim$ mixture Gaussian$(0.05, 0.9)$ where the TEE is simulated with the following parameters

$$\theta^1 = (\delta, a_0, a_1, b_1) = (0.2, 0.1, 0.1, 0.8)$$

.

The parameter vectors for the GTEE with $f(h_t) = h_t$, $f(h_t) = \log(h_t)$ and $f(h_t) = \sqrt{h_t}$ are respectively given by:

$$\theta^h = \theta^{\log} = \theta^{sq} = (\delta, a_0, a_1, b_1, \sigma^2_\omega) = (0.2, 0.2, 0.1, 0.8, 0.001)$$

.

Parameters (σ_0^2, σ_1^2) are not set in the simulation since they are implied by the respective test and its associated estimation procedure. The initial hidden state vector and its covariance are given by

$$\mathbf{x_0} = \begin{bmatrix} h_0 \\ \beta_{0,0} \\ \beta_{1,0} \end{bmatrix} = \begin{bmatrix} \text{Var}(y) \\ \bar{y} \\ \hat{\beta}_{1,0} \end{bmatrix}, \quad \mathbf{P_{0|0}} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.05 \end{bmatrix}.$$

where $\bar{y}$ and $\text{Var}(y)$ are the sample mean and covariance of the simulated returns y. The initial $\hat{\beta}_{1,0}$ is the AR(1) coefficient that results after fitting an AR(1) model to y. Choosing the initial values as statistics that can be calculated from the observations yields a more reproducible methodology in practice. Moreover, because the statistics always misspecify the true initial values, the simulation study can assess how long the burn-in period may need to be when estimating efficiency in real data. In accordance with the size of the data that was collected, $n = 4900$

returns are generated for a given path. The study finds that although n observations are generated, the first 200 $\beta_{1,k}$ estimates should be excused as part of the burn-in period. Thus, only 4700 estimates are compared to the true value of $\beta_{1,k}$. After rounding, 5000 S&P500 and ALSI data points were collected but allowing for a burn-in period of 250 data points roughly gives the 4700 data points to be estimated in the simulation.

Table (3.1) shows the the average MSE for each filter when applied to the different tests. The results confirm findings in the literature; the MCC-KF and EMCC-KF are more robust than the KF and EKF in the presence of extremely non-Gaussian perturbations. What has not been frequently tested in the literature, on the other hand, is how well the correntropy filter collaborates with the QMLE technique to estimate system parameters. The average QMLE estimates associated with each test are shown in Table (3.2) . Figure 3.1 compare the average $\{\hat{\beta}_{1,k}\}_{k=1}^{n}$ estimates from both filters against the actual $\{\beta_{1,k}\}_{k=1}^{n}$ path for the log-GTEE. It is clear that the MCCKF produces more robust estimates of $\{\beta_{1,k}\}_{k=1}^{n}$.

Average MSE over 100 simulations				
Filter	TEE	GTEE (h_t)	GTEE ($\log(h_t)$)	GTEE ($\sqrt{h_t}$)
KF	0.0037	–	–	–
MCC-KF	0.0035	–	–	–
EKF	–	0.0042	0.0054	0.0048
EMCC-KF	–	0.0040	0.0042	0.0040

Tab. 3.1: Performance of MCC-KF compared to KF

Calculating Gaussian confidence intervals for the Kalman filter is simple, since the Wald standard errors can be constructed in conjunction with the filters' error covariances. These filters contain most of the $\{\beta_{1,k}\}_{k=1}^{n}$ on average, as shown by Figure (3.2). Deriving confidence intervals for the MCC-KF, however, is more difficult since there are no distributional assumptions associated with the MCC-KF. In fact, results from the simulations indicate that the covariance from MCC-KF is usually larger than KF covariance. This paper intends to use the results from the simulation study to arrive at an approximate formula for calculating symmetric two-sided 99% confidence intervals associated with the MCC-KF. The intervals will be calculated as follows:

$$CI(\hat{x}_{k|k})_{99\%} = \hat{x}_{k|k} \pm crit\sqrt{P_{k|k}}$$

where $CI(\hat{x}_{k|k})_{99\%}$ is the 99% confidence interval for the posterior estimate $\hat{x}_{k|k}$, $crit$ is the associated 99% critical value and $P_{k|k}$ is the covariance of the estimation error of $\hat{x}_{k|k}$. Cox and Hinkley (1979) interpret confidence intervals (in their case the 90% confidence interval) in terms of samples: ”Were this procedure to be repeated on numerous samples, the fraction of calculated confidence intervals (which would differ for each sample) that encompass the true population parameter would tend toward 90%”. In the simulated data, a search through different critical values and their resulting bounds iskn first performed. The smallest critical value in the data that produces intervals encompassing the true $\{\beta_{1,k}\}_{k=1}^{n}$ for 99 percent of the observations in an average single path is 2. Considering that the returns generated from this simulation study are highly non-Gaussian, setting the critical value to 2 is a conservative choice for real data.

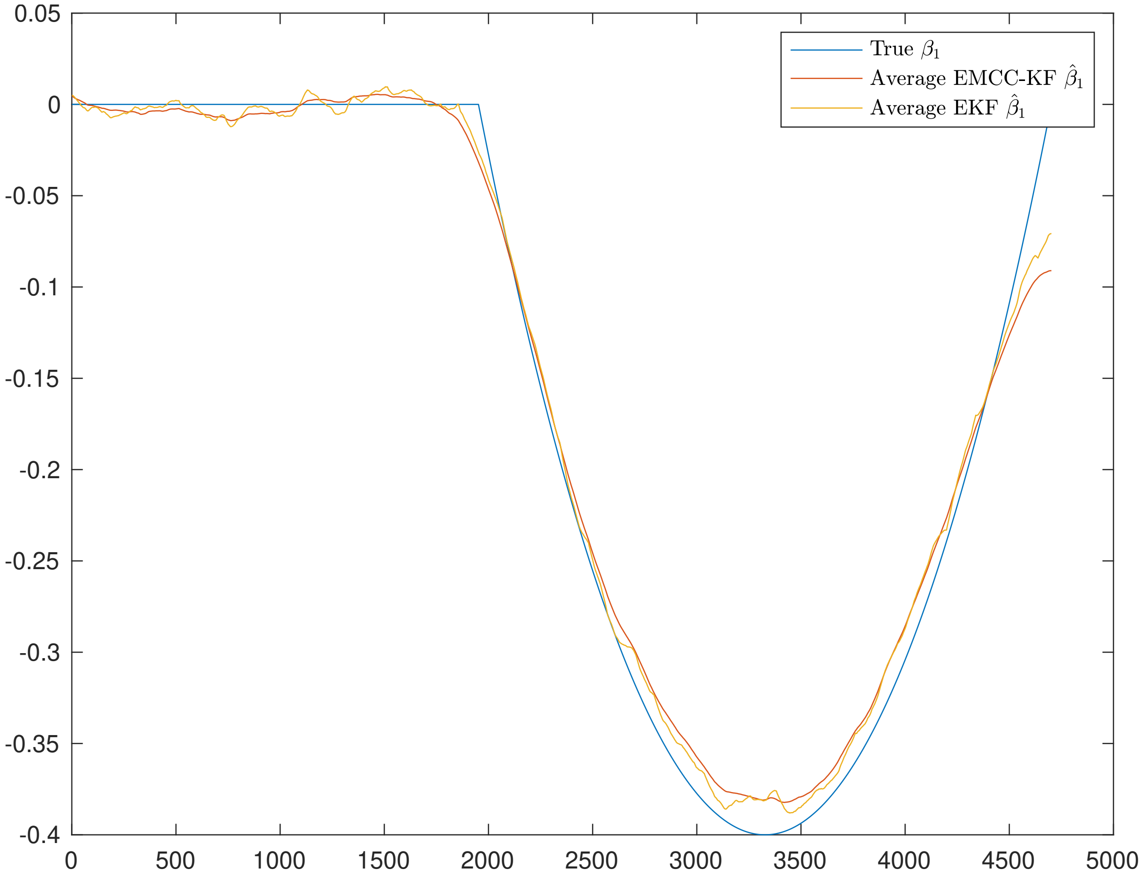

Fig. 3.1: Robustness of the MCC-KF for GTEE($f_v = \log h_t$)

