Macquarie University

Department of Statistics

Log quantile differences and the temporal aggregation of alpha-stable moving average processes

Adrian Walter Barker

Submitted for fulfilment of the requirement for the degree of Doctor of Philosophy (Statistics)

July 2015

Supervisor: Dr. Nino Kordzakhia

Contents

List of Tables

List of Figures

Abstract

The modelling of the time-varying volatility of financial market asset log returns has attracted considerable interest from researchers and market participants. Prominent amongst these models are the generalized autoregressive conditional heteroskedastic (GARCH) models and the stochastic volatility (SV) models. Generally, such models use the conditional variance as the measure of dispersion. This thesis advocates for the use of the log quantile difference (LQD) as an alternative measure of dispersion where the variance of the intraday log returns does not exist.

Use of the LQD rather than the variance can present analytical and computational challenges. In this thesis we show that the impact can be mitigated by assuming that the log returns are from an alpha-stable moving average (SMA) process. The formulae derived for the LQD of the temporal aggregation of an SMA process allow the LQD shape to be examined as a function of aggregation level.

Asymptotically normal estimators are proposed for the LQD of the temporal aggregation of an SMA process, which require asymptotically normal estimators of the SMA process. The quantile-based stable distribution parameter estimators of McCulloch (1986) are adapted for use from an SMA process rather than an independent process. Traditionally such estimators have been calculated at the standard quantile levels originally proposed by McCulloch (1986). In this thesis, the quantile levels are identified which optimise estimators from a selection of SMA processes. We find that in many cases these optimised quantile-based estimators significantly outperform the quantile-based estimators using the standard quantile levels. Improved evaluations of the maximum likelihood estimator asymptotics are made to calculate the relative asymptotic efficiency of the optimal quantile-based estimators. Methods for order identification of an SMA process are developed and studied.

An extension to the SV model is proposed, which we call the stable stochastic volatility

(SSV) model, where the conditional distribution of the daily log returns is alpha-stable. Estimation of the SSV model parameters is done utilising the LQD estimators of the temporal aggregation of the intraday log return process together with an allowance for measurement error.

The methods proposed in this thesis are illustrated in an empirical study carried out on ASX200 index data from 2009 and 2010. A similar study was carried out by the author on the same data in Barker (2014).

Declaration

I hereby certify that this thesis has not been submitted for a higher degree at any other university or institution.

Signed:_____

Adrian Walter Barker

March, 2015

Acknowledgements

I wish to thank a number of people who have helped me in the writing of this thesis.

To my first supervisor, Associate Professor Andrzej Kozek, who encouraged me to explore the world of statistics, to ask questions and look for answers. To my second and final supervisor, Dr Nino Kordzakhia, who guided me through the process of creating a thesis from a collection of ideas.

As a part-time student for the duration of this thesis, one of the main challenges was to manage the competing demands for time from my employment and my research. To my employers, and especially to Stuart Wilson and Svetlana Rokhlin, who were always flexible and sympathetic in their demands for my time.

To my friends Xiufang (Leah) Li, Kin Wah (Mandy) Cheng, Marghanita da Cruz, Nora Sek-Zenian and others who, each in their own way, supported and encouraged me throughout.

And finally, to my late mother Jill, my father Gordon and my sister Wendy, without whom none of this would have been possible and to whom I dedicate this thesis.

Notation and Acronyms

α	Index of stability parameter of a stable distribution
β	Skewness parameter of a stable distribution
γ	Scale parameter of a stable distribution
δ	Location parameter of a stable distribution
$\zeta_p^{(r)}$	The log quantile difference at aggregation level r and quantile levels $p = (p_1, p_2)$
$\widetilde{\zeta}_p^{(r)}$	The parametric estimator of $\zeta_p^{(r)}$
$\widehat{\zeta}_p^{(r)}$	The non-parametric estimator of $\zeta_p^{(r)}$
$\ddot{\zeta}_p^{(r)}$	The parametric estimator of $\zeta_p^{(r)}$ using alternative quantile levels p_M for quantile-based estimation of stable distribution parameters
$\check{\zeta}_p^{(r)}$	The analytic estimator of $\zeta_p^{(r)}$
θ	Moving average parameter of the base process
$\varkappa$	Combined vector of the stable distribution parameter ω and the moving average process parameter θ
$\lambda(h)$	The autocovariance function
ξ_p	The pth quantile
$\rho(h)$	The autocorrelation function
ϕ_j	Autoregressive parameter of the volatility process
ψ_j	Moving average parameter of the volatility process
χ	Combined vector of the stable distribution parameter α the log quantile difference $\zeta_p^{(1)}$ and the moving average process parameter θ
ω	Vector of stable distribution parameters $(\alpha, \beta, \gamma, \delta)$

ACF	Autocorrelation function
AICC	Akaike information criterion with bias correction
ANOVA	Analysis of variance
AR	Autoregressive
ARCH	Autoregressive conditional heteroscedastic
ARMA	Autoregressive moving average
ARFIMA	Autoregressive fractionally integrated moving average
ARIMA	Autoregressive integrated moving average
ASX	Australian stock exchange
$B_{\widehat{\rho}_1}(n, a; \alpha, \beta)$	The a^{th} percentile of the small sample distribution of $\widehat{\rho}(1)$ from an IID sample of length n and distribution $S_\alpha(\beta, 1, 0)$
$B^{ASY}_{\widehat{\rho}_1}(n, a; \alpha, \beta)$	The a^{th} percentile of the asymptotic distribution of $\widehat{\rho}(1)$ from an IID sample of length n and distribution $S_\alpha(\beta, 1, 0)$
$B_{Q_s}(n, a; \alpha, \beta)$	The a^{th} percentile of the small sample distribution of the Q_s statistic from an IID sample of length n and distribution $S_\alpha(\beta, 1, 0)$
CML	Corrected maximum likelihood
CMM	Classical method of moments
d	Index of day
D	Number of trading days in a sample
$\{e_t\}$	Innovations of the base process
EGARCH	Exponential generalised autoregressive conditional heteroscedastic
GARCH	Generalised autoregressive conditional heteroscedastic
GMM	Generalised method of moments
HARCH	Heterogeneous autoregressive conditional heteroscedastic
IGARCH	Integrated generalised autoregressive conditional heteroscedastic
IID	Independent and identically distributed
IMA	Integrated moving average
k	Autoregressive order of the volatility process
LQD	Log quantile difference
MA	Moving average
ML	Maximum likelihood

m	Moving average order of the volatility process
$N\left(\mu, \sigma^2\right)$	Normal distribution with mean μ and variance σ^2
p	Quantile level
p_M	Quantile levels used in the quantile-based estimation of stable distribution parameters
q	Moving average order of the base process
QML	Quasi maximum likelihood
Q_s	Box-Pierce Q-statistic calculated over s lags of the autocorrelation function
r	Aggregation level
RML	Restricted maximum likelihood
$S_\alpha^0\left(\beta, \gamma, \delta\right)$	Stable distribution parameterisation, Nolan (1998)
$\left\{S_t^{(r)}\right\}$	Aggregated process
SLAD	Self weighted least absolute deviation estimation
SMA	Stable moving average
SR-SARV	Square-root stochastic autoregressive volatility
SSV	Stable stochastic volatility
SV	Stochastic volatility
t	Index of time
$\{u_t\}$	Innovations of the true realised volatility process
$U\left(a, b\right)$	Uniform distribution over the domain $[a, b]$
$\{v_t\}$	Innovations of the estimated realised volatility process
V_ζ	The asymptotic covariance matrix of the empirical log quantile difference estimator
V_ξ	The asymptotic covariance matrix of the empirical quantile estimator
V_θ	The asymptotic covariance matrix of the SLAD estimator
V_ω	The asymptotic covariance matrix of the stable distribution parameter estimator

VaR	Value at Risk
VARIMA	Vector autoregressive integrated moving average
VARMA	Vector autoregressive moving average
$\{w_t\}$	Measurement error process
$\{X_t\}$	Base process

Chapter 1

Introduction

1.1 Background

The modelling of financial market data has been the subject of a large amount of research over many decades. The attraction of such research is easy to understand, given the volume and value of daily financial market transactions modelling and the potential benefit should one be able to create an advantage over other market participants through a superior model. A frequently observed characteristic of financial market data is that of volatility clustering, (Pagan (1996), Cont (2001) and Zumbach (2011)). Volatility clustering occurs when high or low volatility trading days cluster together. This can be described quantitatively as the presence of a high positive autocorrelation in the volatility time series. Volatility clustering lies in stark contrast to the general absence of autocorrelation in the actual asset returns, (Cont (2001)). The presence of volatility clustering provides opportunities for using a range of models in its modelling and forecasting.

A simple example of the use of volatility forecasts can be found in options trading. The buyer of a call option pays a premium for the right but not obligation to buy an asset at a fixed price at some pre-determined time in the future regardless of the asset price at that time. A trader who judges that the price of an asset is likely to increase enough to cover the cost of the premium, may consider buying a call option to be profitable. Trading in options, rather than the asset itself, can increase the trader's leverage in the market. Similarly, the buyer of a put option pays a premium for the right but not obligation to sell an asset at a fixed price at some time in the future. Buying a put and call option of the same asset

creates a position called the long straddle. The long straddle is profitable if the asset price moves up or down by a large amount, i.e. has high volatility and is unprofitable if the asset price remains steady, i.e. has low volatility. Thus an ability to forecast volatility may assist a trader to calculate the value of a long straddle position and determine whether the current market price is under-valued or over-valued.

A popular type of stochastic volatility model are variations of the autoregressive conditional heteroscedastic (ARCH) and generalised autoregressive conditional heteroscedastic (GARCH) models. The ARCH model was first proposed by Engle (1982) and extended to become GARCH models by Bollerslev (1986). Let P_t denote the price of an asset at time t and define the log return of the asset by

$$X_t = \ln\left(P_t/P_{t-1}\right). \tag{1.1.1}$$

The conditional variance of the log return, σ_t^2, is given by

$$\sigma_t^2 = var\left[X_t \mid \mathcal{F}_{t-1}\right] \tag{1.1.2}$$

where $\mathcal{F}_{t-1}$ is the collection of all information available at time $t-1$.

Definition 1.1.1 ***(Generalised Autoregressive Conditional Heteroscedastic Model).*** *A* GARCH(p,q) *model is defined by the following equations*

$$X_t = \sigma_t \varepsilon_t \tag{1.1.3}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{1.1.4}$$

where $\{\varepsilon_t\}$ *is a sequence of independent, identically distributed (*IID*)* $N(0,1)$ *random variables and* $\alpha_0 > 0, \alpha_i, \beta_j \geq 0$ *are constants which serve as the model parameters.*

That GARCH models exhibit volatility clustering is clear from (1.1.4). Large values for X_{t-i}^2 and σ_{t-j}^2 will generate a large value for σ_t^2 and make it likely that X_t will also have a large absolute value. Conversely small values for X_{t-i}^2 and σ_{t-j}^2 make it likely that X_t will also have a small absolute value. The GARCH models also exhibit the heavy tails commonly seen in financial markets and there is no autocorrelation in the log returns $\{X_t\}$. Estimation of the GARCH model parameters can be achieved through maximisation of the log likelihood function, (e.g. Xekalaki and Degiannakis (2010)).

Despite having the attractive properties discussed above, various limitations of the GARCH model have led to a plethora of GARCH model variants. The exponential GARCH (or EGARCH) model (Nelson (1991)) allows asymmetry in the formula for the conditional variance σ_t^2, which otherwise depends only on the absolute value of X_{t-i} and σ_{t-j}, not the sign. The integrated GARCH (or IGARCH) model (Engle and Bollerslev (1986)) allows for shocks to the conditional variance σ_t^2 to persist indefinitely rather than decay asymptotically. The relationship between the GARCH and IGARCH models is analogous to that between the autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) models. The heterogeneous ARCH (or HARCH) model (Müller et al. (1997)) allows the conditional variance σ_t^2 to be affected differently by past conditional variances evaluated over different time intervals. Descriptions of these and other GARCH model variants can be found in Xekalaki and Degiannakis (2010).

An alternative to the GARCH models is the stochastic volatility (SV) model, sometimes referred to as autoregressive variance (ARV) model. These models were first proposed in Taylor (1982) and Taylor (1986) (Section 3.5).

Definition 1.1.2 ***(Stochastic Volatility Model).*** *The* SV *model is defined by the following equations*

$$X_t = \sigma_t \varepsilon_t \tag{1.1.5}$$

$$\ln \sigma_t = \mu + \sum_{j=1}^{k} \phi_j \ln \sigma_{t-j} + v_t, \tag{1.1.6}$$

where $\{\varepsilon_t\}$ *is a sequence of* IID $N(0,1)$ *random variables,* $\{v_t\}$ *are a sequence of* IID $N\left(0, \sigma_v^2\right)$ *random variables, and* $\{\varepsilon_t\}, \{v_t\}$ *are assumed to be independent of each other.*

If the parameters ϕ_j are chosen so that (1.1.6) has a positive autocorrelation function at low lag values, then volatility clustering is a property of the SV models. The kurtosis of X_t was shown (Taylor (1986)) to be

$$\text{kurtosis}\left[X_t\right] = 3\exp\left(4\sigma_v^2\right). \tag{1.1.7}$$

Thus the distribution of X_t always has a heavier tail than the normal distribution however, unlike GARCH models, the fourth moment of X_t always exists. Popular estimation methods for SV models include quasi maximum likelihood (QML) (Harvey et al. (1994) see also

Anderson and Sorensen (1997)) and generalised method of moments (GMM) (Melino and Turnbull (1990) and Anderson and Sorensen (1996)).

The GARCH and SV models described above include only the daily log returns as observational data. The availability of high-frequency intraday prices offers the potential to develop alternate models for the volatility of financial market asset returns. High-frequency finance is the analysis of financial market data gathered over short time intervals, often trade by trade, either for its own purposes or for possible inferences about the behaviour of financial markets over longer time intervals, (Dacorogna et al. (2001), Tsay (2010) and Hautsch (2012)).

An important concept in high-frequency finance is that of temporal aggregation. The temporal aggregation of the stochastic process $\{X_t\}$ is generally defined as the weighted sum of past and current process values, (Silvestrini and Veredas (2008)). The importance of temporal aggregation is in its role in examining the relationship between models indexed on time intervals of different lengths. Such relationships may allow inference on high frequency data from models derived from low frequency data and vice versa. In this thesis, we consider only a special case of temporal aggregation, sometimes referred to as flow aggregation, where all the weights equal one.

Definition 1.1.3 ***(Temporal Aggregation).*** *The temporal aggregation of the stochastic process $\{X_t\}$ at aggregation level r, is defined by*

$$S_t^{(r)} = \sum_{i=0}^{r-1} X_{t-i}. \tag{1.1.8}$$

We may refer to $\{X_t\}$ as the base process and $\left\{S_t^{(r)}\right\}$ as the aggregated process. A non-overlapping aggregated process is one where each element of the base process contributes to only one element of the aggregated process

$$S_\tau^{(r)} = \sum_{i=0}^{r-1} X_{r\tau-i}. \tag{1.1.9}$$

Important early work on the temporal aggregation of AR processes was done by Amemiya and Wu (1972). They showed that the temporal aggregation of a non-overlapping AR(p) process is an ARMA(p, q) process where $p \leq q$. In addition, they compared the performance of predictors for the aggregated process derived from aggregated process data with

those derived from base process data. This work was continued by Tiao (1972) where it was shown that the predictors were asymptotically equivalent for large aggregation levels unless the base process is non-stationary. The temporal aggregation of non-overlapping ARMA(p, q) processes was shown by Brewer (1973) to converge at large aggregation levels to an ARMA$(p, p+1)$ process if $q > p$ or an ARMA(p, p) process if $q \leq p$. Temporal aggregation of ARIMA processes was considered by Weiss (1984) and Stram and Wei (1986), of vector ARMA (or VARMA) processes by Lütkepohl (1986), of vector ARIMA (or VARIMA) processes by Marcellino (1999) and of autoregressive fractionally integrated moving average (ARFIMA) processes by Tsai and Chan (2005).

Research on temporal aggregation has also been extended to include volatility processes. In Drost and Nijman (1993), three classes of the GARCH model are defined: strong, semi-strong and weak. These classes differ in the requirements for independence of the innovation process $\{\varepsilon_t\}$ in (1.1.3) . Strong GARCH models have an IID innovation process $\{\varepsilon_t\}$. All weak GARCH models are semi-strong GARCH models, all semi-strong GARCH models are strong GARCH models. It was shown that symmetric weak GARCH models, but not strong nor semi-strong GARCH models, are closed under temporal aggregation. In Meddahi and Renault (2005), the square-root stochastic autoregressive volatility (SR-SARV) class of volatility models was defined as an extension of the class of weak GARCH models. The class of SR-SARV models includes SV models. They show that non-overlapping SR-SARV models are closed under temporal aggregation.

After the logarithmic transformations, the base and aggregated lognormal AR processes are related by a type of "non-linear temporal aggregation", (Salazar and Fereira (2011)). The logarithm of a lognormal AR process is a Gaussian AR process. However, the logarithm of the temporal aggregation of a lognormal AR process is not a Gaussian process, though empirically it may appear very similar. A moment-matching method is proposed by Salazar and Fereira (2011), for approximating the aggregated process by a lognormal AR process.