QMLE Estimates									
Test	Filter	Parameters							AIC
TEE		δ	a_0	a_1	b_1	Q_0	Q_1	Q_ω	
	True	0.2	0.1	0.1	0.8	–	–	–	
	KF	0.224	0.1169	0.0932	0.789	5.64e−7	5.85e−5	–	12959
		(0.0774)	(0.0369)	(0.0293)	(0.0541)	(1.01e−6)	(2.70e−5)	–	
	MCC-KF	0.225	0.1169	0.0932	0.789	7.69e−7	5.59e−5	–	12964
		(0.0779)	(0.0363)	(0.0294)	(0.0533)	(1.42e−6)	(2.59e−5)	–	
GTEE $f_v = h_t$		δ	a_0	a_1	b_1	Q_0	Q_1	Q_ω	
	True	0.2	0.2	0.1	0.8	–	–	0.001	
	EKF	0.2265	0.232	0.0926	0.787	1.10e−6	9.09e−5	0.0187	16592
		(0.0241)	(0.0597)	(0.0241)	(0.0427)	(7.44e−6)	(8.96e−5)	(0.0230)	
	EMCC-KF	0.222	0.233	0.0939	0.7862	9.13e−7	9.22e−5	0.0215	16603
		(0.0507)	(0.0622)	(0.0236)	(0.0440)	(2.93e−6)	(5.54e−5)	(0.0243)	
GTEE $f_v = \log(h_t)$		δ	a_0	a_1	b_1	Q_0	Q_1	Q_ω	
	True	0.2	0.2	0.1	0.8	–	–	0.001	
	EKF	0.198	0.243	0.096	0.777	1.21e−6	1.61e−4	0.0238	16591
		(0.0272)	(0.0661)	(0.0484)	(0.0909)	(3.37e−6)	(2.18e−4)	(0.0229)	
	EMCC-KF	0.181	0.247	0.098	0.775	2.11e−6	1.04e−4	0.0227	16601
		(0.0279)	(0.0654)	(0.0478)	(0.0862)	(7.10e−6)	(6.06e−5)	0.0232	
GTEE $f_v = \sqrt{h_t}$		δ	a_0	a_1	b_1	Q_0	Q_1	Q_ω	
	True	0.2	0.2	0.1	0.8	–	–	0.001	
	EKF	0.1998	0.2455	0.0970	0.775	5.91e−7	1.35e−4	0.0240	16585
		(0.0988)	(0.0882)	(0.0301)	(0.0670)	(2.23e−6)	(2.16e−4)	(0.0238)	
	EMCC-KF	0.182	0.250	0.0990	0.773	9.32e−7	1.08e−4	0.0223	16594
		(0.0851)	(0.0921)	(0.0325)	(0.0694)	(2.09e−6)	(8.59e−5)	(0.0247)	

Tab. 3.2: QMLE Estimates for each test

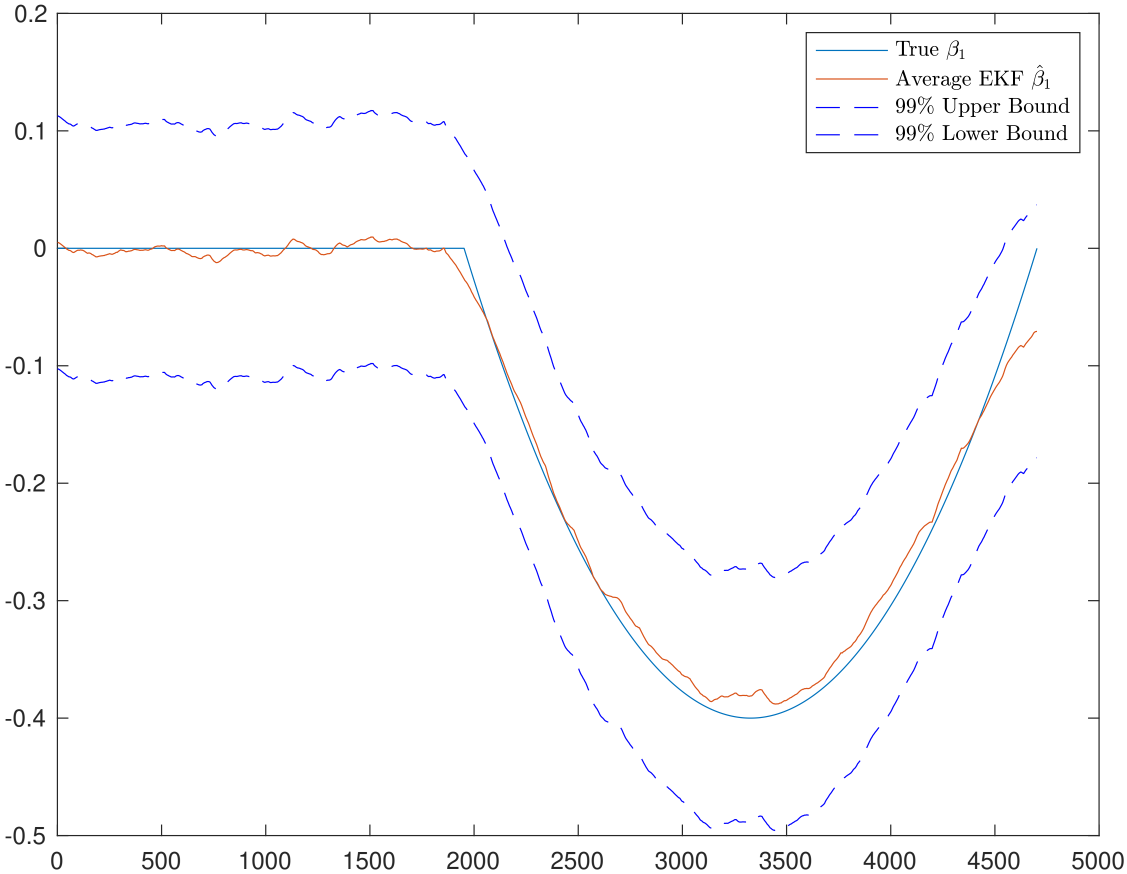

Fig. 3.2: Estimating $\beta_{1,k}$ within 99% confidence interval

Chapter 4

Empirical results

4.1 The Data and their descriptive statistics

4.1.1 Data Description

The data sets used in this investigation are the daily closing prices of the S&P500 and the FTSE/JSE All Share Index (ALSI). The indices represent the price levels of a mature and a younger market respectively. Both time series begin on the 29th of November in 1998 and contain dividend adjusted last prices that run until the 15th of November 2018. The data were obtained from Bloomberg.