A unified framework is proposed by Sbrana and Silvestrini (2013) for modelling both the temporal aggregation and contemporaneous aggregation of financial asset returns based on integrated moving average (IMA) processes. Contemporaneous aggregation is the weighted sum of different stochastic processes at the same points in time.

Where the base process $\{X_t\}$ is the log return of a financial asset, the aggregated process

$\left\{S_t^{(r)}\right\}$ is particularly useful, since the sum of adjacent log returns is equal to the log return over the union of the intervals

$$S_t^{(r)} = \sum_{i=0}^{r-1} X_{t-i} = \sum_{i=0}^{r-1} \ln\left(P_{t-i}/P_{t-i-1}\right) = \ln\left(P_t/P_{t-r}\right). \tag{1.1.10}$$

Therefore if we can derive properties of the aggregated process from the properties of the base process, then we may be able to use the aggregated process properties to provide estimates for the low frequency volatility.

One method for the estimation of low frequency volatility from high-frequency returns is through the use of realised volatility, $R_t^{(r)}$, given by

$$R_t^{(r)} = \sum_{i=0}^{r-1} X_{t-i}^2. \tag{1.1.11}$$

If the base process $\{X_t\}$ is an IID process with zero mean and finite variance, then the expected value of $R_t^{(r)}$ is equal to the variance of the aggregated process $\left\{S_t^{(r)}\right\}$. Cross-product terms may be added to (1.1.11), if the assumption that $\{X_t\}$ is an IID process is relaxed to allow $\{X_t\}$ to be an MA process. Realised volatility was used in French et al. (1987) to estimate the volatility of monthly returns from daily returns and by Anderson et al. (2001a) and Anderson et al. (2001b) to estimate the volatility of daily returns from intraday returns. However, the assumption that the distribution of the high frequency returns has a finite variance is still a contentious topic in financial markets research.

The finite variance hypothesis for financial market returns dates back to the Brownian motion model of Bachelier (1900) and later by Osborne (1959) and others. This model implies a Gaussian distribution for financial market log returns. However, the empirical evidence analysed in Mandelbrot (1963) and Fama (1965) suggested that the distribution of financial market log returns had heavier tails than the Gaussian and that a non-Gaussian stable distribution with infinite variance would be more appropriate, (see Appendix E for the definition and some properties of stable distributions). In subsequent years, many empirical studies have been conducted and other distributions have been proposed, (e.g. Teichmoeller (1971), Fielitz and Smith (1972), Blattberg and Gonedes (1974), Fielitz and Rozelle (1983), Barndorff-Nielsen (1997) and Grabchak and Samorodnitsky (2010)).

Whilst it is generally accepted that the Gaussian distribution is too light-tailed for most financial market returns, whether or not non-Gaussian stable distributions are too heavy-

tailed remains unclear. Measurement of the heaviness of the tail of a distribution is often performed using the tail index.

Definition 1.1.4 ***(Tail Index)*** *Let $\mathcal{R}_u$ denote the class of regularly varying functions at infinity with index u, that is the class of positive functions g such that*

$$\lim_{t\to\infty} \frac{g(tx)}{g(t)} = x^u \tag{1.1.12}$$

for all $x > 0$. The tail index v of a distribution function F is the value such that $1-F \in \mathcal{R}_{-v}$.

A simple example is the Pareto distribution where

$$F(x) = 1 - x^{-v}, \quad x \geq 1 \tag{1.1.13}$$

which has a tail index equal to v. For such distributions, all moments greater than or equal to the tail index are infinite. The tail index of a non-Gaussian stable distribution equals the stability parameter and of a t distribution equals the degrees of freedom parameter. See Rachev and Mittnik (2000) - Section 3.6, for a discussion on tail index estimation.

Several studies have reported that the tail index of financial market returns appears to increase with aggregation level (e.g. Akgiray and Booth (1988) and Dacorogna et al. (2001) - Section 5.4). If the base process $\{X_t\}$ is an IID stable process, then the tail index should be constant across all aggregation levels. To quote E. Schwartz in the foreward of Rachev and Mittnik (2000),

> "This observation actually implies that returns are not independent, and / or not identically distributed, and / or not stable, but it does not automatically rule out that stable processes may underlie generating models."

The models considered in this thesis do not allow for a change in the tail index with aggregation level. The empirical study included in this thesis is not conclusive on the assertion that the tail index does not change with aggregation level, but does provide some evidence to suggest that the intraday log returns are not identically distributed.

If the conditional distribution of financial market returns does not have a finite variance, then the expected value of the realised volatility, as defined in (1.1.11), does not exist. In this thesis, we consider an alternative definition of realised volatility based on quantiles, the expected value of which can exist where the conditional distribution of financial market returns does not have a finite variance.

1.2 Estimation of realised volatility from infinite variance intraday log returns

For infinite variance distributions, an alternative to the variance as a measure of dispersion is the log quantile difference (LQD).

Definition 1.2.1 ***(Log Quantile Difference)***. *Let* $\{X_t\}$ *be a stationary process. Let* ξ_{p_j} *denote the* p_j^{th} *quantile of* $\{X_t\}$. *Define the log quantile difference of* $\{X_t\}$ *at quantile levels* $p = (p_1, p_2)$ *to be*

$$\zeta_p = \ln\left(\xi_{p_2} - \xi_{p_1}\right) \tag{1.2.1}$$

where $0 < p_1 < p_2 < 1$. *We take the logarithm in* (1.2.1), *in order that this measure of dispersion takes values over the entire real line.*

The LQD has a number of advantages over the variance as a measure of dispersion. The LQD is finite for all stable distributions, whereas only the Gaussian stable distribution has a finite variance. The log quantile difference provides multiple measures of volatility for each distribution, through the choice of p, whereas the variance provides only a single measure. The LQD is a more robust measure than the variance, (see for example Wilcox (2005)).

Use of the LQD as a measure of volatility is closely related to the use of value at risk (VaR) as a measure of market risk. Market risk is the risk to a financial asset due to an unexpected movement in the asset's price. The VaR is defined as the maximum loss suffered by a financial asset over a given time period which is not exceeded with a given probability. Thus, forecasts from models of the LQD of a financial asset can be used to estimate the VaR of the asset. See Chapter 7 of Tsay (2010) and included references for more information on VaR.

The main advantage of the variance over the LQD is that the calculations are much simpler. More specifically, for any two independent random variables with finite variance, the sum of the variances is equal to the variance of the sum. The same result is not true in general for LQDs. Inequality relations between the quantiles of some IID random variables and the quantiles of their sum are provided in Watson and Gordon (1986), (see also Liu and David (1989)). The author is not aware of any general equality relations between the quantiles of IID random variables and the quantiles of their sum.

Many of the calculation difficulties related to the use of LQDs can be overcome if we assume that the intraday log returns follow a stable moving average (SMA) process. The use of SMA processes also allows for more realistic modelling where intraday log returns are not IID.

Definition 1.2.2 ***(Stable Moving Average Process).*** *The process* $\{X_t\}$ *is called a stable moving average process of order* q, SMA(q), *if it follows the model equation*

$$X_t = \sum_{j=0}^{q} \theta_j e_{t-j} \tag{1.2.2}$$

where $\theta_0 = 1$ *and* $\{e_t\}$ *is an* IID *sequence of stable random variables such that*

$$e_t \sim S_\alpha^0 \left(\beta_e, \gamma_e, \delta_e\right) \tag{1.2.3}$$

using the S^0 *parameterisation of stable distributions proposed in Nolan (1998). See* $(E.1.3)$ *for the characteristic function of the* S^0 *parameterisation.*

The stability parameter α determines the heaviness of the tails of the distribution, with special cases at $\alpha = 1$ (Cauchy) and $\alpha = 2$ (Gaussian). The parameters β_e, γ_e and δ_e respectively determine the skewness, scale and location of the distribution. A stable random variable has a finite variance only if $\alpha = 2$ and has a finite mean only if $\alpha \geq 1$.

One of the useful properties of stable distributions is that linear combinations of stable random variables with stability parameter α are also stable random variables with stability parameter α. Thus

$$X_t \sim S_\alpha^0 \left(\beta^{(1)}, \gamma^{(1)}, \delta^{(1)}\right) \tag{1.2.4}$$

and

$$S_t^{(r)} \sim S_\alpha^0 \left(\beta^{(r)}, \gamma^{(r)}, \delta^{(r)}\right) \tag{1.2.5}$$

are stable random variables and have the same stability parameter α as e_t. Note that the temporal aggregation $\left\{S_t^{(r)}\right\}$ of the SMA process $\{X_t\}$ is also a SMA process with the same innovation sequence $\{e_t\}$. Formulae for the stable distribution parameters of X_t and $S_t^{(r)}$ are provided in Chapter 2.

These properties of stable distributions allow analytic formula to be derived for the LQD of $S_t^{(r)}$ in terms of the parameters of some SMA processes $\{X_t\}$. Let $\zeta_p^{(r)}$ denote the LQD

of the temporal aggregation $\left\{S_t^{(r)}\right\}$ of the SMA(q) process $\{X_t\}$. In the simple case where $q = 0$, i.e. $\{X_t\}$ is an IID process, it can be shown that

$$\zeta_p^{(r)} = \alpha^{-1} \ln r + \zeta_p^{(1)}. \tag{1.2.6}$$

More general formulae for $\zeta_p^{(r)}$ where $q \neq 0$ are provided in Chapter 2.

Under the assumption that the intraday log returns are from an SMA process, we can use the formulae in Chapter 2 to calculate an estimate of $\zeta_p^{(r)}$ provided we have estimates for the SMA process parameters. Asymptotically normal estimators for θ_j, $j = 0, \ldots, q$ can be calculated using the self-weighted least absolute deviation (SLAD) method described in Zhu and Ling (2012). However, the estimation of stable distribution parameters from an SMA process requires a new method developed in this thesis.

A number of methods have been proposed for the estimation of the parameters of a stable distribution. Quantile-based methods were proposed by Fama and Roll (1971) and McCulloch (1986). Methods based on the sample characteristic function were proposed by Press (1972), Paulson et al. (1975) and Kogon and Williams (1998). Maximum likelihood estimation methods were investigated in Brorsen and Yang (1990) and Nolan (2001). The methods listed above all assume that estimation is from an independent sample. In this thesis, we show how to adapt the quantile-based stable distribution parameter estimators of McCulloch (1986) to construct asymptotically normal estimators of the stable distribution parameters from an SMA process.

Quantile-based stable distribution parameter estimators have traditionally used quantile estimators at the standard quantile levels proposed by McCulloch (1986). However, no reason was provided by McCulloch to explain why those quantile levels should be used in preference to any of an infinite selection of possible alternatives. In this thesis, we attempt to identify the optimal choice of quantile levels for use in quantile-based stable distribution parameter estimation from a range of SMA processes. We find that estimators for α and $\beta^{(1)}$ generated using the standard quantile levels are reasonably efficient from a range of different SMA processes, but are far from optimal in many cases. For example, estimators for α generated using the standard quantile levels from a symmetric IID process are quite poor where α lies outside the range $[0.8, 1.6]$. In addition, we find that the optimal choice of quantile levels for estimating α do not necessarily produce good estimates of $\beta^{(1)}$ and vice versa.

The relative asymptotic efficiency was calculated of the optimal quantile-based stable distribution parameter estimators compared with the maximum likelihood estimators. No analytic formula is available for the asymptotic variance of the maximum likelihood estimators, but instead can be calculated through numerical approximations of the Fisher information matrix integral. Evaluations of the asymptotic variance of the maximum likelihood estimators have been published in DuMouchel (1975), Nolan (2001) and Matsui and Takemura (2006). However, these values are not always consistent and in some cases would mean that the optimal quantile-based estimator is more efficient than the maximum likelihood estimator. Evaluations in Nolan (2001) were found to have insufficiently accounted for the contribution of the tail of the Fisher information matrix integral. We propose a new method for evaluating the tails of the Fisher information matrix integral using the asymptotic properties of the stable distribution and use this method to evaluate a new set of approximations for the asymptotic variance of the maximum likelihood estimators. Using these evaluations, we find that the relative asymptotic efficiency of the optimal quantile-based estimator compared to the maximum likelihood estimator is between 70% and 85% for many values of α and $\beta^{(1)}$.

Standard methods for order identification and goodness of fit testing available in finite variance time series analysis are not necessarily appropriate for SMA processes. However, some methods involving the sample autocorrelation function can be adapted for SMA processes since, somewhat surprisingly, the autocorrelation function is well-defined for an SMA process even though the autocovariance function may be infinite. The asymptotic distribution of the sample autocorrelation function was derived in Davis and Resnick (1986). Unfortunately convergence to the asymptotic distribution for the sample autocorrelation function is very slow, hence the asymptotic distribution does not provide accurate significance levels for any autocorrelation function based statistical tests. For order identification of SMA processes, it was suggested to use significance levels calculated from the sample ACF asymptotic distribution of the Cauchy distribution for all SMA processes, (Adler et al. (1998)). Whilst this suggestion performs well for SMA(1) processes where the value of θ_1 is close to one, it performs poorly for SMA(1) processes where the value of θ_1 is close to zero. In this thesis, we propose the use of significance levels calculated from simulations of the same size as the sample. We show that these significance levels produce similar results to

those in Adler et al. (1998) for SMA(1) processes where the value of θ_1 is close to one and much better results for SMA(1) processes where the value of θ_1 is close to zero. A similar approach is adopted for a Q-statistic test of the independence of SMA process residuals.

That intraday log returns follow an SMA process is inconsistent with the SV model for daily returns which assume a conditionally normal distribution. To address this, we propose an extension to the SV model, which we call the stable stochastic volatility (SSV) model.

Definition 1.2.3 ***(Stable Stochastic Volatility Model)***. *The* SSV *model is defined by the following equations*

$$X_d \sim \Gamma_d \varepsilon_d \tag{1.2.7}$$

$$\ln \Gamma_d = \mu + \sum_{j=1}^{k} \phi_j \ln \Gamma_{d-j} + v_d + \sum_{j=1}^{m} \psi_j v_{d-j} \tag{1.2.8}$$

where $\{\varepsilon_d\}$ *is a sequence of* IID $S_\alpha^0(0,1,0)$ *random variables,* $\{v_d\}$ *are a sequence of* IID $N\left(0,\sigma_v^2\right)$ *random variables. It is assumed that* $\{\varepsilon_d\}$ *and* $\{v_d\}$ *are independent sequences.*

Equations (1.2.7) and (1.2.8) in the SSV model are analogous to equations (1.1.5) and (1.1.6) of the SV model. In the SSV model, we allow $\{\ln \Gamma_d\}$ to be an ARMA(k,m) process, rather than the pure autoregressive process included in the SV model. Note that if $\alpha = 2$ and $m = 0$, then the SV and SSV models are identical. Recent work on SV models with non-Gaussian conditional distributions can be found in Gander and Stephens (2007) and its references. For more information on alpha-stable risk modelling in general, see Peters and Shevchenko (2015).

Let $\{X_{d;t}\}$ denote the intraday log returns indexed over day d and time t. We assume that $\{X_{d;t}\}$ follows the SMA process

$$X_{d;t} = e_{d;t} + \theta_1 e_{d;t-1} \tag{1.2.9}$$

where for each day d, the innovarions $\{e_{d;t}\}$ is an IID sequence of random variables

$$e_{d;t} \sim S_\alpha^0(0, \gamma_d, 0). \tag{1.2.10}$$

All parameters of the intraday log return processes are assumed constant except for the innovations scale parameter, γ_d, which is stochastic from day to day. If the intraday log returns are an SMA process, then the daily log return, being the temporal aggregation of

the intraday log returns, also has a stable distribution with Γ_d as the scale parameter, (see Section 2.2). Thus the scale parameter Γ_d is a function of the parameters of the day's intraday log return process parameters and so constant within that day, but stochastic from day to day, and so can be estimated each day from that day's intraday log return data. These estimates for Γ_d can then be used as data for the estimation of the SSV model parameters in (1.2.8) . Note that the LQD of the temporal aggregation of the intraday log return process and $\ln \Gamma_d$ differ only by a constant. Note also that the GMM and QML methods for estimation of the SV model do not translate directly to the SSV model.

1.3 Thesis outline

In Chapter 2, formulae are provided for the stable distribution parameters of the temporal aggregation of an SMA process. Given the parameters of a stable distribution, the log quantile differences can always be calculated numerically however, for some types of SMA processes we have derived formulae for the log quantile differences of the aggregated process. For these SMA processes, the shape of the log quantile difference as a function of aggregation level is shown to be independent of quantile level. In Section 2.3, the classes of invertible SMA(1) and SMA(2) processes are examined in more detail.

In Chapter 3, we derive a method for estimating the parameters of an SMA process. In Sections 3.2 and 3.3, we adapt the quantile-based estimators of McCulloch (1986), to allow for samples from an SMA process and show how to calculate the asymptotic distribution of these estimators. A new method for the evaluation of the asymptotic distribution of maximum likelihood stable distribution parameter estimators is proposed in Section 3.4. These asymptotic variances are used to calculate the relative asymptotic efficiency of the optimal quantile-based estimators identified in Section 3.5. Simulation results for the quantile-based estimation of stable distribution parameters from an SMA process are presented in Section 3.9.

In Section 3.6, background information is provided on the asymptotic distribution of the sample ACF from an SMA process. Simulations are used to demonstrate the slow convergence of the sample ACF to its asymptotic distribution as well as to provide estimates of quantiles of the sample ACF distribution at realistic sample sizes. In Section 3.7, simulations are used to estimate the significance levels of small sample Q-statistic tests for residual independence.

In Section 3.8, simulations are used to estimate the significance levels of small sample order indentification tests of SMA processes. Comparisons between the misidentification percentages obtained using our simulated significance levels and the asymptotic Cauchy significance levels used in Adler et al. (1998) are included in Tables 3.8.1 and 3.8.2.

In Chapter 4, we use the formulae provided in Chapter 2 and the estimators defined in Chapter 3 to define the parametric estimator for the LQD of the temporal aggregation of an SMA process. Estimators for the stable distribution parameters of the temporal aggregation of an SMA process are defined in Section 4.2. Based on the estimators defined in Section 4.2, estimators for the LQD of the temporal aggregation of an SMA process, referred to as parametric LQD estimators, are defined in Section 4.3. Formulae for the asymptotic distribution of the parametric LQD estimators are derived in Section 4.3. The asymptotic variance of the parametric LQD estimators is dependent on the asymptotic covariance between the stable distribution parameter estimators and the moving average parameter SLAD estimators. Although no formulae are available for this asymptotic covariance, simulations in Section 4.4 are used to show that the asymptotic covariance is fairly small and has little effect on the asymptotic variance of parametric LQD estimators. In Section 4.5, an estimator is defined for the LQD of an SMA process, referred to as the non-parametric LQD estimator, based on the empirical quantile estimates. In Section 4.6, we demonstrate how the choice of quantile levels for the quantile-based stable distribution parameter estimators can effect the relative asymptotic efficiency of the parametric and non-parametric LQD estimators. In Section 4.7, an estimator is defined for the LQD of the temporal aggregation of some types of SMA processes, referred to as the analytic LQD estimator. The analytic LQD estimator is based on formulae derived in Chapter 2. In Section 4.8, we derive formulae for the joint asymptotic distribution of the stability parameter estimator and the non-parametric LQD estimator. Simulation results are presented in Section 4.9.

In Chapter 5, we examine in more detail, the estimation of realised volatility under the assumptions of the SSV model. In Section 5.2. we adapt the more general LQD estimators defined in Chapter 4 to define the mean symmetric parametric LQD estimator which satisfies the assumptions of the SSV model. In Section 5.3, we define the corrected maximum likelihood (CML) method for estimating the parameters of the true, unobservable, LQD process $\left\{\zeta_{d;p}^{(r)}\right\}$ from the estimated LQD process $\left\{\overline{\zeta}_{d;p}^{(r)}\right\}$ and the measurement error variance

estimates. Simulation results are presented in Section 5.4.

In Chapter 6 we apply the methods of this thesis to the ASX200 index of the Australian Stock Exchange (ASX) in 2009 and 2010. A study of the same data was done in Barker (2014), using slightly different methods and without any correction for measurement error. In Section 6.2, we describe the data cleaning tasks performed to deal with missing data, frozen data and the discontinuities in sample return distributions near zero. The base process data used for this analysis was the log returns over thirty second intervals. Estimation of the realised volatility and ssv model parameters are reported in Section 6.3. The results of some diagnostic tests are reported in Section 6.4. Proofs and additional background information are given in the appendicies.

Chapter 2

Log quantile difference of the temporal aggregation of stable moving average processes

2.1 Introduction

Let $\{X_t\}$ be the stable moving average process (Definition 1.2.2) of order q, SMA(q),

$$X_t = \sum_{j=0}^{q} \theta_j e_{t-j} \tag{2.1.1}$$

where $\theta_0 = 1$ and $\{e_t\}$ is an IID sequence of stable random variables such that

$$e_t \sim S^0_\alpha(\beta_e, \gamma_e, \delta_e) \tag{2.1.2}$$

using the S^0 parameterisation of stable random variables in Nolan (1998). Let θ denote the $q+1$ dimensional vector of moving average parameters

$$\theta = (\theta_0, \ldots, \theta_q)'. \tag{2.1.3}$$

Let $\left\{S^{(r)}_t\right\}$ denote the temporal aggregation (Definition 1.1.3) of $\{X_t\}$ at aggregation level r

$$S^{(r)}_t = \sum_{i=0}^{r-1} X_{t-i}. \tag{2.1.4}$$

Note that

$$S^{(1)}_t = X_t. \tag{2.1.5}$$

Let $\xi_{p_j}^{(r)}$ denote the $p_j th$ quantile of $\left\{S_t^{(r)}\right\}$ and $\zeta_p^{(r)}$ the log quantile difference (Definition 1.2.1) of $\left\{S_t^{(r)}\right\}$ at quantile levels $p = (p_1, p_2)'$

$$\zeta_p^{(r)} = \ln\left(\left|\xi_{p_2}^{(r)} - \xi_{p_1}^{(r)}\right|\right). \tag{2.1.6}$$

Similarly, let ξ_{p_j} denote the $p_j th$ quantile of $\{e_t\}$ and ζ_p the log quantile difference of $\{e_t\}$ at quantile levels $p = (p_1, p_2)'$.

In this chapter formulae are derived for the stable distribution parameters of $\left\{S_t^{(r)}\right\}$. An analytic formula is derived for $\zeta_p^{(r)}$ which is applicable to some SMA(q) processes. For these SMA(q) processes, the shape of the log quantile difference as a function of the aggregation level is examined and shown to be dependent on the parameters of the moving average process but not the quantile levels. Invertible SMA(1) and SMA(2) processes are examined in more detail.

2.2 Calculation of the log quantile difference

In this section, we show, under certain conditions on the stable moving average process $\{X_t\}$, that

$$\zeta_p^{(r)} = \alpha^{-1} \ln\left(r\left|\sum_{i=0}^{q}\theta_i\right|^{\alpha} + g_\alpha(\theta)\right) + \zeta_p \tag{2.2.1}$$

where

$$g_\alpha(\theta) = \left(\sum_{i=0}^{q-1}\left|\sum_{j=0}^{i}\theta_j\right|^{\alpha} - q\left|\sum_{i=0}^{q}\theta_i\right|^{\alpha} + \sum_{i=1}^{q}\left|\sum_{j=i}^{q}\theta_j\right|^{\alpha}\right). \tag{2.2.2}$$

We start with a general result which applies to all SMA(q) processes.

Theorem 2.2.1 *The distribution of the aggregated process* $\left\{S_t^{(r)}\right\}$ *is given by*

$$S_t^{(r)} \sim S_\alpha^0\left(\beta^{(r)}, \gamma^{(r)}, \delta^{(r)}\right) \tag{2.2.3}$$

where

$$\beta^{(r)} = \frac{\sum_{j=0}^{r+q-1} sign\left(c_j^{(r)}\right)\left|c_j^{(r)}\right|^{\alpha}}{\sum_{j=0}^{r+q-1}\left|c_j^{(r)}\right|^{\alpha}}\beta_e, \tag{2.2.4}$$

$$\gamma^{(r)} = \left(\sum_{j=0}^{r+q-1}\left|c_j^{(r)}\right|^{\alpha}\right)^{1/\alpha}\gamma_e, \tag{2.2.5}$$

if $\alpha \neq 1$

$$\begin{aligned}\delta^{(r)} &= \left(\sum_{j=0}^{r+q-1} c_j^{(r)}\right)\delta_e + \\ &\tan(\pi\alpha/2)\left[\beta^{(r)}\gamma^{(r)} - \beta_e\gamma_e\left(\sum_{i=0}^{r+q-1} sign\left(c_j^{(r)}\right)\left|c_j^{(r)}\right|\right)\right]\end{aligned} \tag{2.2.6}$$

if $\alpha = 1$

$$\begin{aligned}\delta^{(r)} &= \left(\sum_{j=0}^{r+q-1} c_j^{(r)}\right)\delta_e + \\ &\frac{2}{\pi}\left[\beta^{(r)}\gamma^{(r)}\ln\gamma^{(r)} - \beta_e\gamma_e\left(\sum_{i=0}^{r+q-1} sign\left(c_j^{(r)}\right)\left|c_j^{(r)}\right|\ln\left(\left|c_j^{(r)}\right|\gamma_e\right)\right)\right]\end{aligned} \tag{2.2.7}$$

and

$$c_j^{(r)} = \sum_{k=\max(j-r+1,0)}^{\min(j,q)} \theta_k. \tag{2.2.8}$$

Proof. From the definition of the base process $\{X_t\}$ and the aggregated process $\left\{S_t^{(r)}\right\}$, we have

$$\begin{aligned}S_t^{(r)} &= \sum_{i=0}^{r-1} X_{t-i} \\ &= \sum_{i=0}^{r-1}\sum_{j=0}^{q}\theta_j e_{t-i-j} \\ &= \sum_{j=0}^{r+q-1} c_j^{(r)} e_{t-j}\end{aligned} \tag{2.2.9}$$

where $c_j^{(r)}$ is given by (2.2.8). Thus $S_t^{(r)}$ is a linear combination of IID stable random variables. An application of parts a) and c) of Lemma E.1.1 proves the theorem. ■

Whilst Theorem 2.2.1 provides formulae for the stable distribution parameters of the aggregated process, in general it is not possible to derive from these a formula for the log quantile difference of the aggregated process. However, such a formula can be derived under certain conditions on the base process.

Theorem 2.2.2 *If the stable distribution skewness parameters of the innovations* $\{e_t\}$ *and the aggregated process* $\left\{S_t^{(r)}\right\}$ *are equal, that is*

$$\beta^{(r)} = \beta_e \tag{2.2.10}$$

where $\beta^{(r)}$ *is given by* (2.2.4), *then the log quantile difference* $\zeta_p^{(r)}$ *is given by*

$$\zeta_p^{(r)} = \ln\left(\gamma^{(r)}/\gamma_e\right) + \zeta_p. \tag{2.2.11}$$

Proof. From Theorem 2.2.1, we have that the aggregated process, $\left\{S_t^{(r)}\right\}$, has a stable distribution given by

$$S_t^{(r)} \sim S_\alpha^0\left(\beta^{(r)}, \gamma^{(r)}, \delta^{(r)}\right) \tag{2.2.12}$$

where $\beta^{(r)}, \gamma^{(r)}$ and $\delta^{(r)}$ are as shown in (2.2.4), (2.2.5) and (2.2.6) or (2.2.7). If $\beta^{(r)} = \beta_e$, then $\left\{S_t^{(r)}\right\}$ is a scale and location transformation of the innovations $\{e_t\}$. Thus

$$\frac{S_t^{(r)} - \delta^{(r)}}{\gamma^{(r)}} \sim \frac{e_t - \delta_e}{\gamma_e} \tag{2.2.13}$$

and Lemma E.1.2 gives (2.2.12). ■

The requirement of Theorem 2.2.2 that $\beta^{(r)} = \beta_e$ is satisfied if either

$$\beta_e = 0 \tag{2.2.14}$$

or

$$c_j^{(r)} \geq 0 \quad \text{for } j = 0, \ldots, r+q-1 \tag{2.2.15}$$

Note that

$$c_0^{(r)} = 1 \quad \text{for all } r. \tag{2.2.16}$$

The following corollary to Theorem 2.2.2 holds at all aggregation levels.

Corollary 2.2.1 *If* $\beta_e = 0$ *or* $\theta_k \geq 0$ *for* $k = 0, \ldots, q$ *then*

$$\zeta_p^{(r)} = \ln\left(\gamma^{(r)}/\gamma^{(1)}\right) + \zeta_p^{(1)}. \tag{2.2.17}$$

for any $r = 1, 2, \ldots$.

For $r \geq q$ we can slightly relax the assumptions of Corollary 2.2.1 to obtain the following result.

Corollary 2.2.2 *If* $\beta = 0$ *or*

$$\sum_{j=0}^{i} \theta_j \geq 0 \quad \textit{for } i = 0, \ldots, q-1 \textit{ and } \sum_{j=i}^{q} \theta_j \geq 0 \quad \textit{for } i = 1, \ldots, q \tag{2.2.18}$$

then for any $r \geq q$ *the log quantile difference* $\zeta_p^{(r)}$ *is given by the formula in* (2.2.1).

Proof. For any $r \geq q$, the condition (2.2.18) is sufficient for all the $c_j^{(r)}$ terms to be non-negative. Note that $\sum_{j=1}^{q} \theta_j \geq 0$ implies that $\sum_{j=0}^{q} \theta_j > 0$. Substitution of the formula for $\gamma^{(r)}$ in (2.2.5) into (2.2.11) yields (2.2.1). ■

The simple nature of (2.2.1) as a function of the aggregation level r, allows the derivation of some interesting results regarding the properties of $\zeta_p^{(r)}$.

Although for our purposes the formula for $\zeta_p^{(r)}$ in (2.2.1) is only valid for integer values of $r \geq q$, nonetheless it is a function of r which is well defined for all real positive values of r. Formally we can take partial derivatives of $\zeta_p^{(r)}$ with respect to $\ln r$, to get for $r \geq q$

$$\frac{\partial}{\partial \ln r}\zeta_p^{(r)} = \alpha^{-1}\frac{r\left|\sum_{i=0}^{q}\theta_i\right|^{\alpha}}{r\left|\sum_{i=0}^{q}\theta_i\right|^{\alpha} + g_\alpha(\theta)} \tag{2.2.19}$$

and

$$\frac{\partial^2}{(\partial \ln r)^2}\zeta_p^{(r)} = \alpha^{-1}\frac{r\left|\sum_{i=0}^{q}\theta_i\right|^{\alpha} g_\alpha(\theta)}{\left(r\left|\sum_{i=0}^{q}\theta_i\right|^{\alpha} + g_\alpha(\theta)\right)^2} \tag{2.2.20}$$

and draw conclusions on the shape of $\zeta_p^{(r)}$ as a function of $\ln r$.

Corollary 2.2.3 *If the assumptions of Corollary 2.2.2 hold then*

$$\lim_{r\to\infty}\frac{\partial}{\partial \ln r}\zeta_p^{(r)} = \alpha^{-1}. \tag{2.2.21}$$

For $r \geq q$,

$$sign\left(\frac{\partial^2}{(\partial \ln r)^2}\zeta_p^{(r)}\right) = sign\left(g_\alpha(\theta)\right) \tag{2.2.22}$$

and therefore

$$\begin{aligned} &\text{if } g_\alpha(\theta) > 0 \text{ then } \zeta_p^{(r)} \text{ is convex in } \ln r, \\ &\text{if } g_\alpha(\theta) = 0 \text{ then } \zeta_p^{(r)} \text{ is linear in } \ln r, \\ &\text{if } g_\alpha(\theta) < 0 \text{ then } \zeta_p^{(r)} \text{ is concave in } \ln r. \end{aligned} \tag{2.2.23}$$

Remark 2.2.1 *The derivatives in* (2.2.19) *and* (2.2.20) *and therefore the results of Corollary 2.2.3 do not depend on p for all $r \geq q$ and all α.*

Remark 2.2.2 *A general formula for the characteristic function of the temporal aggregation of a moving average process is derived in McKenzie (1988). This formula can be used in an alternative proof of Theorem 2.2.1.*

Remark 2.2.3 *The function $g_\alpha(\theta)$ is finite and does not depend on the aggregation level r. Thus for large r we can approximate* (2.2.1) *to get*

$$\zeta_p^{(r)} \approx \alpha^{-1} \ln r + \ln\left(\left|\sum_{i=0}^{q} \theta_i\right|\right) + \zeta_p \tag{2.2.24}$$

which describes a linear relationship between $\zeta_p^{(r)}$ and $\ln r$.

Remark 2.2.4 *If $\beta^{(r)} \neq \beta_e$, .then* (2.2.13) *in the proof of Theorem 2.2.2 does not hold. General equality relations for the quantiles of sums of random variables in terms of the quantiles of the summands are difficult to achieve. (Watson and Gordon (1986), Liu and David (1989))*

Remark 2.2.5 *In the special case where the base process $\{X_t\}$ is* IID*, we have*

$$\beta^{(r)} = \beta_e, \tag{2.2.25}$$

$$\gamma^{(r)} = r^{1/\alpha}\gamma_e, \tag{2.2.26}$$

$$\delta^{(r)} = \begin{cases} r\delta_e + \tan(\pi\alpha/2)\,\beta_e\gamma_e\left(r^{1/\alpha} - r\right) & \text{if } \alpha \neq 1 \\ r\delta_e + \dfrac{2}{\pi}\beta_e\gamma_e r \ln r & \text{if } \alpha = 1 \end{cases} \tag{2.2.27}$$

and the expression for $\zeta_p^{(r)}$ in (2.2.1) *reduces to*

$$\begin{aligned} \zeta_p^{(r)} &= \alpha^{-1}\ln r + \zeta_p \\ &= \alpha^{-1}\ln r + \ln\gamma + \zeta_p^*. \end{aligned} \tag{2.2.28}$$

where ζ_p^* is the log quantile difference of $S_\alpha^0(\beta_e, 1, 0)$ at quantile level p. Note that the expressions for $\delta^{(r)}$ in (2.2.27) *are different from those derived in Section 2.2 of Chan et al. (2008) which the author believes to be in error.*

Remark 2.2.6 *Using the results of Lemma E.1.2, it can seen that $\zeta_p^{(r)}$ does not depend on δ_e.*

2.3 Invertible SMA(1) and SMA(2) processes

In order to improve our understanding of the results in Section 2.2, in this section we explore further their application to the classes of invertible SMA(1) and SMA(2) processes.