Comparability between the indices is essential in this investigation. Considering that the S&P500 represents many of the largest the common stocks listed on the New York Stock Exchange (NYSE) and the NASDAQ, a more encompassing index such as the ALSI would be a more comparable reflection of the South African stock market than the Top 40 index. Moreover FTSE Russell (2019) constructs the ALSI so that it is both investable and tradable. The two indices included in this paper are also both weighted using free-float market capitalisation.

4.1.2 Summary Statistics

The daily prices from the S&P500 (S_t) and the ALSI (X_t) are differenced and converted to continuously compounded log-returns,

$$r_t = \log(S_t/S_{t-1}), \qquad y_t = \log(X_t/X_{t-1}).$$

A summary of the returns from each index is provided in Table (4.1). For most periods, the returns registered by the S&P500 are smaller in absolute magnitude than those found on the ALSI. This explains why the S&P500 has a smaller inter-quartile range than the ALSI. However, the S&P500 experienced considerably greater volatility during the global financial crisis of 2008. In fact, between October 2008 and May 2009, r_t and y_t recorded their respective minimums and maximums within the sample period. The global financial crisis disproportionately affected the indices, with the S&P500 registering a lower minimum daily return (-9.47%) and a higher maximum return (10.96%) than the ALSI. There are many clusters in both time series but the most significant ones occur during the crisis of 2008. Figures (4.1) and (4.2) show that there may have been clustering in the the ALSI in 2006 and in the S&P500 in 2011 . The presence of clusters in the data suggest that r_t and y_t are suitable for GARCH modelling. Furthermore, the sample kurtosis found in both the ALSI and the S&P500 is substantially greater than the normal value of three. Unconditionally leptokurtic returns and volatility clustering are features of the data that would be suitably described by the GARCH(p, q) model.

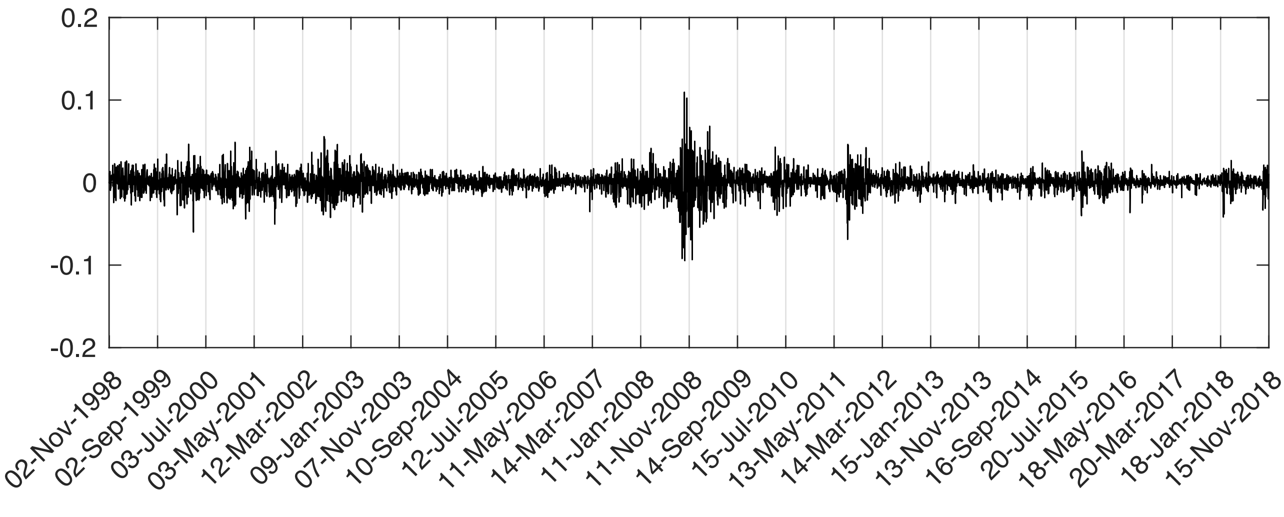

Fig. 4.1: S&P500 Daily Returns

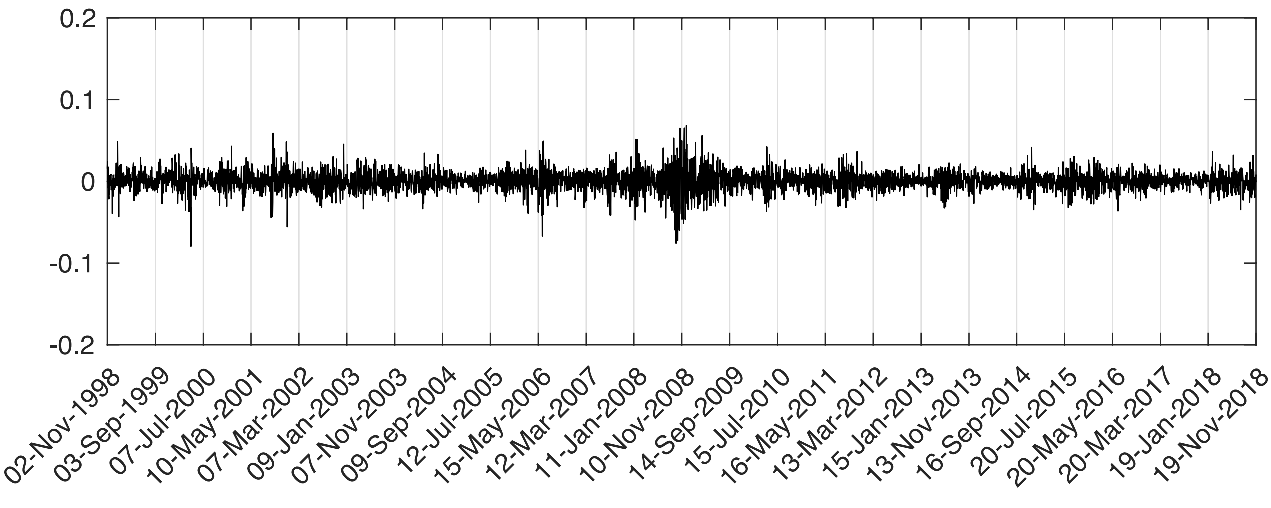

Fig. 4.2: FTSE/JSE All Share Daily Returns

Index	Min	1st Qu	Median	Mean	3rd Qu	Max	Std	Skew	Ex Kurt
S&P500	-0.0947	-0.00492	0.000522	0.000183	0.00582	0.110	0.0120	-0.216	11.2
ALSI	-0.0795	-0.00594	0.000679	0.000459	0.00711	0.0683	0.0119	-0.168	6.54

Tab. 4.1: Summary Statistics of S&P500 and ALSi indices

4.1.3 Testing for Dependency Structures

Firstly, r_t and y_t are tested for stationarity. The Dickey-Fuller Test fails to find a unit root in either returns series (at the 1% significance level), which suggests that the time series are stationary. Thereafter, the data are checked for possible autocorrelations. Figure (4.3) displays the sample ACF's for the returns and squared returns of the S&P500 and the ALSI indices. The top correlograms indicate that r_t and y_t exhibit significant first-order autocorrelation. Fitting an AR(1) model to the S&P500 returns results in an AR(1) coefficient that is negative and significant at the 1% level. On the other hand, when the same model is fitted the ALSI returns, a positive and significant AR(1) coefficient is estimated.