More specifically, we identify sub-regions within the invertibility region of the SMA(1) and SMA(2) process where $g_\alpha(\theta)$ is either positive, zero or negative for various values of α.

An invertible SMA(q) process is one where all the roots of the moving average polynomial (Definition C.1.8)

$$1 + \theta_1 z + \cdots + \theta_q z^q = 0 \tag{2.3.1}$$

lie outside the complex unit circle, $|z| > 1$. The subset of $\mathbb{R}^q$ comprising the invertible parameters of an SMA(q) process is referred to as the invertibility region. The invertibility region of the SMA(1) process is the set

$$\{\theta_1 : |\theta_1| < 1\}. \tag{2.3.2}$$

The invertibility region of the SMA(2) process is the set

$$\{(\theta_1, \theta_2) : \theta_2 < 1 \text{ and } \theta_1 + \theta_2 > -1 \text{ and } \theta_1 - \theta_2 < 1\}. \tag{2.3.3}$$

Expressions for the invertibility region of higher order SMA(q) processes can be found in Wise (1956).

To assist with this analysis we divide the invertibility region of the SMA(2) process into five sub-regions as shown in Figure 2.3.1. These sub-regions are defined as open sets, so that the entire invertibility region consists of the union of the five sub-regions, the borders between them and the origin. The inequalities defining these sub-regions are listed in (2.3.4).

$$\begin{aligned}
&\text{Sub-Region 1} \quad \{(\theta_1, \theta_2) : \theta_1 < -1 \text{ and } \theta_2 < 1 \text{ and } \theta_1 + \theta_2 > -1\} \\
&\text{Sub-Region 2} \quad \{(\theta_1, \theta_2) : \theta_1 > -1 \text{ and } \theta_2 > 0 \text{ and } \theta_1 + \theta_2 < 0\} \\
&\text{Sub-Region 3} \quad \{(\theta_1, \theta_2) : \theta_2 > 0 \text{ and } \theta_2 < 1 \text{ and } \theta_1 + \theta_2 > 0 \text{ and } \theta_1 - \theta_2 < 1\} \\
&\text{Sub-Region 4} \quad \{(\theta_1, \theta_2) : \theta_2 < 0 \text{ and } -1 < \theta_1 + \theta_2 < 0 \text{ and } \theta_1 - \theta_2 < 1\} \\
&\text{Sub-Region 5} \quad \{(\theta_1, \theta_2) : \theta_2 < 0 \text{ and } \theta_1 + \theta_2 > 0 \text{ and } \theta_1 - \theta_2 < 1\}
\end{aligned} \tag{2.3.4}$$

This analysis of SMA(2) processes encompasses an analysis of SMA(1) processes and IID stable processes. The invertibility region of the SMA(1) process consists of the border between sub-regions 2 and 4, the border between sub-regions 3 and 5 and the origin. The invertibility region of the IID stable process is located at the origin.

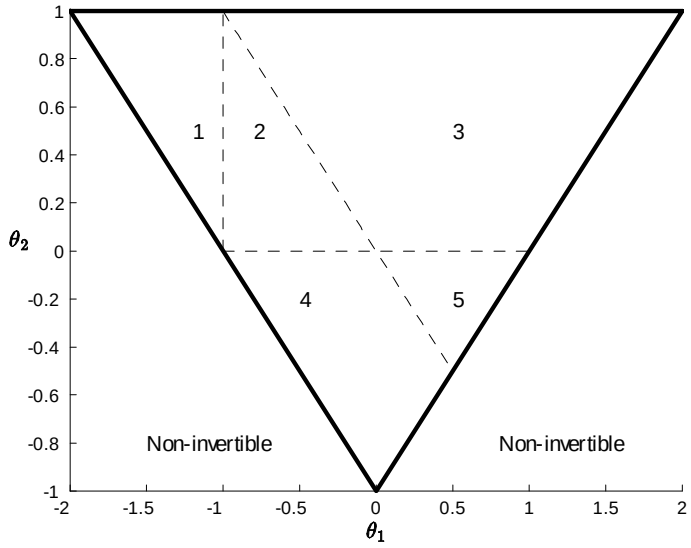


Figure 2.3.1: The five sub-regions of the invertibility region of the SMA(2) process.

Remark 2.3.1 *For an invertible* SMA(2) *process, the set of values of* (θ_1, θ_2) *which satisfy* (2.2.18) *in Corollary 2.2.2 consists of sub-region 3 and its borders with sub-regions 2 and 5.*

A sub-region is said to be positive, zero or negative for a given α if $g_\alpha(\theta)$ is respectively positive, zero or negative for all points in the sub-region. A sub-region is said to be mixed for a given α if there exist some points in the sub-region for which $g_\alpha(\theta)$ is positive and other points for which $g_\alpha(\theta)$ is negative. Similar descriptions are used to describe the borders between the sub-regions.

Theorem 2.3.1 *The categorisation of the sub-regions of the invertibility region of the* SMA(2) *process according to the values of* $g_\alpha(\theta)$ *in the sub-region is listed in Table 2.3.1. The categorisation of the sub-regions of the invertibility region of the* SMA(2) *process according to the values of* $g_\alpha(\theta)$ *in the sub-region is listed in Table 2.3.1. At the origin,* $g_\alpha(\theta) = 0$ *for all* α.

Proof. See Appendix A.1. ■

Corollary 2.3.1 *For an invertible* SMA(1) *process , the function* $g_\alpha(\theta)$ *satisfies the follow-*

Sub-Region	$0 < \alpha < 1$	$\alpha = 1$	$1 < \alpha \leq 2$
Positive	All	1,2,4 and 5	1 and 4
Zero	None	3	None
Negative	None	None	3
Mixed	None	None	2 and 5

Table 2.3.1: Categorisation of the sub-regions of the invertibility region of the SMA(2) process into positive, zero, negative and mixed sub-regions with respect to the sign of $g_\alpha(\theta)$.

Border	$0 < \alpha < 1$	$\alpha = 1$	$1 < \alpha \leq 2$
Positive	All	(1,2), (2,4) and (4,5)	(1,2), (2,4) and (4,5)
Zero	None	(2,3) and (3,5)	None
Negative	None	None	(2,3) and (3,5)

Table 2.3.2: Categorisation of the sub-region borders of the invertibility region of the SMA(2) process into positive, zero and negative borders with respect to the sign of $g_\alpha(\theta)$. We use (a,b) to denote the border between sub-regions a and b.

ing relations

$$g_\alpha(\theta) \text{ is } \begin{cases} > 0 & \text{if } \theta_1 < 0 \text{ or } \theta_1 > 0 \text{ and } \alpha < 1 \\ = 0 & \text{if } \theta_1 = 0 \text{ or } \theta_1 > 0 \text{ and } \alpha = 1 \\ < 0 & \text{if } \theta_1 > 0 \text{ and } \alpha > 1 \end{cases} . \tag{2.3.5}$$

Proof. This corollary follows immediately from the results in Theorem 2.3.1 for the origin and for the borders between sub-regions $(2,4)$ and $(3,5)$. ■

It is perhaps helpful to see the results of Theorem 2.3.1 in graphical form as provided in Figure 2.3.2. Figure 2.3.2(a) is applicable to $g_\alpha(\theta)$ for any $\alpha \in (0,1)$. Whilst Figures 2.3.2(c) and 2.3.2(d) appear similar, the locations of the respective green lines, i.e. the sets

$$D_\alpha = \{\theta : g_\alpha(\theta) = 0\} \tag{2.3.6}$$

are not the same.

Remark 2.3.2 *For an* SMA(2) *process it is straightforward to show that*

$$D_2 = \{\theta : \theta_1 + 2\theta_2 + \theta_1\theta_2 = 0\}. \tag{2.3.7}$$

For $\alpha \in (1,2)$, closed form expressions for D_α have not been obtained except to note that D_α contains the points $\theta = (1,0,0)'$ and $\theta = (1,-1,1)'$. Strictly $\theta = (1,-1,1)'$ is on the border of but not in the invertibility region.

To illustrate the behaviour of $\zeta_p^{(r)}$ where θ lie in different sub-regions of the invertibility region, we present plots of $\zeta_p^{(r)}$ for $p = (0.50, 0.95)'$, $\beta = 0$ and various combinations of α, θ_1 and θ_2 in Figure 2.3.3. The dotted parallel lines in Figure 2.3.3 have a slope $1/\alpha$.

As shown in Corollary 2.2.3, for each choice of α, θ in Figure 2.3.3, it can be seen that the plot of $\zeta_p^{(r)}$ against $\ln r$ is concave, linear or convex wherever $g_\alpha(\theta)$ is negative, zero or positive and that the sign of $g_\alpha(\theta)$ agrees with the results in Table 2.3.1. In all cases the derivative $\partial\zeta_p^{(r)}/\partial \ln r$ approaches $1/\alpha$ with increasing r.

The convergence of the derivative $\partial\zeta_p^{(r)}/\partial \ln r$ to $1/\alpha$ can be much slower in the positive sub-regions than in the negative sub-regions. The example shown in Figure 2.3.3(d) for $\alpha = 2$ and sub-region 1, still has a derivative $\partial\zeta_p^{(r)}/\partial \ln r$ much less than $1/\alpha$ at an aggregation level of $\exp(3.8) \approx 45$.

To conclude this chapter, we show that two different invertible SMA(2) processes can have the same log quantile differences at all aggregation levels and all quantile levels, i.e. the log quantile difference does not necessarily identify an SMA(2) process.

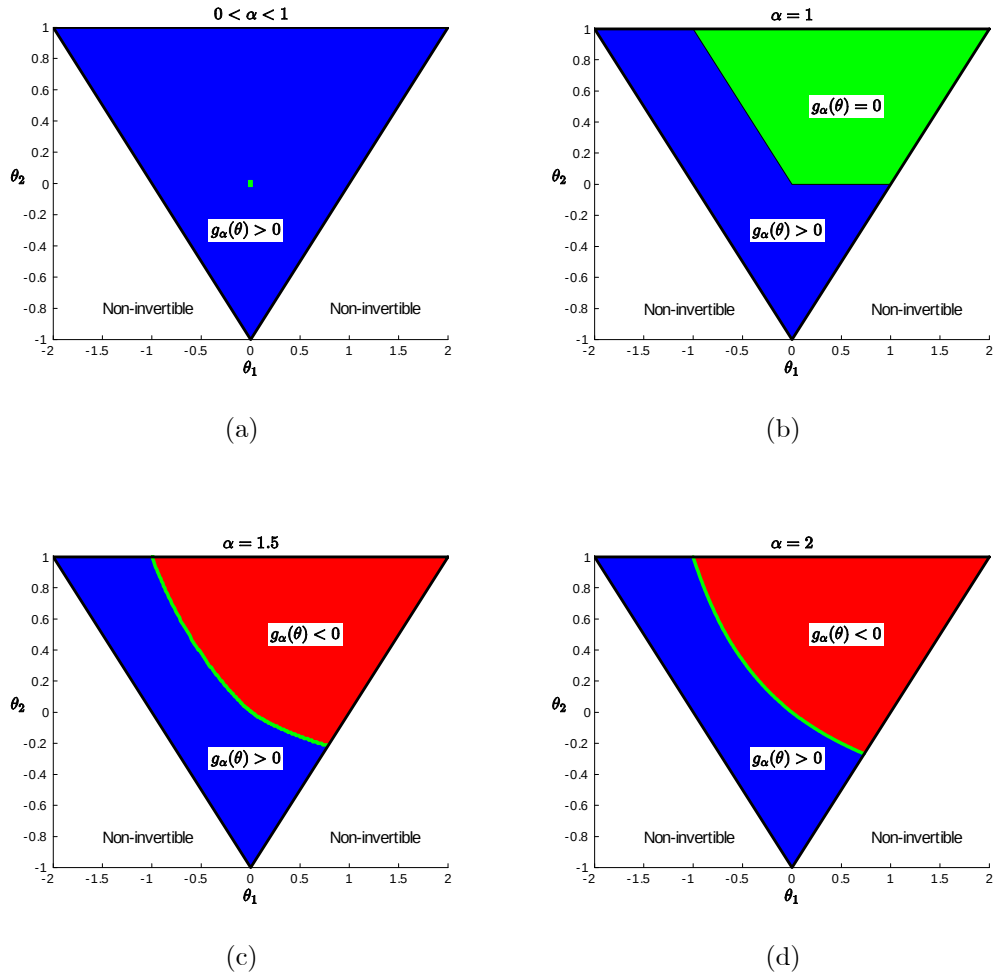


Figure 2.3.2: A graphical display of the categorisation of the invertibility region of the SMA(2) process with respect to the sign of $g_\alpha(\theta)$ into positive (blue), zero (green) and negative (red) regions for (a) $0 < \alpha < 1.0$, (b) $\alpha = 1.0$, (c) $\alpha = 1.5$ and (d) $\alpha = 2.0$.

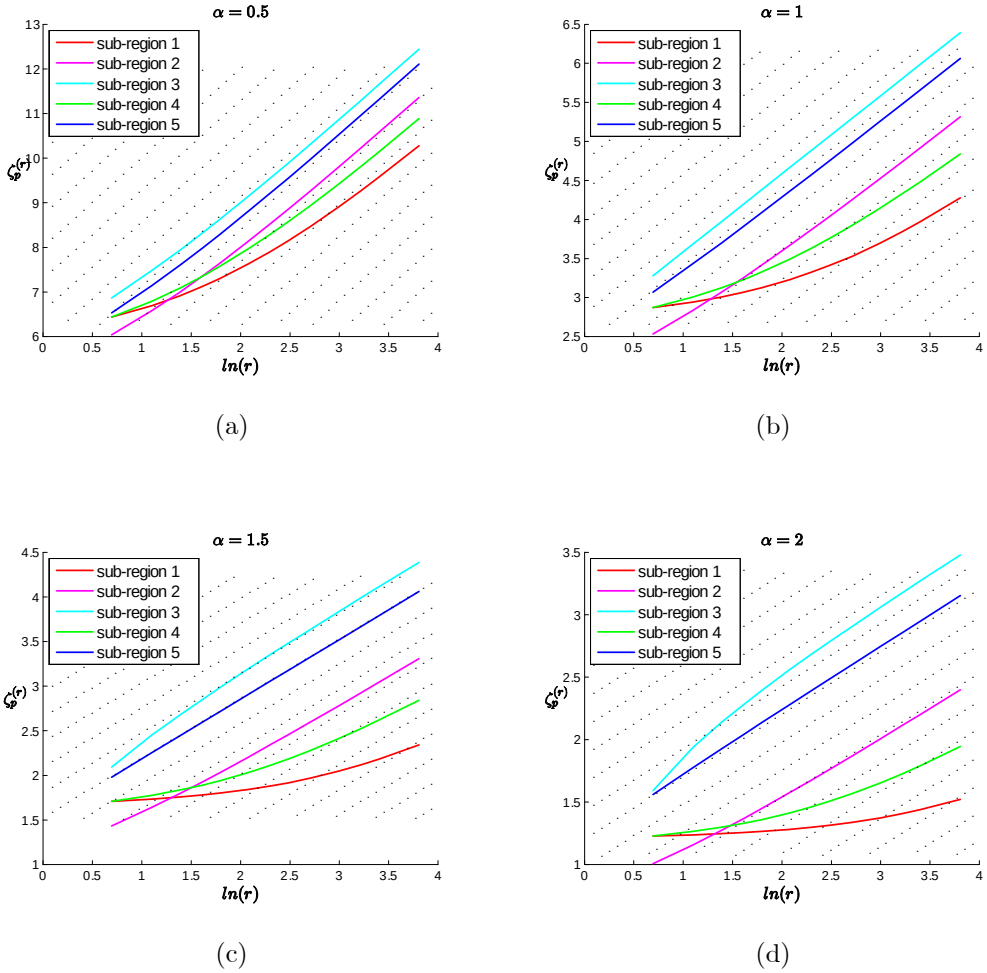


Figure 2.3.3: Plots of the log quantile difference values at quantile levels $p = (0.50, 0.95)$ of an aggregated symmetric SMA(2) process for various values of α with parameters in each of the sub-regions of the invertibility region. For sub-region 1 (θ_1,θ_2) = (-1.4,0.6), sub-region 2 (θ_1,θ_2) = (-0.5,0.2), sub-region 3 (θ_1,θ_2) = (0.2,0.9), sub-region 4 (θ_1,θ_2) = (-0.2,-0.4) and sub-region 5 (θ_1,θ_2) = (0.7,-0.2). The dotted lines are at a slope of $1/\alpha$.

Theorem 2.3.2 *Let* $\{X_t\}$ *be the* SMA(2) *process*

$$X_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} \tag{2.3.8}$$

where

$$e_t \sim S_\alpha^0 (0, \gamma_e, \delta_e) . \tag{2.3.9}$$

If $\{X_t\}$ *is invertible and its parameters satisfy*

$$\theta_1 < 0 \text{ and } \theta_1 - \theta_2 \neq 1, \tag{2.3.10}$$

then there exists an invertible SMA(2) *process,* $\{Y_t\}$ *, given by*

$$Y_t = u_t + \psi_1 u_{t-1} + \psi_2 u_{t-2} \tag{2.3.11}$$

and

$$u_t \sim S_\alpha^0 (0, \gamma, \delta) \tag{2.3.12}$$

which has the same log quantile differences as $\{X_t\}$ *for all aggregation levels* $r \geq 2$ *and all quantile levels and where*

$$\theta_1 \neq \psi_1 \text{ and } \theta_2 \neq \psi_2. \tag{2.3.13}$$

Proof. This theorem is proved by demonstration, choosing

$$\psi_1 = -1 + \theta_2 \text{ and } \psi_2 = 1 + \theta_1. \tag{2.3.14}$$

We note that if $\theta_1 - \theta_2 \neq 1$, then $\theta_1 \neq \psi_1$ and $\theta_2 \neq \psi_2$.

For $\{Y_t\}$ to be invertible, the parameters ψ_1, ψ_2 need to satisfy the inequalities in (2.3.3) . By assumption $\theta_1 < 0$ and therefore

$$\psi_2 = 1 + \theta_1 < 1. \tag{2.3.15}$$

The process $\{X_t\}$ is assumed invertible and therefore

$$\psi_1 + \psi_2 = \theta_1 + \theta_2 > -1. \tag{2.3.16}$$

Further $\theta_2 < 1$ and $\theta_1 > -2$, therefore

$$\begin{aligned} \psi_2 - \psi_1 &= \theta_2 - \theta_1 - 2 \\ &< -\theta_1 - 1 \\ &< 1 \end{aligned} \tag{2.3.17}$$

and consequently $\{Y_t\}$ is invertible.

Since $\{X_t\}$ and $\{Y_t\}$ both satisfy (2.2.18) it follows from Corollary 2.2.2 that for $r \geq 2$

$$\zeta_{X;p}^{(r)} = \alpha^{-1} \ln \left(r \left| 1 + \theta_1 + \theta_2 \right|^{\alpha} + g_{\alpha} \left(\theta_1, \theta_2 \right) \right) + \zeta_p, \tag{2.3.18}$$

$$\zeta_{Y;p}^{(r)} = \alpha^{-1} \ln \left(r \left| 1 + \psi_1 + \psi_2 \right|^{\alpha} + g_{\alpha} \left(\psi_1, \psi_2 \right) \right) + \zeta_p \tag{2.3.19}$$

where $\zeta_{X;p}^{(r)}, \zeta_{Y;p}^{(r)}$ are the log quantile differences of $\{X_t\}$ and $\{Y_t\}$ respectively at aggregation level r and quantile levels p_1 and p_2. Note that ζ_p is the log quantile difference of both $\{e_t\}$ and $\{u_t\}$ at quantile level p. Since

$$\begin{aligned} g_{\alpha} \left(\theta \right) &= 1 + \left| 1 + \theta_1 \right|^{\alpha} - 2 \left| 1 + \theta_1 + \theta_2 \right|^{\alpha} + \left| \theta_1 + \theta_2 \right|^{\alpha} + \left| \theta_2 \right|^{\alpha} \\ &= 1 + \left| \psi_2 \right|^{\alpha} - 2 \left| 1 + \psi_1 + \psi_2 \right|^{\alpha} + \left| \psi_1 + \psi_2 \right|^{\alpha} + \left| 1 + \psi_1 \right|^{\alpha} \\ &= g_{\alpha} \left(\psi \right), \end{aligned} \tag{2.3.20}$$

it follows that

$$\zeta_{X;p}^{(r)} = \zeta_{Y;p}^{(r)} \tag{2.3.21}$$

and the theorem is proved. ■

Chapter 3

Estimation of the parameters of a stable moving average process

3.1 Introduction

Let $\{X_t\}$ be the SMA(q) process (Definition 1.2.2),

$$X_t = \sum_{j=0}^{q} \theta_j e_{t-j} \tag{3.1.1}$$

where $\theta_0 = 1$ and $\{e_t\}$ is an IID sequence of stable random variables such that

$$e_t \sim S_\alpha^0 (\beta_e, \gamma_e, \delta_e) \tag{3.1.2}$$

using the S^0 parameterisation of stable random variables in Nolan (1998). Let θ denote the $q+1$ dimensional vector of moving average parameters

$$\theta = (\theta_0, \ldots, \theta_q)' . \tag{3.1.3}$$

In Theorem 2.2.1 it was shown that

$$X_t \sim S_\alpha^0 \left(\beta^{(1)}, \gamma^{(1)}, \delta^{(1)}\right) \tag{3.1.4}$$

and formulae for $\beta^{(1)}$, $\gamma^{(1)}$ and $\delta^{(1)}$ were provided. For brevity, in this chapter, we drop the "(1)" superscripts from the stable distribution parameter symbols in 3.1.4 and instead say that

$$X_t \sim S_\alpha^0 (\beta, \gamma, \delta) \tag{3.1.5}$$

Let

$$\omega = (\alpha, \beta, \gamma, \delta)' \tag{3.1.6}$$

denote the vector of stable distribution parameters of X_t. In this chapter we construct a method for the asymptotically normal estimation of the parameters ω from an SMA process.

A number of methods have been proposed for the estimation of the parameters of a stable distribution. A method based on sample quantiles was proposed by Fama and Roll (1971) which was simple to implement, but was only applicable to symmetric stable distributions with $\alpha \geq 1$ and contained a bias. This method was extended by McCulloch (1986) to cover asymmetric stable distributions and is asymptotically unbiased. A more recent discussion on the related use of indirect inference for the estimation of stable distributions is included in Garcia et al. (2011). Methods based on the sample characteristic function were proposed by Press (1972), Paulson et al. (1975) and Kogon and Williams (1998). Maximum likelihood estimation methods were investigated in Brorsen and Yang (1990) and Nolan (2001). The methods listed above assume an independent sample is available. In this thesis, we require a method which can be used to estimate stable distribution parameters from a dependent sample, more specifically a sample from an SMA process.

In Section 3.2, we provide the asymptotic distribution of empirical quantile estimators from an SMA process based on results from Sen (1972). In Section 3.3 we show how to use these quantile estimators to adapt the quantile-based stable distribution parameter estimators of McCulloch (1986) to construct asymptotically normal estimators of the stable distribution parameters from an SMA process.

The quantile-based stable distribution parameter estimators have traditionally used quantile estimators at the standard quantile levels proposed by McCulloch (1986). In Section 3.5, we show that more efficient quantile-based stable distribution parameter estimators can be constructed from quantile estimators at other quantile levels. The optimal choice of quantile levels is shown to be dependent on the stable distribution parameters. The relative asymptotic efficiency of these optimal quantile-based stable distribution parameter estimators is calculated against the maximum likelihood stable distribution parameter estimators. A new method for the evaluation of the asymptotic distribution of maximum likelihood stable distribution parameter estimators is proposed in Section 3.4.

Many of the standard methods available for the estimation of finite variance ARMA

processes, such as Gaussian maximum likelihood, least squares, Yule-Walker and Hannan-Risannen etc, are not valid for the estimation of infinite variance ARMA processes. Alternative methods have been proposed by several authors, but the estimators from many of these methods have complicated asymptotic distributions which makes them very difficult to use. (Hannan and Kanter (1977), Davis et al. (1992), Mikosch et al. (1995) and Davis (1996)). A self-weighted least absolute deviation (SLAD) estimator was proposed by Ling (2005) for infinite variance autoregressive processes and was shown to have an asymptotically normal distribution. The SLAD estimator was extended to cover infinite variance ARMA processes in Pan et al. (2007). The asymptotic properties of the global SLAD estimator were established by Zhu and Ling (2012) and it is the global SLAD estimator, hereafter referred to simply as the SLAD estimator, which is used for the estimation of θ from SMA processes throughout this thesis. Some of the theory behind SLAD estimators is included in Appendix C.3, together with a discussion on the various configuration choices which were made for this thesis. Note that selection of a weight function for the SLAD estimator ration requires either knowledge or an estimate of the stable distribution parameter α.

Standard methods for order identification and goodness of fit testing available in finite variance time series analysis are not necessarily appropriate for SMA processes. However, some methods involving the sample autocorrelation function can be adapted for SMA processes since, somewhat surprisingly, the autocorrelation function is well-defined for an SMA process, even though the autocovariance function may be infinite. The asymptotic distribution of the sample autocorrelation function was derived by Davis and Resnick (1986). Details are provided in Section 3.6. Unfortunately convergence to the asymptotic distribution for the sample autocorrelation function is very slow, hence the asymptotic distribution does not provide accurate significance levels for any autocorrelation function based statistical tests. In Sections 3.7 and 3.8, we show how simulations can be used to provide accurate significance levels for statistical tests which identify the order of an SMA process and test for the independence of residuals.

The results of simulations are reported in Section 3.9 which demonstrate the use of the methods described in this chapter. We also investigate the correlation between the estimators for ω and θ.

3.2 Quantile estimation from a stable moving average process

For any real-valued random variable X on a probability space $(\Omega, \mathcal{A}, P)$, there is an associated distribution function $F : \mathbb{R} \to [0,1]$ defined by

$$F(x) \equiv P(X \leq x), \tag{3.2.1}$$

from which we can construct a quantile function $Q : [0,1] \to \mathbb{R}$ defined by

$$Q(p) \equiv \inf\{x : F(x) \geq p\} \tag{3.2.2}$$

and a density function $f : \mathbb{R} \to \mathbb{R}_+$ defined by

$$F(x) \equiv \int_{-\infty}^{x} f(s)\, ds. \tag{3.2.3}$$

Often, it is convenient to denote the *pth* quantile of a distribution function by

$$\xi_p \equiv Q(p). \tag{3.2.4}$$

Let $\{x_j\}_{j=1}^n$ be a sample drawn from the distribution function F, then we define the empirical distribution function and the empirical quantile estimator.

Definition 3.2.1 ***(Empirical Distribution Function).*** *The empirical distribution function is given by*

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{j=1}^{n} I_{(-\infty,x]}(x_j). \tag{3.2.5}$$

Definition 3.2.2 ***(Empirical Quantile Estimator).*** *The empirical quantile estimator is given by*

$$\widehat{\xi}_{n;p} = \inf\left\{x : \widehat{F}_n(x) \geq p\right\}. \tag{3.2.6}$$

Where convenient, we may use $\widehat{\xi}_p$ instead of $\widehat{\xi}_{n;p}$ to denote the empirical quantile estimator.

The asymptotic distribution of the empirical quantile estimator from an IID process is well known, (Appendix D). The asymptotic distribution of the empirical quantile estimator, where the sample is taken from a possibly non-stationary m-dependent process was derived by Sen (1968). Further work in this area has been done by, amongst others: Dutta and Sen (1971) on autoregressive processes, Sen (1972) on ϕ-mixing processes, Oberhofer and Haupt (2005) on non-stationary processes and Dominicy et al. (2013) on S-mixing processes. In this section we apply these earlier results to find the joint asymptotic distribution of multiple empirical quantile estimators from a stable moving average process.

Definition 3.2.3 *(ϕ - mixing process). Let $\{X_t\}$ be a stationary process defined on a probability space $\{\Omega, \mathcal{A}, P\}$. Let $\mathcal{F}_n$ denote the σ - field generated by $\{X_t : t \leq n\}$ and $\mathcal{G}_n$ denote the σ - field generated by $\{X_t : t \geq n\}$. The ϕ_n measure of dependence between $\mathcal{F}_n$ and $\mathcal{G}_n$ is given by*

$$\phi_n = \sup\{|P(B \mid A) - P(B)| : A \in \mathcal{F}_0, P(A) > 0, B \in \mathcal{G}_n\}. \tag{3.2.7}$$

If

$$\lim_{n\to\infty} \phi_n = 0, \tag{3.2.8}$$

then the process $\{X_t\}$ is said to be ϕ - mixing.

An SMA(q) process is ϕ - mixing, since $\mathcal{F}_0$ and $\mathcal{G}_n$ are independent for all $n > q$ and therefore $\phi_n = 0$ for all $n > q$. A general result for the asymptotic distribution of the empirical quantile estimator from a ϕ - mixing process is given in Theorem 3.1 of Sen (1972). We adapt that result in the following theorem for the specific case of an SMA(q) process.

Theorem 3.2.1 *If $\{X_t\}$ is an* SMA(q) *process and $p \in (0,1)$, then*

$$\sqrt{n}\left(\widehat{\xi}_p - \xi_p\right) \xrightarrow{d} N(0, V_\xi) \tag{3.2.9}$$

where

$$V_\xi = \frac{\sum_{h=-q}^{q}\left(G_h\left(\xi_p\right) - p^2\right)}{f^2\left(\xi_p\right)} \tag{3.2.10}$$

and

$$G_h(\xi) = P\left(\left\{X_t \leq \xi_p\right\} \cap \left\{X_{t+h} \leq \xi_p\right\}\right). \tag{3.2.11}$$

Proof. An SMA(q) process is a stationary ϕ - mixing process with a continuous density function f at ξ_p satisfying

$$\sum_{n=1}^{\infty} \phi_n^{1/2} < \infty \tag{3.2.12}$$

and

$$0 < f\left(\xi_p\right) < \infty \tag{3.2.13}$$

and therefore we can apply Theorem 3.1 of Sen (1972) to prove this theorem. Note that for an SMA(q) process

$$G_h\left(\xi_p\right) = p^2, \quad \forall\, |h| > q. \tag{3.2.14}$$

■

In Sen (1972), the results for empirical quantile estimators from scalar ϕ - mixing processes are extended to cover multivariate ϕ - mixing processes. A general result for the asymptotic distribution of the empirical quantile estimator from a multivariate ϕ - mixing process is given in Theorem 6.5 of Sen (1972). We use that result in the following theorem for the specific case of the joint asymptotic distribution of multiple empirical quantile estimators from an SMA(q) process.

Theorem 3.2.2 *Let*

$$\boldsymbol{\xi} = (\xi_1, \ldots, \xi_k)' \tag{3.2.15}$$

denote the vector of true quantiles of an SMA(q) *process* $\{X_t\}$ *at the vector of quantile levels*

$$p = (p_1, \ldots, p_k)' \tag{3.2.16}$$

where $0 < p_1, \ldots, p_k < 1$ *and*

$$\widehat{\boldsymbol{\xi}} = \left(\widehat{\xi}_1, \ldots, \widehat{\xi}_k\right)' \tag{3.2.17}$$

denote the corresponding vector of empirical quantiles. Then

$$\sqrt{n}\left(\widehat{\boldsymbol{\xi}} - \boldsymbol{\xi}\right) \xrightarrow{d} N\left(0, V_{\boldsymbol{\xi}}\right) \tag{3.2.18}$$

where

$$V_{\boldsymbol{\xi}} = (v_{ij}), \tag{3.2.19}$$

$$v_{ij} = \frac{\sum_{h=-q}^{q}\left(G_h\left(\xi_i, \xi_j\right) - p_i p_j\right)}{f\left(\xi_i\right) f\left(\xi_j\right)} \tag{3.2.20}$$

and

$$G_h\left(\xi_i, \xi_j\right) = P\left(\left\{X_t \leq \xi_i\right\} \cap \left\{X_{t+h} \leq \xi_j\right\}\right). \tag{3.2.21}$$

Proof. Apply Theorem 6.5 of Sen (1972) to the process

$$\mathbf{X}_t = (X_{1;t}, \ldots, X_{k;t})' \tag{3.2.22}$$

where $X_{i;t}$ for i = $1, \ldots, k$ are identical copies of the SMA(q) process $\{X_t\}$. ■

In order to calculate the asymptotic variance, $V_{\boldsymbol{\xi}}$ from (3.2.19), it is necessary to calculate the joint probabilities, $G_h\left(\xi_i, \xi_j\right)$, for each $|h| \leq q$. For $h = 0$, $G_h\left(\xi_i, \xi_j\right)$ simplifies to become

$$G_0\left(\xi_i, \xi_j\right) = \min\left(p_i, p_j\right) \tag{3.2.23}$$

For $h \neq 0$, the evaluation of (3.2.21) whilst theoretically possible is computationally difficult for many SMA(q) processes. For an IID process we get

$$G_h\left(\xi_i, \xi_j\right) = 0, \quad \text{for } h \neq 0. \tag{3.2.24}$$

Thus for IID processes, the asymptotic covariance matrix of Theorem 3.2.2 reduces to the same asymptotic covariance matrix as in Theorem D.1.3.