The bottom correlograms for the squared returns suggest that neither r_t nor y_t are serially independent time series. By applying the Ljung-Box test of non-correlation to the squared returns series, the data is tested for ARCH effects. The Ljung-Box statistic for each index is significant at the 5% level, indicating that there might be volatility clustering in each returns series.

Volatility clustering is characteristic of returns that are generated by the (G)TEE. Furthermore, while the leptokurtic behaviour of each index's returns may be a consequence of conditional heteroskedasticity, it is plausible that the kurtosis is being influenced by a conditionally non-Gaussian distribution as well. This may be particularly true for the S&P500 returns, whose unconditional kurtosis (11.2317) greatly exceeds that of the ALSI and the normal distribution. The data may therefore be appropriately modelled by both the TEE and the GTEE, where each framework is estimated by the standard KF and the MCCKF.

4.2 Testing adaptive market efficiency

The analysis presented in the previous section sustains that the TEE and GTEE models may be suitably applied to the data. In this section, the $\{\beta_{1,k}\}_{k=1}^{n}$ paths are extracted from r_t and y_t, whereafter their shape is interpreted. Four tests will be applied to each index - the standard TEE and the GTEE with the following feedback functions: $f_v = h_t$, $f_v = \sqrt{h_t}$ and lastly the GTEE with $f_v = \log h_t$.

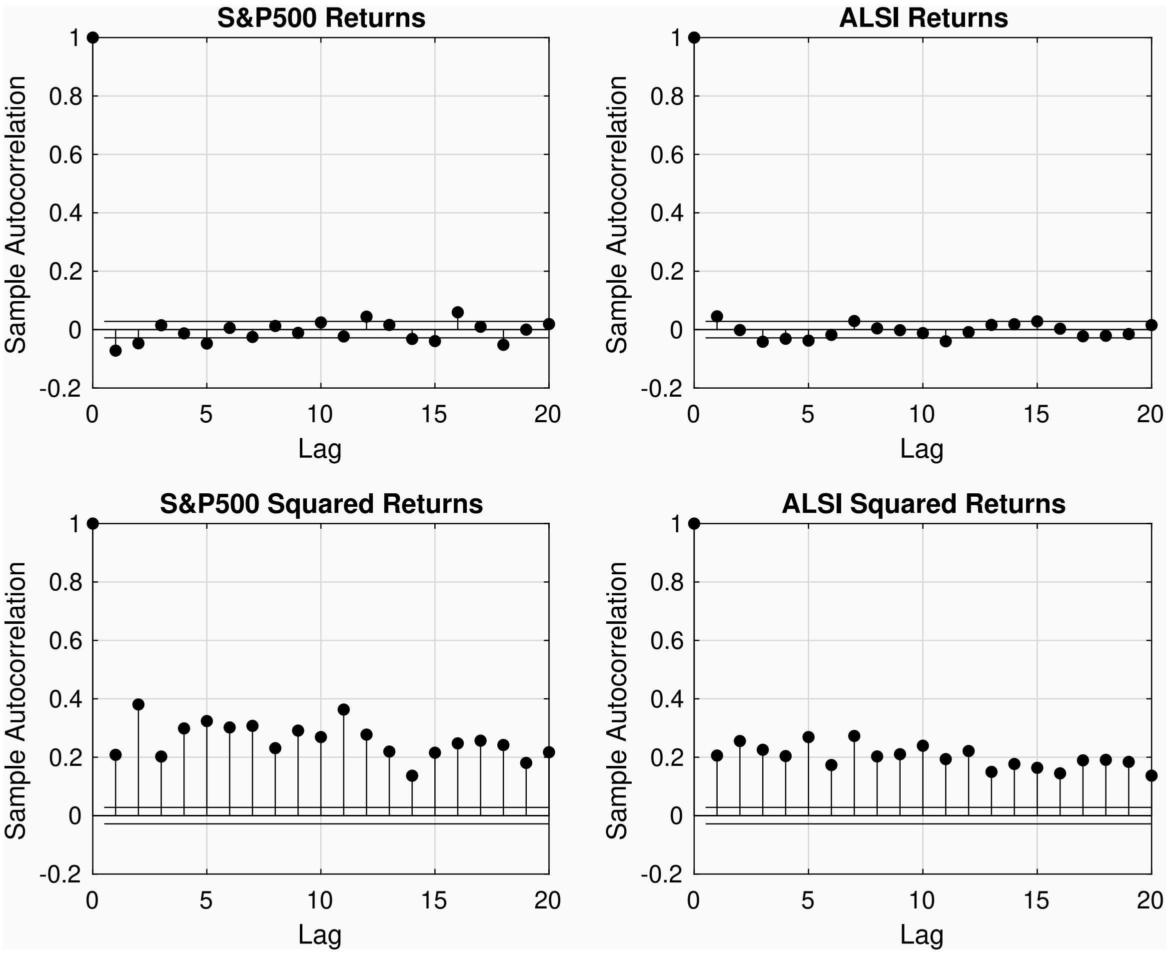

Fig. 4.3: Correlograms for Indices Returns and Squared Returns

Table (4.2) shows the AIC values corresponding to each test and applied to the two indices. The standard Kalman filter TEE has the lowest AIC value when applied to the ALSI index. On the other hand, the GTEE ($f_v = \sqrt{h_t}$) with an associated MCC-KF achieves the lowest AIC among the tests of efficiency for the S&P500. Since this particular volatility function was preferred over the default $f_v = h_t$, we may infer that it is more plausible that there is a linear relationship between returns and volatility rather than variance on the American stock markets.

Test of efficiency	S&P500			ALSI		
	(E)KF	(E)MCC-KF	Best-fit	(E)KF	(E)MCC-KF	Best-fit
TEE	13951.65	13592.09	KF	14755.00*	14762.04	KF
GTEE ($f_v = h_t$)	13953.64	13951.38	EMCC-KF	14757.11	14761.53	EKF
GTEE ($f_v = \log(h_t)$)	13954.29	13952.39	EMCC-KF	14758.96	14763.33	EKF
GTEE ($f_v = \sqrt{h_t}$)	13952.28*	13949.42	EMCC-KF	14758.00	14762.57	EKF

Tab. 4.2: AIC values for TEE and GTEE for S&P500 and ALSI

The parameter estimates (and t-statistics) for the best performing tests are given in Table (4.3). At a significance level of 1%, the assumption that $\hat{\delta} = 0$ for either the ALSI or S&P500 returns cannot be sustained. In fact, the estimated value of $\hat{\delta}$ for the S&P500 and the ALSI suggests that there is a positive relationship between the expected returns of the indices under investigation and their expected volatility. It is reasonable, therefore, to conclude that investors in the American markets and the JSE are risk-averse and they require a higher return premium when they expect volatility to increase. Note that $\hat{\alpha} + \hat{\beta} < 1$, which is consistent with the stationarity of the data. In addition, $\hat{\alpha} + \hat{\beta}$ is close to 1 which indicates strong persistence of shocks for both series. Regardless, $\hat{\alpha} + \hat{\beta}$ is higher for the ALSI than for the S&P500. A close inspection of figures (4.2) and (4.1) reveals that volatility shocks for the ALSI seem to decay less rapidly than for the S&P500, which sustains the turbulence of 2008.

Figures (4.4) and (4.5) show the estimated $\beta_{1,k}$ for the ALSI and S&P500 respectively along with their respective 99% confidence intervals. From the start of the sample period until August 2004, the $\{\hat{\beta}_{1,k}\}_{k=1}^{n}$ for the ALSI is significantly greater than zero and declining. Thereafter, there is no evidence that $\{\beta_{1,k}\}_{k=1}^{n}$ is significantly non-zero. Thus, Figures (4.4) implies that the JSE may have been inefficient between November 1998 and August 2004 and was then efficient from then till November 2018. These findings go against the literature — an earlier study by Jefferis and Smith (2005) reported that the JSE was efficient between January 1990 and June 2001 while Kulikova *et al.* (2019) arrived at the same conclusions but for the Top 40 index starting in 2002 going up to 2012. The divergence from previous findings may be explained by the longer time period under which efficiency is tested in this paper.

Index	$\hat{a}_0$	$\hat{a}_1$	$\hat{b}_1$	$\hat{\delta}$	$\hat{\sigma}_0^2$	$\hat{\sigma}_1^2$	$\hat{\sigma}_\omega^2$
S&P500	0.0269	0.1198	0.8588	0.1521	7.35e−6	2.21e−6	8.86e−11
	(10.33)	(34.13)	(3.02)	(40.18)	(0.934)	(0.309)	(2.54e−8)
ALSI	0.0277	0.1008	0.8799	0.0405	4.73e−11	3.17e−6	−
	(4.23)	(18.61)	(98.51)	(19.98)	(2.07e−5)	(1.311)	(−)

Tab. 4.3: Summary Statistics of S&P500 and ALSi indices

Contrary to the aforementioned results, the American stock market does not exhibit market inefficiency within the sample period. Nevertheless, the degree of efficiency may be changing as evidenced by the non-constant $\{\beta_{1,k}\}_{k=1}^n$ estimates. From 1998, $\{\hat{\beta}_{1,k}\}_{k=1}^n$ gradually decreases until November 2008 where the market comes closest to inefficiency. This period coincides with the beginning of the Financial Crisis, which provides a compelling explanation to the shape of $\{\hat{\beta}_{1,k}\}_{k=1}^n$. In the years that follow the trough in 2008, $\hat{\beta}_{1,k}$ makes its approach towards the zero value.