Suppose X_t is the SMA(1) process defined by

$$X_t = e_t + \theta_1 e_{t-1}, \tag{3.2.25}$$

where $\{e_t\}$ is an IID sequence of random variables such that

$$e_t \sim S_\alpha^0\left(\beta_e, \gamma_e, \delta_e\right). \tag{3.2.26}$$

Let f_e and F_e denote respectively the density and distribution functions of $\{e_t\}$. Then

$$\begin{aligned} G_1\left(\xi_i, \xi_j\right) &= P\left(\left\{e_t + \theta_1 e_{t-1} \leq \xi_i\right\} \cap \left\{e_{t+1} + \theta_1 e_t \leq \xi_j\right\}\right) \\ &= P\left(\left\{e_{t-1} \leq \frac{\xi_i - e_t}{\theta_1}\right\} \cap \left\{e_{t+1} \leq \xi_j - \theta_1 e_t\right\}\right) \\ &= \int_{-\infty}^{\infty} F_e\left(\frac{\xi_i - u}{\theta_1}\right) F_e\left(\xi_j - \theta_1 u\right) f_e\left(u\right) du, \end{aligned} \tag{3.2.27}$$

which can be evaluated numerically. Note that

$$G_1\left(\xi_i, \xi_j\right) = G_{-1}\left(\xi_j, \xi_i\right). \tag{3.2.28}$$

For higher order SMA(q) processes, the evaluation of $G_h\left(\xi_i, \xi_j\right)$ becomes computationally difficult, involving a $q - 1 + h$ dimensional integral however, the estimation of $G_h\left(\xi_i, \xi_j\right)$ is straightforward. Let $\{x_t\}_{t=1}^n$ be a sample of size n from the stationary SMA(q) process $\{X_t\}$. We define the estimator $\widehat{G}_h\left(\xi_i, \xi_j\right)$ as

$$\widehat{G}_h\left(\xi_i, \xi_j\right) = (n-h)^{-1} \sum_{t=1} I\left\{x_t \leq \xi_i\right\} I\left\{x_{t+h} \leq \xi_j\right\}, \quad \text{for } |h| > 1 \tag{3.2.29}$$

and it is clear that $\widehat{G}_h\left(\xi_i, \xi_j\right)$ is a consistent estimator of $G_h\left(\xi_i, \xi_j\right)$.

3.3 Quantile-based estimation of stable distribution parameters from a stable moving average process

In this section we investigate an extension of the quantile-based method of McCulloch (1986), which we adapt for estimation from an SMA(q) process with distribution $S_\alpha^0\left(\beta, \gamma, \delta\right)$.

Let ξ_p denote the *pth* quantile of the stable distribution $S^0_\alpha(\beta,\gamma,\delta)$ and define the following statistics

$$v_\alpha = \frac{\xi_{0.95}-\xi_{0.05}}{\xi_{0.75}-\xi_{0.25}}, \tag{3.3.1}$$

$$v_\beta = \frac{\xi_{0.95}+\xi_{0.05}-2\xi_{0.50}}{\xi_{0.95}-\xi_{0.05}}. \tag{3.3.2}$$

From Lemma E.1.3, it can be shown that the statistics v_α and v_β do not depend on γ and δ and we can consider them as functions solely of α and β

$$v_\alpha = \phi_1(\alpha,\beta), \tag{3.3.3}$$

$$v_\beta = \phi_2(\alpha,\beta). \tag{3.3.4}$$

It can be seen that $\phi_1(\alpha,\beta)$ is a strictly decreasing function of α for each β and that $\phi_2(\alpha,\beta)$ is a strictly decreasing function of β for each α. The relationships (3.3.3) and (3.3.4) can be inverted to give

$$\alpha = \psi_1(v_\alpha, v_\beta), \tag{3.3.5}$$

$$\beta = \psi_2(v_\alpha, v_\beta). \tag{3.3.6}$$

No analytic formula is available for the functions ϕ_1, ϕ_2, ψ_1 and ψ_2. Our approach to the numerical evaluation of these functions is discussed later in this section.

Let $\widehat{\xi}_p$ denote a consistent estimator for ξ_p. Substituting the estimators $\widehat{\xi}_p$ into (3.3.1) and (3.3.2) gives consistent estimators for v_α, v_β,

$$\widehat{v}_\alpha = \frac{\widehat{\xi}_{0.95}-\widehat{\xi}_{0.05}}{\widehat{\xi}_{0.75}-\widehat{\xi}_{0.25}}, \tag{3.3.7}$$

$$\widehat{v}_\beta = \frac{\widehat{\xi}_{0.95}+\widehat{\xi}_{0.05}-2\widehat{\xi}_{0.50}}{\widehat{\xi}_{0.95}-\widehat{\xi}_{0.05}}. \tag{3.3.8}$$

Consistent estimators for the parameters α,β can then be calculated using

$$\widehat{\alpha} = \psi_1(\widehat{v}_\alpha, \widehat{v}_\beta), \tag{3.3.9}$$

$$\widehat{\beta} = \psi_2(\widehat{v}_\alpha, \widehat{v}_\beta). \tag{3.3.10}$$

We can use the results of Lemma E.1.3 to define estimators for γ and δ by

$$\widehat{\gamma} = \frac{\widehat{\xi}_{0.75}-\widehat{\xi}_{0.25}}{\widehat{\xi}^*_{0.75}-\widehat{\xi}^*_{0.25}} \tag{3.3.11}$$

and

$$\widehat{\delta} = \widehat{\xi}_{0.50} - \widehat{\gamma}\widehat{\xi}^{*}_{0.50}. \tag{3.3.12}$$

The estimators in (3.3.11) and (3.3.12) are similar to those defined in McCulloch (1986). The differences are due to McCulloch's choice of parameterisation for the stable distribution, which includes discontinuities at $\alpha = 1$.

Let

$$p_M = (0.05, 0.25, 0.50, 0.75, 0.95) \tag{3.3.13}$$

denote the quantile levels used to define the statistics in (3.3.1) and (3.3.2) and let

$$\xi_M = (\xi_{0.05}, \xi_{0.25}, \xi_{0.50}, \xi_{0.75}, \xi_{0.95})' \tag{3.3.14}$$

denote the quantiles at those levels. Other quantile levels not included in p_M could have been used to define $\widehat{\gamma}$ and $\widehat{\delta}$, although it seems natural to choose from the same quantile levels used to define $\widehat{\alpha}$ and $\widehat{\beta}$. Indeed, other choices of quantile levels are also available to define $\widehat{\alpha}$ and $\widehat{\beta}$ and it is possible that a different choice of quantile levels would produce better estimates. In Section 3.5, an investigation is conducted into the optimal choice of quantile levels for quantile-based stable distribution parameter estimation.

Let

$$\omega = (\alpha, \beta, \gamma, \delta)' \tag{3.3.15}$$

denote the vector of stable distribution parameters and

$$\widehat{\omega} = \left(\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma}, \widehat{\delta}\right)' \tag{3.3.16}$$

denote the vector of estimators defined above. An asymptotic distribution for $\widehat{\omega}$ was derived in McCulloch (1986) under the assumption of an independent sample. The assumption of an independent sample allows the use of empirical quantile estimators to form a consistent, asymptotically normal, estimator for ξ_M. However, we have seen that a consistent, asymptotically normal, estimator for ξ_M can also be defined from an SMA sample using the methods described in Section 3.2. Thus, we can use the same approach as was taken in McCulloch (1986) to derive an asymptotic distribution for $\widehat{\omega}$ under the assumption that the sample is an SMA process.

Theorem 3.3.1 *Let*

$$\widehat{\xi}_M = \left(\widehat{\xi}_{0.05}, \widehat{\xi}_{0.25}, \widehat{\xi}_{0.50}, \widehat{\xi}_{0.75}, \widehat{\xi}_{0.95}\right)' \tag{3.3.17}$$

be the empirical quantile estimators of ξ_M *from an* SMA *process, then as the sample size* $n \to \infty$ *we have from Theorem 3.2.2 we have*

$$\sqrt{n}\left(\widehat{\xi}_M - \xi_M\right) \xrightarrow{d} N\left(0, V_{\widehat{\xi}_M}\right). \tag{3.3.18}$$

The asymptotic distribution of the stable distribution parameter $\widehat{\omega}$ *is given by*

$$\sqrt{n}\left(\widehat{\omega} - \omega\right) \xrightarrow{d} N\left(0, V_{\widehat{\omega}}\right), \tag{3.3.19}$$

where

$$V_{\widehat{\omega}} = D_{\widehat{\omega}} V_{\widehat{\xi}_M} D'_{\widehat{\omega}} \tag{3.3.20}$$

and $D_{\widehat{\omega}}$ *is the* 4×5 *matrix given by*

$$D_{\widehat{\omega}} = \left(\frac{\partial \widehat{\omega}_i}{\widehat{\xi}_{M;j}} \mid_{\widehat{\xi}_M = \xi_M} \right)_{i=1,4;j=1,5}. \tag{3.3.21}$$

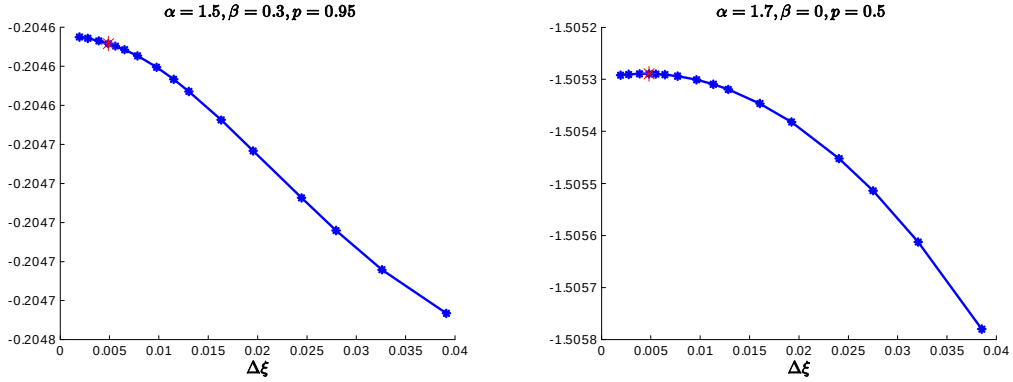
Proof. The proof of this theorem follows the same approach as was taken in McCulloch (1986), which is essentially a use of Theorem B.2.3. ■

The formula for the asymptotic covariance matrix $V_{\widehat{\omega}}$ in (3.3.20) has the same partial derivative matrix $D_{\widehat{\omega}}$ as that in McCulloch (1986) however, the asymptotic covariance matrix, $V_{\widehat{\xi}_M}$, of the quantile estimators is different from that in McCulloch (1986) due to sample being from an SMA process rather than from an IID process.

Remark 3.3.1 *If the true innovation values of the* SMA *process were observable, then they would form a sample from an* IID *process from which quantile-based stable distribution parameter estimates could be calculated. However, the estimated residuals of an* SMA *process are only approximately stable and do not necessarily provide unbiased estimates for the true innovation quantiles.*

A general analytic formula is not available for the calculation of the partial derivatives in (3.3.21). It is suggested in McCulloch (1986) that the partial derivatives can be estimated "by means of small perturbations of the population quantiles", but no specific recommendations regarding the size of these perturbations are given. To limit the scope of our investigation into this matter, we restrict ourselves to perturbations given by

$$\Delta\xi = \frac{\widehat{\xi}_{0.75} - \widehat{\xi}_{0.25}}{C} \tag{3.3.22}$$



(a) (b)

Figure 3.3.1: Estimates for (a) $\dfrac{\partial\widehat{\alpha}}{\partial\widehat{\xi}_{0.95}}$ where $\alpha = 1.5$ and $\beta = 0.3$ and (b) $\dfrac{\partial\widehat{\beta}}{\partial\widehat{\xi}_{0.50}}$ where $\alpha = 1.7$ and $\beta = 0.0$. The estimates for C = 400 are those indicated by a red '*'.

for some $C > 0$ and assume that the same perturbation is applied to each quantile estimator. Let $\widehat{\xi} = \left(\widehat{\xi}_{0.05}, \widehat{\xi}_{0.25}, \widehat{\xi}_{0.50}, \widehat{\xi}_{0.75}, \widehat{\xi}_{0.95}\right)$ be the sample quantiles. Let $\widehat{\alpha}_p^+$ be the estimate of α derived from the set of quantiles where $\widehat{\xi}_p$ is replaced by $\widehat{\xi}_p + \Delta\xi$ and $\widehat{\alpha}_p^-$ be the estimate of α derived from the set of quantiles where $\widehat{\xi}_p$ is replaced by $\widehat{\xi}_p - \Delta\xi$. Similarly, we define $\widehat{\beta}_p^+, \widehat{\beta}_p^-$, etc. Our estimate for $\dfrac{\partial\widehat{\alpha}}{\partial\xi_p}$ is defined to be

$$\widehat{\frac{\partial\widehat{\alpha}}{\partial\widehat{\xi}_p}} = \frac{\widehat{\alpha}_p^+ - \widehat{\alpha}_p^-}{2\Delta\xi} \tag{3.3.23}$$

with similar definitions for $\widehat{\dfrac{\partial\widehat{\beta}}{\partial\xi_p}}$, $\widehat{\dfrac{\partial\widehat{\gamma}}{\partial\xi_p}}$ and $\widehat{\dfrac{\partial\widehat{\delta}}{\partial\xi_p}}$.

Estimates for each of the partial derivative estimators were calculated for various stable distributions. Examples of these calculations are presented in Figure 3.3.1 for values of C between 50 and 1000. The optimal choice for C is not obvious, given we do not have any true values for the partial derivatives, although in general the value of the partial derivative estimates does not change greatly for C between 50 and 1000. A slightly lower value of C and hence slightly larger perturbation can help to smooth the partial derivatives and avoid occasional numerical abberations. Unless otherwise stated, in the remainder of this thesis we use $C = 400$ to calculate the partial derivative estimates.

Remark 3.3.2 *The quantile-based method for stable distribution parameter estimation can easily be adapted for the estimation of the parameters of a stable distribution which is known to be symmetric. A symmetric stable distribution has parameters β and δ and the statistic ν_β equal to zero. The estimator $\widehat{\nu}_\alpha$ is still calculated as per* (3.3.7)*, but $\widehat{\nu}_\beta$ is set to zero. Estimates for α and γ are then calculated as described above. In calculating the asymptotic covariance matrix $V_{\widehat{\omega}}$, it is necessary to set to zero all elements of the partial derivative matrix $D_{\widehat{\omega}}$ corresponding to partial derivatives of $\widehat{\beta}$ and $\widehat{\delta}$.*

With some minor modifications by the author, the MATLAB package STBL_CODE was used throughout this thesis to generate sequences of stable random variable, calculate values of the stable density, distribution and quantile functions. To implement stable distribution parameter estimation, a lookup table was generated for ψ_1 and ψ_2 with 184 values of v_α and 86 values of v_β. Interpolation is used to calculate the values of ψ_1 and ψ_2 for those values of v_α and v_β which do not exactly match the lookup table values. Spline interpolation is used in preference to linear interpolation, except for α close to 2, where spline interpolation occasionally performs poorly. All partial derivatives in Figure 3.3.1 were calculated using spline interpolation. Were linear interpolation used to calculate the derivatives in Figure 3.3.1, then the resulting plots would show discontinuities in the first derivative at points where the values of v_α and v_β move between cells in the lookup tables.

3.4 Evaluation of the asymptotic standard deviation of the maximum likelihood estimator of the stable distribution parameters

In this section, we evaluate the asymptotic standard deviation (ASD) of the maximum likelihood (ML) estimator of the stable distribution parameters, ω. We use these evaluations in Section 3.5 to investigate the asymptotic efficiency of McCulloch's quantile-based estimators defined in Section 3.3. Since the ML estimator is the minimum variance asymptotically normal estimator, it provides a convenient benchmark for the asymptotic efficiency of other estimators. Throughout this section, we assume that the ML estimator is taken from an IID sample.

The asymptotic covariance matrix of the ML estimator, V_ω^{ML}, is the inverse of the 4×4

Fisher information matrix $I = (I_{ij})$ where

$$I_{ij} = \int_{-\infty}^{\infty} g_{ij}(x)\, dx \tag{3.4.1}$$

where

$$g_{ij} = \frac{\partial f}{\partial \omega_i} \frac{\partial f}{\partial \omega_j} \frac{1}{f} \tag{3.4.2}$$

and f is the stable distribution density function and

$$(\omega_1, \omega_2, \omega_3, \omega_4) = (\alpha, \beta, \gamma, \delta)\,. \tag{3.4.3}$$

No general formula exists for the ASD of the ML estimator, therefore evaluation must be done numerically. Evaluations previously reported in DuMouchel (1975) and Nolan (2001) both have shortcomings. In DuMouchel (1975), the ASDs near $\alpha = 1$ are unreliable due to the choice of parameterisation. In Nolan (2001), the ASDs at low values of α are higher than the true ASDs due to the narrow interval used when evaluating the Fisher information matrix integrals, Nolan (2014). Concerns about the accuracy of the evaluations listed in Nolan (2001) were raised in Fan (2006). Evaluation of the Fisher information matrix for the special case of symmetric stable distributions was reported in Matsui and Takemura (2006).

To evaluate the Fisher information matrix integrals, we adopt a similar approach to that taken in Nolan (2001), but also include an approximation for the integral tails.

1. A grid of n points $\{x_j\}_{j=1}^n$ on the real line is chosen using the quantiles $\left\{\xi_{p_j}\right\}_{j=1}^n$ of the stable distribution function at quantile levels $\{p_j\}_{j=1}^n$ where

   $$p_1 = 0.00002, p_2 = 0.00004, \ldots, p_n = 0.99998. \tag{3.4.4}$$

   For this choice of quantile levels $n = 49,999$.