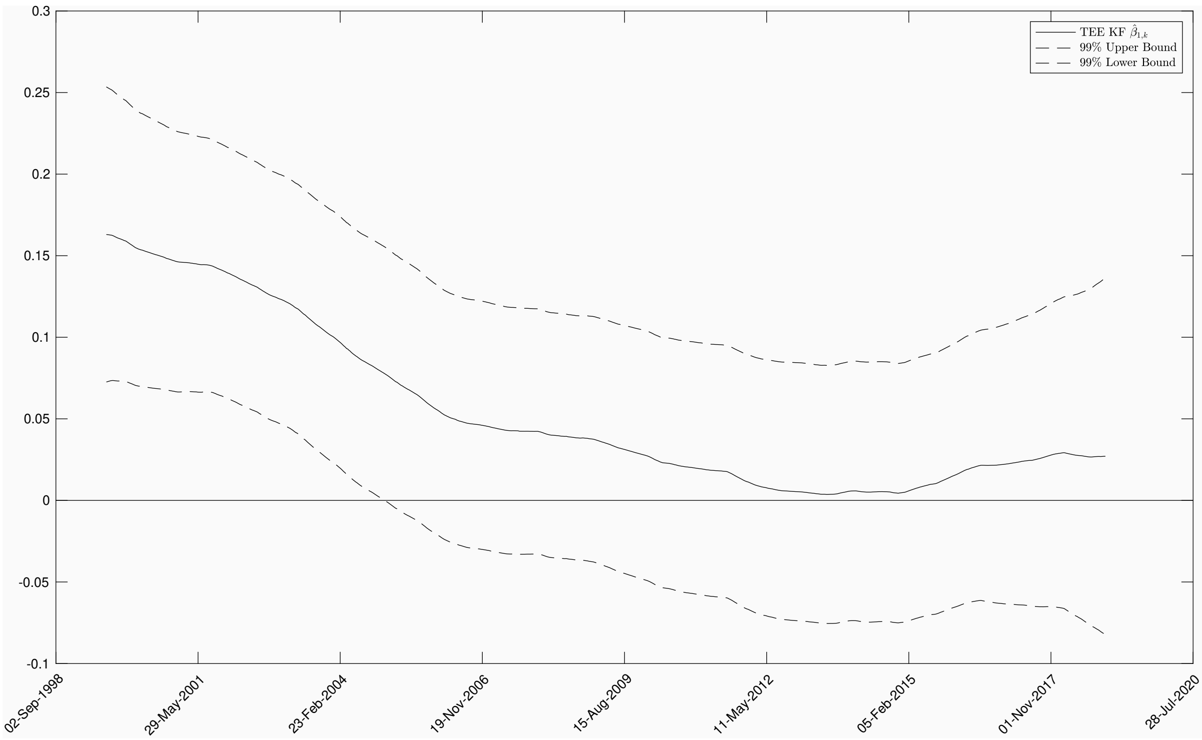

Fig. 4.4: $\hat{\beta}_{1,k}$ for ALSI

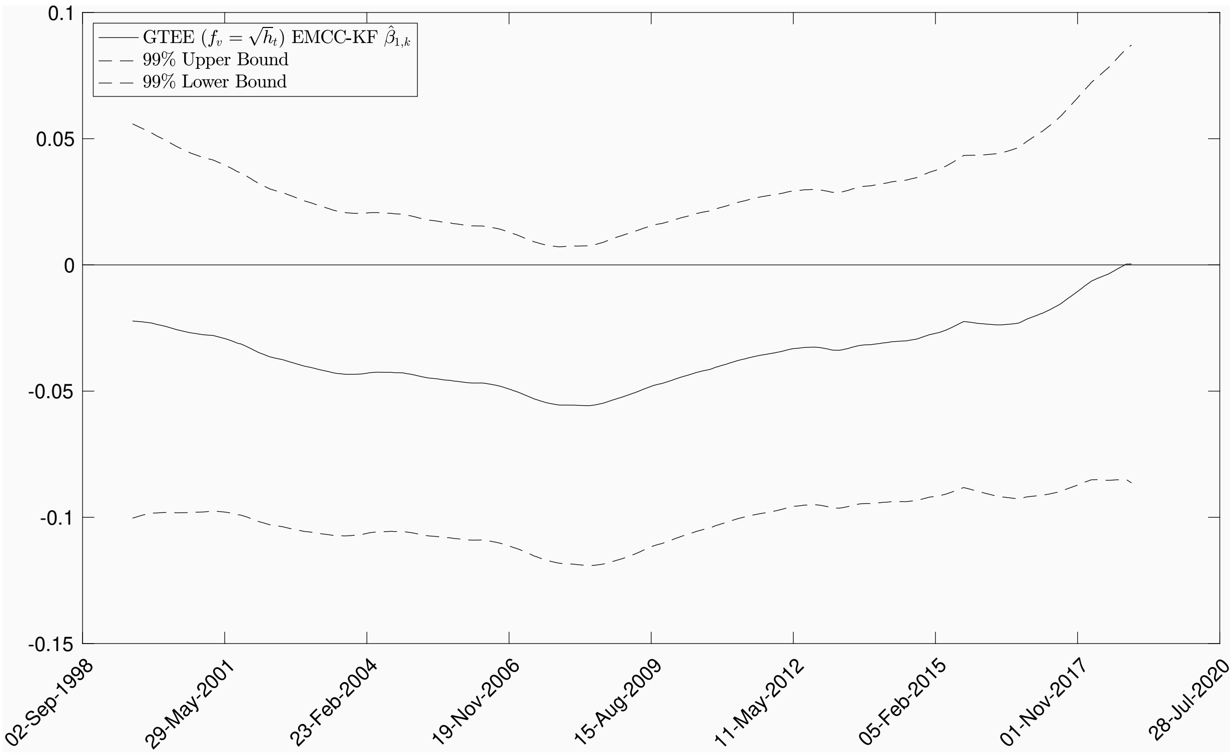

Fig. 4.5: $\hat{\beta}_{1,k}$ for $S\&P500$

4.3 Testing goodness-of-fit of the TEE and GTEE

After finding ARCH effects in Section 4.1.3 and then correctively fitting a GARCH based model, it is important to check that the model residuals are free from conditional heteroskedasticity. It is also imperative to investigate whether the serial correlation shown in Figure 4.3 was removed by the models. In this section, the goodness-of-fit of the best performing test for each index is evaluated.

A series of model residuals $\{e_t\}_{t=1}^{n}$ is calculated by subtracting the expected return from the realised return

$$e_t = y_t - \left(\hat{\beta}_{0,t} + \hat{\beta}_{1,t} y_{t-1} + \hat{\delta}\hat{h}_t\right)$$

for the ALSI and

$$e_t = r_t - \left(\hat{\beta}_{0,t} + \hat{\beta}_{1,t} r_{t-1} + \hat{\delta}\sqrt{\hat{h}_t}\right)$$

for the S&P500, where $\{\hat{\beta}_{0,t}\}_{t=1}^{n}$, $\{\hat{\beta}_{1,t}\}_{t=1}^{n}$, $\hat{\delta}$ and $\{\hat{h_t}\}_{t=1}^{n}$ are outputs from each index's respective test of efficiency. The standardised residuals

$$e_t/\sqrt{\hat{h}_t}$$

are then tested for autocorrelation and heteroskedasticity. The images in the upper segment of Figures A.1 and A.2 plot the indices' respective residual time series and their histograms. Although the standardised residuals for the ALSI appear normal, with a sample mean of -0.0055 and sample standard deviation which is close to unity (0.9981), the normality assumption is rejected for by the Jaque-Bera test and the Kolmogorov-Smirnov test for normality. On the other hand, the Kolmogorov-Smirnov test cannot reject the assumption that the residuals are t-distributed i.e $p = 0.44$. Thus, the tests imply that the innovations of the returns process may be non-Gaussian.

The bottom diagrams in Figures A.1 and A.2 show the correlograms of the standardised residuals and for the squared standardised residuals. These diagrams provide no evidence that the ALSI or S&P500 residuals are dependent in time. Formal tests of correlation are also unable to detect any time-dependancy. Applying the Ljung-Box test of the 12th order on the residuals and their squares reveals that the null hypothesis of no correlation cannot be rejected at any reasonable significance level. Table A.1 summarises these test results.

Chapter 5

Conclusion

In this paper, the weak form adaptive levels of efficiency were tested for the JSE and the American stock markets. The TEE and three versions of the GTEE were first tested in a simulation study, where it was shown that the MCC-KF is more robust than the KF at estimating time-dependent autocorrelation coefficients in the presence of non-Gaussian errors. Previous studies on the topic support this finding. What is not often discussed in the literature is how well these filters perform when used in conjunction with the QMLE technique. In this paper's simulation study, the KF consistently recorded lower AIC's than the MCC-KF regardless of the test being applied. This implies that in most cases, the KF is the preferred filter when estimating unknown parameters with the QMLE procedure.

The tests were then applied to the ALSI and the S&P500, each test being implemented with a KF and a MCC-KF. Implementing the GTEE using an MCC-KF yielded a superior AIC for the S&P500 while the standard Kalman filter TEE achieved a lower AIC for the ALSI. The American stock markets were found to be efficient throughout the sample period whereas the South African stock markets showed signs of inefficiency in the earlier part of the data whereafter no evidence of inefficiency could be found.

Estimating stochastic volatility models with the MCC-KF was not considered in this paper but this is left as a topic for future research. Applying the MCC-KF to the study of younger capital markets than those examined in this paper could also form part of future literature. It is possible that any estimation improvement that these modern filters offer would be accentuated in younger markets where infrequent trading can create non-Gaussian returns. Transaction costs would have to be considered in any further studies.

Bibliography

Ausín, M. C. and Galeano, P. (2007). Bayesian estimation of the gaussian mixture garch model, Computational Statistics & Data Analysis **51**(5): 2636–2652.

Bai, X., Russell, J. R. and Tiao, G. C. (2003). Kurtosis of garch and stochastic volatility models with non-normal innovations, Journal of Econometrics **114**(2): 349–360.

Bilik, I. and Tabrikian, J. (2010). Maneuvering target tracking in the presence of glint using the nonlinear gaussian mixture kalman filter, IEEE Transactions on Aerospace and Electronic Systems **46**(1).

Black, F. (1976). Studies of stock price volatility changes.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, Journal of econometrics **31**(3): 307–327.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return, The review of economics and statistics pp. 542–547.

Bollerslev, T., Engle, R. F. and Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances, Journal of political Economy **96**(1): 116–131.

Bollerslev, T. and Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, Econometric reviews **11**(2): 143–172.

Chen, B., Liu, X., Zhao, H. and Principe, J. C. (2017). Maximum correntropy kalman filter, Automatica **76**: 70–77.

Cinar, G. T. and Príncipe, J. C. (2012). Hidden state estimation using the Correntropy Filter with fixed point update and adaptive kernel size, The 2012 International Joint Conference on Neural Networks (IJCNN), pp. 1–6.

Cox, D. R. and Hinkley, D. V. (1979). Theoretical statistics, Chapman and Hall/CRC.

Emerson, R., Hall, S. G. and Zalewska-Mitura, A. (1997). Evolving market efficiency with an application to some bulgarian shares, Economics of planning **30**(2): 75–90.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation, Econometrica: Journal of the Econometric Society pp. 987–1007.

Engle, R., Lilien, D. and Robins, R. (1987). Estimating time varying risk premia in the term structure: The ARCH-M model, Econometrica pp. 391–407.

Fama, E. (1965). The behavior of stock-market prices, The Journal of Business **38**(1): 34–105.