2. At each point x_j approximations for the partial derivatives in (3.4.2) are calculated from changes to the density function due to small perturbations of each of the stable distribution parameters. The size of perturbation used was 0.002, except where $\alpha = 1$ which required the use of larger perturbations to avoid numerical problems.

3. The sum over the grid of points $\{x_j\}$ of the product of the partial derivatives and the density function weighted by the distance between the points is used to approximate

the body

$$B_{ij} = \int_{\xi_{p_1}}^{\xi_{p_n}} g_{ij}(x)\, dx \tag{3.4.5}$$

of the each of the integrals I_{ij}.

Inevitably, any choice for the grid of points $\{x_j\}$, does not account for the contribution to

$$I_{ij} = N_{ij} + B_{ij} + P_{ij} \tag{3.4.6}$$

by the negative,

$$N_{ij} = \int_{-\infty}^{\xi_{p_1}} g_{ij}(x)\, dx \tag{3.4.7}$$

and positive,

$$P_{ij} = \int_{\xi_{p_n}}^{\infty} g_{ij}(x)\, dx \tag{3.4.8}$$

tails of the integral. Moreover, a single choice of quantile levels in (3.4.4) does not produce approximations of I_{ij} to the same precision for all stable distributions. For high values of α, more extreme quantile levels need to be included in (3.4.4) than for low values of α.

An approximation of the tails N_{ij} and P_{ij} can be obtained through use of the following tail approximation of the stable density function. The statement $h_1(x) \simeq h_2(x)$ as $x \to a$ means $\lim_{x\to a} h_1(x)/h_2(x) = 1$.

Theorem 3.4.1 ***(Tail density approximation - Nolan (2015), Theorem 1.12)***. *For* $0 < \alpha < 2$ *and* $-1 < \beta < 1$, *the tail properties of the stable distribution density function,* f, *are given by*

$$f(x) \simeq C_{\alpha,\beta,\gamma} x^{-(\alpha+1)} \quad \text{as } x \to \infty \tag{3.4.9}$$

and

$$f(x) \simeq C_{\alpha,-\beta,\gamma} |x|^{-(\alpha+1)} \quad \text{as } x \to -\infty \tag{3.4.10}$$

where

$$C_{\alpha,\beta,\gamma} = \alpha\gamma^{\alpha} \sin\left(\frac{\pi\alpha}{2}\right) \frac{\Gamma(\alpha)}{\pi} (1+\beta). \tag{3.4.11}$$

The author acknowledges Professor Nolan's assistance in drawing the author's attention to Theorem 3.4.1, (Nolan (2014)). The following corollary of Theorem 3.4.1 provides formulae for tail properties of the partial derivatives, $\dfrac{\partial f}{\partial \omega_i}$.

Corollary 3.4.1 *For $0 < \alpha < 2$ and $-1 < \beta < 1$, the tail properties as $x \to \infty$ of the partial derivatives, $\frac{\partial f}{\partial \alpha}, \frac{\partial f}{\partial \beta}, \frac{\partial f}{\partial \gamma}$ and $\frac{\partial f}{\partial \delta}$ are given by*

$$\begin{aligned}
\frac{\partial f}{\partial \alpha} &\simeq x^{-(\alpha+1)} \left(\frac{\partial C_{\alpha,\beta,\gamma}}{\partial \alpha} - C_{\alpha,\beta,\gamma} \ln x \right) && (3.4.12)\\
\frac{\partial f}{\partial \beta} &\simeq \frac{1}{1+\beta} f && (3.4.13)\\
\frac{\partial f}{\partial \gamma} &\simeq \frac{\alpha}{\gamma} f && (3.4.14)\\
\frac{\partial f}{\partial \delta} &\simeq \frac{(\alpha+1)}{x} f && (3.4.15)
\end{aligned}$$

and as $x \to -\infty$ by

$$\begin{aligned}
\frac{\partial f}{\partial \alpha} &\simeq |x|^{-(\alpha+1)} \left(\frac{\partial C_{\alpha,-\beta,\gamma}}{\partial \alpha} - C_{\alpha,-\beta,\gamma} \ln |x| \right) && (3.4.16)\\
\frac{\partial f}{\partial \beta} &\simeq \frac{1}{1-\beta} f && (3.4.17)\\
\frac{\partial f}{\partial \gamma} &\simeq \frac{\alpha}{\gamma} f && (3.4.18)\\
\frac{\partial f}{\partial \delta} &\simeq \frac{(\alpha+1)}{x} f && (3.4.19)
\end{aligned}$$

Note that the derivation of the formula for $\frac{\partial f}{\partial \delta}$ utilises the fact that $\frac{\partial f}{\partial \delta} = -\frac{\partial f}{\partial x}$. The following corollary of Theorem 3.4.1 provides formulae for tail properties of the Fisher information functions g_{ij}.

Corollary 3.4.2 *For $0 < \alpha < 2$ and $-1 < \beta < 1$, the tail properties as $x \to \infty$ and as $x \to -\infty$ of the derivatives of the Fisher information functions, g_{ij}, are given by*

$$\frac{d \ln |g_{ij}|}{d \ln |x|} \simeq b_{ij} \tag{3.4.20}$$

where

$$b_{ij} = \begin{cases} -(\alpha+1) & i,j \neq 4 \\ -(\alpha+2) & i = 4, j \neq 4 \text{ and } i \neq 4, j = 4 \\ -(\alpha+3) & i,j = 4 \end{cases} . \tag{3.4.21}$$

The results of Corollary 3.4.2 suggest an approximation of the Fisher information matrix

functions over the tails of the distribution by the following linear relationships

$$\ln\left(\left|g_{ij}\left(x\right)\right|\right) \approx \ln\left(\left|g_{ij}\left(\xi_{p_n}\right)\right|\right) + b_{ij}\left(\ln|x| - \ln\left|\xi_{p_n}\right|\right), \quad x \in \left[\xi_{p_n}, \infty\right), \tag{3.4.22}$$

$$\ln\left(\left|g_{ij}\left(x\right)\right|\right) \approx \ln\left(\left|g_{ij}\left(\xi_{p_1}\right)\right|\right) + b_{ij}\left(\ln|x| - \ln\left|\xi_{p_1}\right|\right), \quad x \in \left(-\infty, \xi_{p_1}\right], \tag{3.4.23}$$

where p_1, p_n are the quantile levels choosen in (3.4.4).

Use of the approximations for g_{ij} in (3.4.22) and (3.4.23) allow analytic approximations for tails of the Fisher information matrix integrals given by

$$P_{ij} \approx -g_{ij}\left(\xi_{p_n}\right)\frac{\xi_{p_n}}{(b_{ij}+1)} \tag{3.4.24}$$

and

$$N_{ij} \approx -g_{ij}\left(\xi_{p_1}\right)\frac{\left|\xi_{p_1}\right|}{(b_{ij}+1)}. \tag{3.4.25}$$

Let

$$T_{ij} = \frac{N_{ij}+P_{ij}}{I_{ij}} * 100\% \tag{3.4.26}$$

denote the relative percentage size of the Fisher information matrix integral tails compared to the complete integral. For clarity, we write I_{11} as $I_{\alpha\alpha}$, I_{12} as $I_{\alpha\beta}$ etc and similarly for N_{ij}, P_{ij}, T_{ij} and g_{ij}. For the choice of quantile levels $\{p_j\}_{j=1}^{n}$ in (3.4.4), T_{ij} is generally less than 1%, except for $T_{\beta\beta}$ in highly skewed distributions which can approach 5%. For symmetric distributions, $T_{\alpha\alpha}$ is the highest of the T_{ij} percentages. The approximations for $g_{\alpha\alpha}$ in (3.4.22) and (3.4.23) are the least accurate of all the g_{ij} functions due to the inclusion of the $\ln x$ term in (3.4.12). Fortunately, the accuracy of the approximations for $g_{\alpha\alpha}$ increases with α, as does T_{ij}, and therefore the accuracy of the approximations for $I_{\alpha\alpha}$ are largely independent of α for symmetric distributions.

In Table 3.4.1, we list our evaluations of the elements of the Fisher information matrix for selected symmetric stable distributions. Our evaluations are close to those listed in Table 6 of Matsui and Takemura (2006). The greatest differences occur for $I_{\alpha\alpha}$ and all our evaluations of $I_{\alpha\alpha}$ are within 0.2% of those listed in Table 6 of Matsui and Takemura (2006).

Taking the square roots of diagonal elements of the inverse of the Fisher information matrix gives the ASDs, $\sigma_\alpha, \sigma_\beta, \sigma_\gamma$ and σ_δ of the stable distribution parameters α, β, γ and

α	$I_{\delta\delta}$	$I_{\gamma\gamma}$	$I_{\alpha\alpha}$	$I_{\alpha\gamma}$
1.9	0.4727	1.6127	0.8838	-0.3962
	(0.4727)	(1.6127)	(0.8846)	(-0.3963)
1.8	0.4552	1.3898	0.5931	-0.3138
	(0.4552)	(1.3898)	(0.5937)	(-0.3138)
1.7	0.4424	1.2189	0.5023	-0.2692
	(0.4424)	(1.2189)	(0.5028)	(-0.2692)
1.6	0.4334	1.0775	0.4722	-0.2395
	(0.4334)	(1.0775)	(0.4726)	(-0.2396)
1.5	0.4281	0.9556	0.4732	-0.2173
	(0.4281)	(0.9556)	(0.4737)	(-0.2174)
1.4	0.4270	0.8475	0.4968	-0.1992
	(0.4270)	(0.8475)	(0.4973)	(-0.1992)
1.3	0.4310	0.7498	0.5419	-0.1832
	(0.4310)	(0.7498)	(0.5424)	(-0.1832)
1.2	0.4419	0.6603	0.6114	-0.1679
	(0.4419)	(0.6603)	(0.6119)	(-0.1679)
1.1	0.4630	0.5775	0.7127	-0.1523
	(0.4630)	(0.5774)	(0.7132)	(-0.1523)
1.0	0.5000	0.5000	0.8593	-0.1352
	(0.5000)	(0.5000)	(0.8590)	(-0.1352)
0.9	0.5641	0.4272	1.0718	-0.1154
	(0.5641)	(0.4272)	(1.0721)	(-0.1154)
0.8	0.6800	0.3586	1.3928	-0.0914
	(0.6800)	(0.3586)	(1.3928)	(-0.0913)
0.7	0.9094	0.2937	1.8982	-0.0612
	(0.9094)	(0.2937)	(1.8974)	(-0.0611)
0.6	1.4445	0.2325	2.7439	-0.0222
	(1.4446)	(0.2325)	(2.7414)	(-0.0220)
0.5	3.1162	0.1753	4.2819	-0.0292
	(3.1167)	(0.1753)	(4.2748)	(0.0295)

Table 3.4.1: Evaluation of the Fisher information matrix integrals for a symmetric stable distribution, $\beta = 0$. For clarity, we write $I_{44} = I_{\delta\delta}$, $I_{33} = I_{\gamma\gamma}$, $I_{11} = I_{\alpha\alpha}$ and $I_{12} = I_{\alpha\gamma}$. Values in () are from Table 6 of Matsui and Takemura (2006).

δ respectively. In Tables 3.4.2 and 3.4.3, we list our evaluations of $\sigma_\alpha, \sigma_\beta, \sigma_\gamma$ and σ_δ for selected values of α, β. Each of $\sigma_\alpha, \sigma_\beta, \sigma_\gamma$ and σ_δ is symmetric in β about zero.

Remark 3.4.1 *Given the technology available at the time, many of the* ASD *evaluations listed in DuMouchel (1975) are surprisingly close to those in Tables 3.4.2 and 3.4.3. Differences for σ_δ are due to the different stable distribution parameterisation used by DuMouchel. Also, as noted in McCulloch (1986), DuMouchel's values for σ_α and σ_γ for $\beta = 0.5$ at $\alpha = 1.1$ "seem to be out of line".*

Remark 3.4.2 *The* ASD *evaluations in Tables 3.4.2 and 3.4.3 are almost all less than those in the appendix of Nolan (2001). In Nolan (2001) a uniform boundary of $[-50, 50]$ was used to evaluate the body B_{ij} of the Fisher information matrix integrals and the tails N_{ij}, P_{ij} were set to zero, (Nolan (2014)). For higher values of α, this choice of boundary provides reasonable evaluations. However, for lower values of α it results in significantly low evaluations for the Fisher information matrix integrals and consequently significantly high evaluations for the* ASD*s. Differences between these two sets of evaluations are minimal at $\alpha = 1.9$ but increase as α decreases. The relative magnitude of these differences is greater than 4% for σ_α where $\alpha < 1.5$, for $\sigma_\beta, \sigma_\gamma$ where $\alpha < 0.8$ and for σ_δ where $\alpha < 0.6$. The relative magnitude of these differences is largely unaffected by the value of β.*

	σ_α			σ_β		
α	$\beta = 0.0$	$\beta = 0.5$	$\beta = 0.9$	$\beta = 0.0$	$\beta = 0.5$	$\beta = 0.9$
1.9	1.128	1.098	0.992	9.060	8.471	5.448
	(1.134)	(1.104)	(0.998)	(9.082)	(8.504)	(5.517)
1.8	1.384	1.350	1.226	5.671	5.216	3.255
	(1.397)	(1.363)	(1.237)	(5.686)	(5.235)	(3.281)
1.7	1.503	1.468	1.341	4.318	3.912	2.349
	(1.526)	(1.491)	(1.360)	(4.332)	(3.928)	(2.367)
1.6	1.545	1.511	1.387	3.551	3.179	1.842
	(1.583)	(1.548)	(1.417)	(3.565)	(3.194)	(1.857)
1.5	1.536	1.504	1.385	3.037	2.695	1.516
	(1.593)	(1.558)	(1.430)	(3.054)	(2.712)	(1.529)
1.4	1.491	1.460	1.348	2.658	2.344	1.286
	(1.571)	(1.536)	(1.412)	(2.677)	(2.362)	(1.299)
1.3	1.418	1.389	1.284	2.361	2.073	1.116
	(1.526)	(1.492)	(1.371)	(2.383)	(2.094)	(1.129)
1.2	1.326	1.298	1.201	2.117	1.854	0.984
	(1.465)	(1.432)	(1.315)	(2.144)	(1.878)	(0.998)
1.1	1.219	1.193	1.104	1.911	1.671	0.878
	(1.393)	(1.360)	(1.246)	(1.942)	(1.699)	(0.894)
1.0	1.103	1.081	0.999	1.732	1.515	0.791
	(1.309)	(1.240)	(1.166)	(1.780)	(1.560)	(0.817)
0.9	0.980	0.958	0.884	1.575	1.379	0.719
	(1.226)	(1.194)	(1.086)	(1.619)	(1.418)	(0.739)
0.8	0.855	0.834	0.768	1.436	1.260	0.657
	(1.132)	(1.100)	(0.995)	(1.488)	(1.305)	(0.681)
0.7	0.728	0.710	0.652	1.312	1.155	0.604
	(1.029)	(0.997)	(0.897)	(1.374)	(1.210)	(0.633)
0.6	0.604	0.587	0.538	1.203	1.063	0.559
	(0.912)	(0.883)	(0.789)	(1.278)	(1.129)	(0.594)
0.5	0.484	0.470	0.430	1.110	0.985	0.520
	(0.776)	(0.749)	(0.665)	(1.201)	(1.065)	(0.564)

Table 3.4.2: Evaluation of the asymptotic standard deviations, σ_α and σ_β, for selected values of α and β. Values in () are from Nolan (2001).

	σ_γ			σ_δ		
α	$\beta = 0.0$	$\beta = 0.5$	$\beta = 0.9$	$\beta = 0.0$	$\beta = 0.5$	$\beta = 0.9$
1.9	0.835	0.827	0.800	1.682	1.680	1.675
	(0.835)	(0.827)	(0.800)	(1.683)	(1.681)	(1.676)
1.8	0.904	0.891	0.848	1.742	1.739	1.733
	(0.904)	(0.891)	(0.848)	(1.744)	(1.741)	(1.735)
1.7	0.965	0.946	0.889	1.764	1.761	1.753
	(0.965)	(0.947)	(0.889)	(1.766)	(1.763)	(1.757)
1.6	1.023	0.999	0.928	1.764	1.762	1.757
	(1.024)	(1.000)	(0.928)	(1.766)	(1.765)	(1.762)
1.5	1.081	1.052	0.966	1.747	1.749	1.751
	(1.082)	(1.052)	(0.966)	(1.750)	(1.753)	(1.759)
1.4	1.141	1.106	1.006	1.716	1.725	1.741
	(1.142)	(1.107)	(1.006)	(1.718)	(1.730)	(1.750)
1.3	1.206	1.165	1.049	1.670	1.691	1.728
	(1.206)	(1.165)	(1.050)	(1.673)	(1.696)	(1.739)
1.2	1.276	1.229	1.098	1.610	1.647	1.713
	(1.276)	(1.229)	(1.099)	(1.613)	(1.653)	(1.725)
1.1	1.355	1.301	1.155	1.536	1.595	1.698
	(1.355)	(1.302)	(1.160)	(1.538)	(1.601)	(1.708)
1.0	1.445	1.385	1.224	1.445	1.533	1.681
	(1.459)	(1.401)	(1.251)	(1.438)	(1.534)	(1.688)
0.9	1.553	1.486	1.311	1.338	1.463	1.664
	(1.563)	(1.500)	(1.343)	(1.338)	(1.467)	(1.669)
0.8	1.684	1.611	1.424	1.214	1.385	1.645
	(1.717)	(1.652)	(1.496)	(1.214)	(1.388)	(1.647)
0.7	1.852	1.773	1.577	1.073	1.301	1.624
	(1.939)	(1.875)	(1.731)	(1.076)	(1.304)	(1.631)
0.6	2.075	1.992	1.791	0.922	1.214	1.599
	(2.283)	(2.226)	(2.101)	(0.933)	(1.226)	(1.633)
0.5	2.390	2.305	2.106	0.769	1.127	1.567
	(2.840)	(2.797)	(2.687)	(0.799)	(1.175)	(1.676)

Table 3.4.3: Evaluation of the asymptotic standard deviations, σ_γ and σ_δ, for selected values of α and β. Values in () are from Nolan (2001).