Francq, C. and Zakoian, J.-M. (2011). GARCH models: structure, statistical inference and financial applications, John Wiley & Sons.

Francq, C., Zakoïan, J.-M. et al. (2004). Maximum likelihood estimation of pure garch and arma-garch processes, Bernoulli **10**(4): 605–637.

FTSE Russell (2019).

Ghahramani, M. and Thavaneswaran, A. (2008). A note on garch model identification, Computers & Mathematics with Applications **55**(11): 2469–2475.

Hall, S. (1990). A note on the estimation of GARCH-M models using the Kalman filter, Bank of England Discussion Paper (32).

Hall, S. (1991). An application of the stochastic GARCH-in-Mean model to risk premia in the London Metal Exchange, The Manchester School **59**(S1): 57–71.

Harvey, A. (1989). Forecasting Structural Time Series Models and the Kalman Filter, Cambridge University Press, Cambridge.

Harvey, A. C. (1987). Applications of the Kalman filter in econometrics, Advances in Econometrics, Proc Fifth World Congress of the Econometric Society, Vol. 1, pp. 285–313.

Harvey, A., Ruiz, E. and Shephard, N. (1994). Multivariate stochastic variance models, The Review of Economic Studies **61**(2): 247–264.

Holder, N. (2017). Testing adaptive market efficiency under the assumption of stochastic volatility, PhD thesis, University of Cape Town.

Ito, M., Noda, A. and Wada, T. (2014). International stock market efficiency: a non-bayesian time-varying model approach, Applied Economics **46**(23): 2744–2754.

Ito, M., Noda, A. and Wada, T. (2016). The evolution of stock market efficiency in the us: a non-bayesian time-varying model approach, Applied Economics **48**(7): 621–635.

Izanloo, R., Fakoorian, S. A., Yazdi, H. S. and Simon, D. (2016). Kalman filtering based on the maximum correntropy criterion in the presence of non-Gaussian noise, 2016 Annual Conference on Information Science and Systems (CISS), pp. 500–505.

Jefferis, K. and Smith, G. (2005). The changing efficiency of african stock markets, South African Journal of Economics **73**(1): 54–67.

Jorion, P. (1988). On jump processes in the foreign exchange and stock markets, The Review of Financial Studies **1**(4): 427–445.

Kailath, T., Sayed, A. H. and Hassibi, B. (2000). Linear Estimation, Prentice Hall, New Jersey.

Kim, J. H., Shamsuddin, A. and Lim, K.-P. (2011). Stock return predictability and the adaptive markets hypothesis: Evidence from century-long U.S. data, Journal of Empirical Finance **18**(5): 868–879.

Koopman, S. J. and Hol Uspensky, E. (2002). The stochastic volatility in mean model: empirical evidence from international stock markets, Journal of applied Econometrics **17**(6): 667–689.

Kulikova, M. and Taylor, D. (2014). Estimating adaptive market efficiency using the kalman filter, ACQuFRR Working Paper Series, University of Cape Town .

Kulikova, M. V. (2016). Square-root algorithms for maximum correntropy estimation of linear discrete-time systems in presence of non-gaussian noise, arXiv preprint arXiv:1610.00257 .

Kulikova, M. V. (2017). Square-root algorithms for maximum correntropy estimation of linear discrete-time systems in presence of non-gaussian noise, Systems & Control Letters **108**: 8–15.

Kulikova, Taylor and Kulikov (2019). A nonlinear bayesian filtering approach to estimating adaptive market efficiency, Russian Journal of Numerical Analysis and Mathematical Modelling .

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**(1): 69–108.

Lim, K.-P., Luo, W. and Kim, J. H. (2013). Are us stock index returns predictable? evidence from automatic autocorrelation-based tests, Applied Economics **45**(8): 953–962.

Liu, X., Qu, H., Zhao, J. and Chen, B. (2016). Extended kalman filter under maximum correntropy criterion, 2016 International Joint Conference on Neural Networks (IJCNN), IEEE, pp. 1733–1737.

Lo, A. (2004). The adaptive markets hypothesis, The Journal of Portfolio Management **30**(5): 15–29.

Malkiel, B. G. (1989). Efficient market hypothesis, Finance, Springer, pp. 127–134.

Malkiel, B., Mullainathan, S. and Stangle, B. (2005). Market efficiency versus behavioral finance, Journal of Applied Corporate Finance **17**(3): 124–136.

Mandelbrot, B. (1972). Certain speculative prices (1963), The Journal of Business **45**(4): 542–543.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach, Econometrica: Journal of the Econometric Society pp. 347–370.

Pesaran, M. H. (1987). The limits to rational expectations, number 339.3 PESl.

Roberts, H. V. (1967). Statistical versus clinical prediction of the stock market.

Ruiz, E. (1994). Quasi-maximum likelihood estimation of stochastic volatility models, Journal of Econometrics **63**: 289–306.

Schweppe, F. C. (1965). Evaluation of likelihood functions for Gaussian signals, IEEE Trans. Inf. Theory **IT-11**: 61–70.

Simon, H. A. (1955). A behavioral model of rational choice, The quarterly journal of economics **69**(1): 99–118.

Ţiţan, A. G. (2015). The efficient market hypothesis: review of specialized literature and empirical research, Procedia Economics and Finance **32**: 442–449.

Urquhart, A. and McGroarty, F. (2014). Calendar effects, market conditions and the Adaptive Market Hypothesis: Evidence from long-run U.S. data, International Review of Financial Analysis **35**: 154–166.

Wang, H., Li, H., Zhang, W., Zuo, J. and Wang, H. (2018). Maximum correntropy derivative-free robust kalman filter and smoother, IEEE Access **6**: 70794–70807.

Ward, M. and Muller, C. (2012). Empirical testing of the capm on the jse, Investment Analysts Journal **41**(76): 1–12.

Wilcox, B. A. and Hamano, F. (2017). Kalman's expanding influence in the econometrics discipline, Proceedings of the 20th IFAC World Congress, Toulouse, France, pp. 644–651.

Yang, Y. and Huang, G. (2017). Map-based localization under adversarial attacks, Proc. of the International Symposium on Robotics Research, Puerto Varas, Chile.

Zalewska-Mitura, A. and Hall, S. (1999). Examining the first stages of market performance: A test for evolving market efficiency, Economics Letters **64**(1): 1–12.

Appendix A

Appendix

A.1 Residual plots

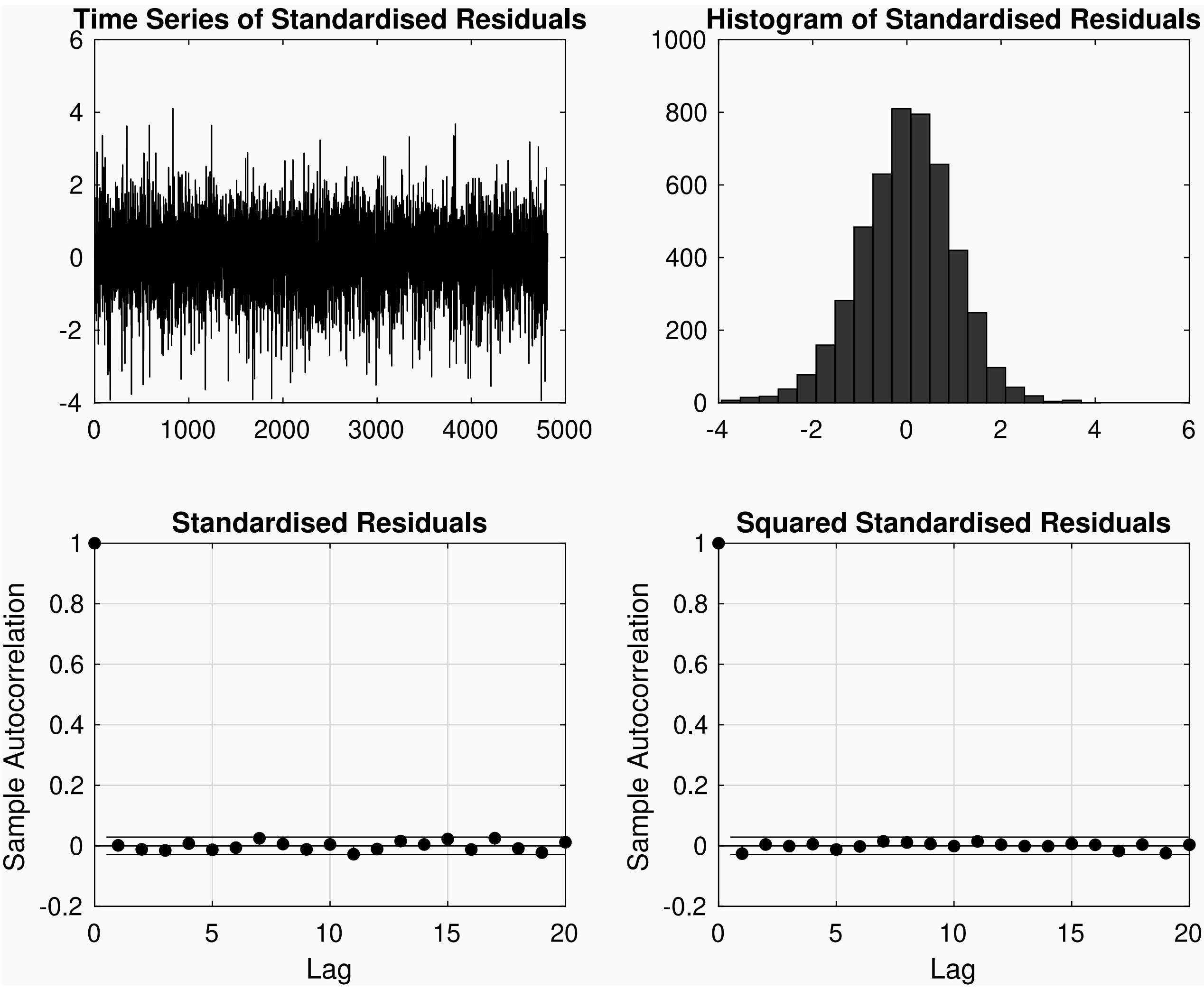

Fig. A.1: TEE residuals for the ALSI

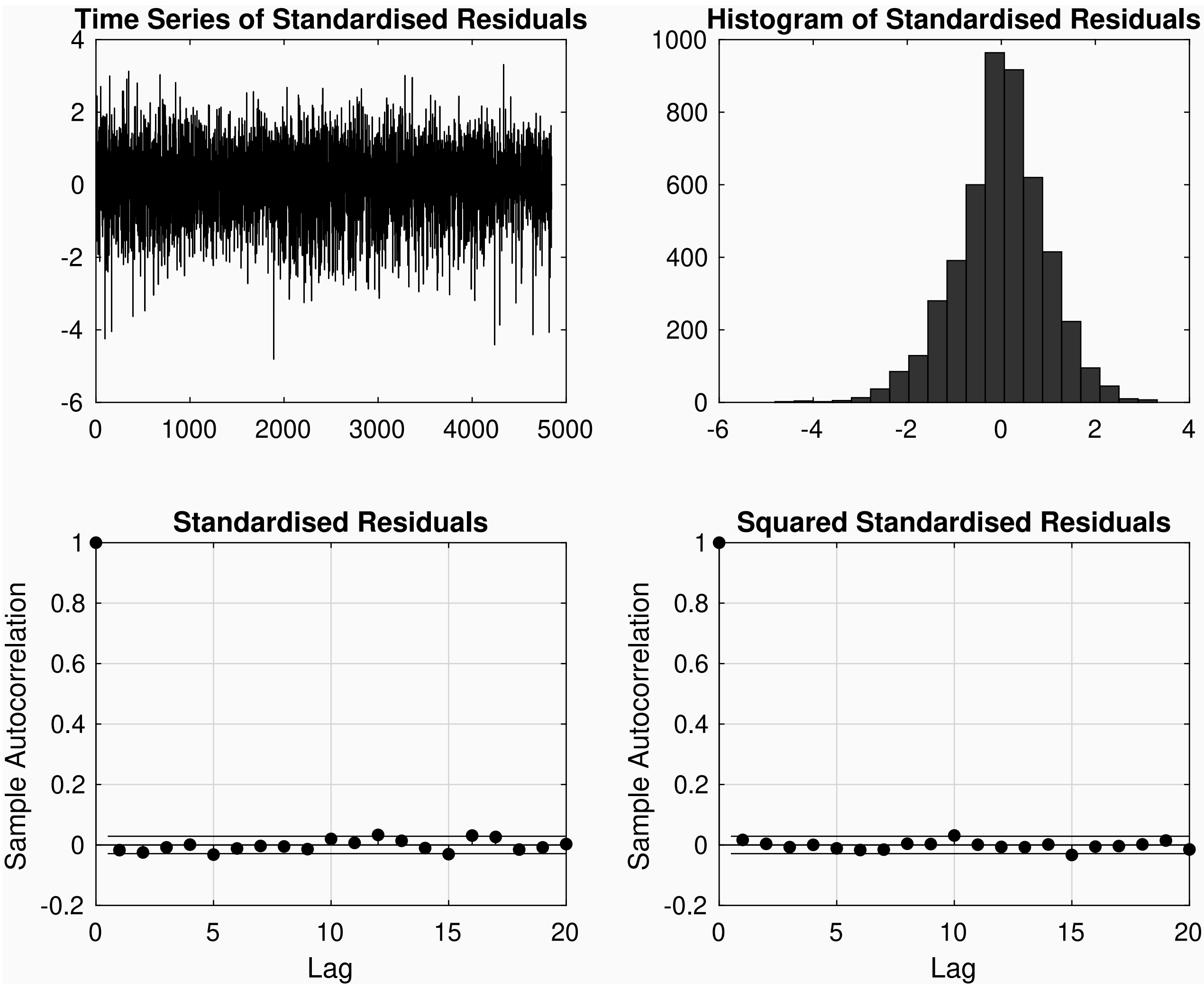

Fig. A.2: GTEE ($f_v = \sqrt{h_t}$) residuals for the S&P500

A.2 Tests in the residuals

Index	Statistic	Value	p-value
S&P500	$Q(12)$	13.59	0.257
	$Q^2(12)$	9.91	0.538
ALSI	$Q(12)$	10.71	0.468
	$Q^2(12)$	7.16	0.768

Tab. A.1: Results from Ljung-Box tests of dependency