3.5 Optimal choice of quantile levels, p_M, for stable distribution parameter estimation

In McCulloch (1986), quantile estimators at quantile levels

$$p_M = (0.05, 0.25, 0.50, 0.75, 0.95) \tag{3.5.1}$$

were used to construct the statistics v_α and v_β (3.3.1 and 3.3.2) which were then used to construct estimators for the stable distribution parameters α and β. Hereafter, we refer to the quantile levels in (3.5.1) as the standard quantile levels for quantile-based stable distribution parameter estimation, or standard quantile levels. The properties of v_α and v_β which allow construction of estimators for α and β are their independence of γ and δ and their invertibility as functions of α and β. These properties are also shared by similar statistics constructed using other choices for the quantile levels in (3.5.1) which could be used to construct other, possibly more efficient, estimators for α and β. The use of quantile levels other than the standard quantile levels was mentioned in Garcia et al. (2011) where the concern was mainly regarding the robustness of quantile-based estimates of α when applied to non stable distributions.

Initially, we restrict our investigation to quantile levels p_M of the form

$$p_M = (p_{M_1}, p_{M_2}, p_{M_3}, p_{M_4}, p_{M_5}) \tag{3.5.2}$$

where

$$0 < p_{M_1} < p_{M_2} < 0.50, \tag{3.5.3}$$

$$p_{M_3} = 0.50, \tag{3.5.4}$$

$$p_{M_4} = 1 - p_{M_2}, \tag{3.5.5}$$

$$p_{M_5} = 1 - p_{M_1}. \tag{3.5.6}$$

Note that the standard quantile levels satisfy the above restrictions where

$$(p_{M_1}, p_{M_2}) = (0.05, 0.25)\,. \tag{3.5.7}$$

We define statistics $v_{\alpha;p_M}$ and $v_{\beta;p_M}$ by

$$v_{\alpha;p_M} = \frac{\xi_{p_{M_5}} - \xi_{p_{M_1}}}{\xi_{p_{M_4}} - \xi_{p_{M_2}}}, \tag{3.5.8}$$

$$v_{\beta;p_M} = \frac{\xi_{p_{M_5}} + \xi_{p_{M_1}} - 2\xi_{p_{M_3}}}{\xi_{p_{M_5}} - \xi_{p_{M_1}}}, \tag{3.5.9}$$

where $\xi_{p_{M_j}}$ is the $p_{M_j}^{th}$ quantile of the distribution. Estimators $\widehat{\alpha}_{p_M}$ and $\widehat{\beta}_{p_M}$ for α and β are constructed by substituting empirical quantile estimators into (3.5.8) and (3.5.9), similar to the method described in Section 3.3.

We wish to identify the optimal value of p_M, which minimises the asymptotic variance of the estimators $\widehat{\alpha}_{p_M}$ and $\widehat{\beta}_{p_M}$. For any single value of p_M, the creation of a lookup table to evaluate the functions ψ_1 and ψ_2 in (3.3.5) and (3.3.6) significantly improves the computation speed in calculating estimators $\widehat{\alpha}_{p_M}$ and $\widehat{\beta}_{p_M}$ and their asymptotic variances. However, the creation of an accurate lookup table is itself a time-consuming exercise and so it is not practical to create such a lookup table for each of the many values of p_M examined whilst searching for the optimal value of p_M. Instead, we use the simplex optimisation method to evaluate the functions ψ_1 and ψ_2, (Nelder and Mead (1965)). In order to avoid numerical problems, it was necessary to evaluate the partial derivatives $D_{\widehat{\omega}}$ from the average over a range of larger perturbations rather than from a single smaller perturbation as described in Section 3.3. Consequently the asymptotic variance evaluations listed in this section are slightly different to those listed elsewhere in this thesis, where lookup tables for the functions ψ_1 and ψ_2 were available.

Let $V_{\widehat{\alpha};p_M}$ and $V_{\widehat{\beta};p_M}$ denote the asymptotic variances of the estimators $\widehat{\alpha}_{p_M}$ and $\widehat{\beta}_{p_M}$ respectively. In general, we see that both $V_{\widehat{\alpha};p_M}$ and $V_{\widehat{\alpha};p_M}$ are more sensitive to changes in p_{M_1} than to changes in p_{M_2}, (Figure 3.5.1).

Optimal values of p_{M_1} and p_{M_2}, subject to (3.5.3) $-$ (3.5.6), for the estimation of α and β are listed in: i) Table 3.5.1 from a symmetric IID process; ii) Table 3.5.2 from an asymmetric IID process and iii) Table 3.5.3 from a symmetric SMA(1) process. The relative asymptotic efficiencies with respect to the ML estimator is given by

$$R_{\alpha;p_M}^{ML} = \frac{V_\alpha^{ML}}{V_{\alpha;p_M}}, \quad R_{\beta;p_M}^{ML} = \frac{V_\beta^{ML}}{V_{\beta;p_M}} \tag{3.5.10}$$

where $V_\alpha^{ML}, V_\beta^{ML}$ are the asymptotic variances of the ML estimators of α, β respectively, taken from Table 3.4.2.

The optimal values of p_{M_1} and p_{M_2} for the estimation of α are not optimal for the estimation of β. Moreover the optimal values of p_{M_1} and p_{M_2} for the estimation of α and β are non-constant functions of both α and β. Note that $V_{\alpha;p_M}$ and $V_{\beta;p_M}$ are symmetrical in β about zero.

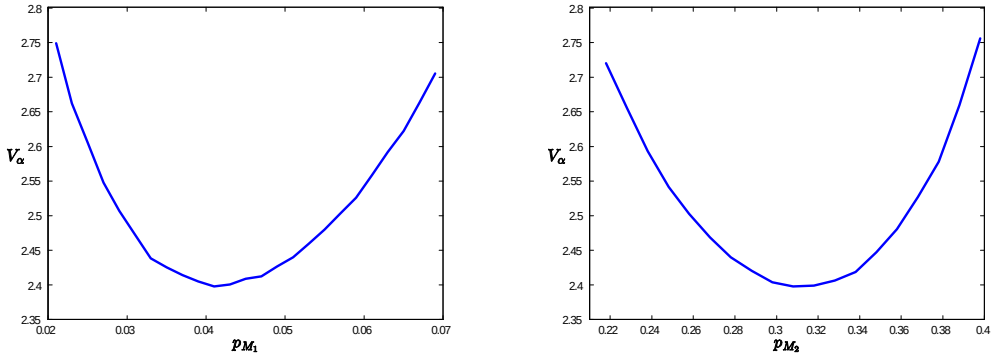
Estimators using the standard p_M are close to optimal over a reasonable range of values

	Estimation of α				Estimation of β			
			$R^{ML}_{\alpha;p_M}$				$R^{ML}_{\beta;p_M}$	
α	(p_{M_1}, p_{M_2})	$V_{\alpha;p_M}$	(i)	(ii)	(p_{M_1}, p_{M_2})	$V_{\beta;p_M}$	(i)	(ii)
0.8	(0.057,0.385)	1.026	0.712	0.504	(0.232,0.396)	2.393	0.861	0.242
0.9	(0.052,0.366)	1.346	0.714	0.557	(0.197,0.377)	2.909	0.853	0.296
1.0	(0.049,0.350)	1.694	0.718	0.608	(0.169,0.350)	3.551	0.845	0.359
1.1	(0.044,0.327)	2.048	0.726	0.654	(0.144,0.324)	4.342	0.841	0.432
1.2	(0.041,0.308)	2.394	0.735	0.688	(0.121,0.295)	5.356	0.837	0.516
1.3	(0.036,0.290)	2.702	0.744	0.703	(0.101,0.268)	6.698	0.832	0.612
1.4	(0.031,0.270)	2.935	0.757	0.683	(0.080,0.238)	8.533	0.828	0.710
1.5	(0.027,0.255)	3.053	0.773	0.612	(0.065,0.231)	11.208	0.823	0.791
1.6	(0.022,0.237)	3.028	0.788	0.490	(0.049,0.207)	15.462	0.815	0.814
1.7	(0.015,0.220)	2.801	0.806	0.340	(0.035,0.178)	23.186	0.804	0.750
1.8	(0.010,0.207)	2.335	0.820	0.201	(0.021,0.167)	40.938	0.786	0.597
1.9	(0.004,0.193)	1.540	0.826	0.091	(0.009,0.142)	109.842	0.747	0.378

Table 3.5.1: Optimal values of p_{M_1} and p_{M_2} for the the estimation of α and β from an IID process where α takes selected values and $\beta = 0.0$. Asymptotic variances of the optimal estimators are denoted by $V_{\alpha;p_M}$ and $V_{\beta;p_M}$ respectively. Asymptotic efficiencies relative to ML estimators are calculated for (i) the optimal value of p_M and (ii) the standard value of $p_M = (0.05, 0.25, 0.50, 0.75, 0.95)$.

	Estimation of α				Estimation of β			
			$R^{ML}_{\alpha;p_M}$				$R^{ML}_{\beta;p_M}$	
α	(p_{M_1}, p_{M_2})	$V_{\alpha;p_M}$	(i)	(ii)	(p_{M_1}, p_{M_2})	$V_{\beta;p_M}$	(i)	(ii)
0.8	(0.065,0.337)	2.205	0.315	0.272	(0.095,0.333)	2.095	0.758	0.463
0.9	(0.060,0.328)	2.669	0.344	0.310	(0.085,0.317)	2.394	0.795	0.545
1.0	(0.053,0.332)	3.100	0.377	0.351	(0.076,0.302)	2.805	0.818	0.636
1.1	(0.048,0.315)	3.480	0.409	0.389	(0.067,0.281)	3.344	0.835	0.733
1.2	(0.041,0.307)	3.800	0.444	0.424	(0.057,0.270)	4.104	0.838	0.812
1.3	(0.039,0.297)	4.051	0.476	0.449	(0.049,0.244)	5.164	0.832	0.830
1.4	(0.033,0.284)	4.201	0.507	0.455	(0.039,0.216)	6.701	0.820	0.761
1.5	(0.025,0.274)	4.222	0.536	0.434	(0.031,0.185)	9.050	0.802	0.631
1.6	(0.021,0.251)	4.058	0.563	0.378	(0.024,0.161)	12.950	0.780	0.484
1.7	(0.016,0.237)	3.660	0.589	0.288	(0.016,0.133)	20.347	0.752	0.344
1.8	(0.011,0.214)	2.953	0.617	0.183	(0.010,0.116)	38.033	0.715	0.218
1.9	(0.006,0.199)	1.814	0.664	0.086	(0.004,0.071)	107.878	0.665	0.115

Table 3.5.2: Optimal values of p_{M_1} and p_{M_2} for the the estimation of α and β from an IID process where α takes selected values and $\beta = 0.5$. Asymptotic variances of the optimal estimators are denoted by $V_{\alpha;p_M}$ and $V_{\beta;p_M}$ respectively. Asymptotic efficiencies relative to ML estimators are calculated for (i) the optimal value of p_M and (ii) the standard value of $p_M = (0.05, 0.25, 0.50, 0.75, 0.95)$.



(a) (b)

Figure 3.5.1: Asymptotic variance V_α from an IID sample where $\alpha = 1.2$ and $\beta = 0.0$. In (a) p_{M_2} is fixed at 0.308. In (b) p_{M_1} is fixed at 0.041.

for α and β. For some values of α and β, estimators using the optimal p_M for the estimation of α provide a poor estimate of β and estimators using the optimal p_M for the estimation of β provide a poor estimate of α. The standard p_M may provide a good compromise, (e.g. $\alpha = 1.4$ and $\beta = 0.0$ in Table 3.5.4). However, for some values of α and β, the optimal p_M for the estimation of α and the optimal p_M for the estimation of β both provide better estimates of both α and β than the standard p_M, (e.g. $\alpha = 1.8$ and $\beta = 0$ in Table 3.5.5).

From Table 3.5.2, it can be seen that estimators using the optimal p_M subject to (3.5.3)−(3.5.6) do not provide particularly efficient estimators of α from asymmetric IID processes. Relaxing those constraints to allow p_M to take on all values such that

$$0 < p_{M_1} < p_{M_2} < p_{M_4} < p_{M_5} < 1, \tag{3.5.11}$$

$$p_{M_1} < p_{M_3} < p_{M_5} \tag{3.5.12}$$

$$p_{M_3} = 0.50, \tag{3.5.13}$$

enables additional improvement in the asymptotic efficiency of estimators using the optimal p_M. Note that the application of restrictions (3.5.11) − (3.5.13) does not preclude values of p_M where $p_{M_2} > p_{M_3}$ or $p_{M_4} < p_{M_3}$. Optimal values of $p_{M_1}, p_{M_2}, p_{M_4}$ and p_{M_5} subject to (3.5.11) − (3.5.13) , for the estimation of α are listed in Table 3.5.6 from an asymmetric IID process.

	Estimation of α			Estimation of β		
		$V_{\alpha;p_M}$			$V_{\beta;p_M}$	
α	(p_{M_1}, p_{M_2})	(i)	(ii)	(p_{M_1}, p_{M_2})	(i)	(ii)
0.8	(0.076,0.391)	1.445	2.111	(0.253,0.397)	3.525	14.081
0.9	(0.068,0.371)	1.840	2.417	(0.217,0.371)	4.184	13.553
1.0	(0.062,0.351)	2.258	2.699	(0.187,0.351)	4.989	13.206
1.1	(0.057,0.329)	2.663	2.955	(0.160,0.337)	5.977	12.988
1.2	(0.050,0.309)	3.045	3.199	(0.135,0.314)	7.230	12.948
1.3	(0.044,0.295)	3.374	3.462	(0.109,0.320)	8.872	13.152
1.4	(0.038,0.266)	3.619	3.805	(0.090,0.291)	11.105	13.836
1.5	(0.030,0.253)	3.730	4.351	(0.071,0.292)	14.430	15.514
1.6	(0.024,0.236)	3.671	5.316	(0.054,0.259)	19.452	19.546
1.7	(0.018,0.235)	3.389	7.017	(0.039,0.243)	28.693	29.662
1.8	(0.010,0.208)	2.822	9.819	(0.024,0.229)	49.777	60.811
1.9	(0.005,0.199)	1.857	14.130	(0.011,0.218)	130.626	233.403

Table 3.5.3: Optimal values of p_{M_1} and p_{M_2} for the the estimation of α and β from an SMA(1) process where α takes selected values and $\beta = 0$, $\theta_1 = 0.4$. Asymptotic variances of the estimators are calculated for (i) the optimal value of p_M and (ii) the standard value of $p_M = (0.05, 0.25, 0.50, 0.75, 0.95)$.

	Optimal α estimation	Optimal β estimation	Standard
(p_{M_1}, p_{M_2})	$(0.031, 0.270)$	$(0.080, 0.238)$	$(0.050, 0.250)$
$V_{\alpha;p_M}$	2.935	4.976	3.255
$V_{\beta;p_M}$	14.207	8.533	9.955

Table 3.5.4: Asymptotic variances $V_{\alpha;p_M}$ and $V_{\beta:p_M}$ for $\alpha = 1.4$ and $\beta = 0$ at selected values of p_{M_1} and p_{M_2}. The standard quantile levels provide good estimates for both α and β.

	Optimal α estimation	Optimal β estimation	Standard
(p_{M_1}, p_{M_2})	$(0.010, 0.207)$	$(0.021, 0.167)$	$(0.050, 0.250)$
$V_{\alpha;p_M}$	2.335	3.246	9.506
$V_{\beta;p_M}$	51.593	40.938	53.903

Table 3.5.5: Asymptotic variances $V_{\alpha;p_M}$ and $V_{\beta:p_M}$ for $\alpha = 1.8$ and $\beta = 0$ at selected values of p_{M_1} and p_{M_2}. The standard quantile levels provide poor estimates for both α and β.

	Estimation of α				
			$R^{MLE}_{\alpha;p_M}$		
α	$(p_{M_1}, p_{M_2}, p_{M_4}, p_{M_5})$	$V_{\alpha;p_M}$	(i)	(ii)	(iii)
0.8	(0.039,0.196,0.438,0.918)	0.978	0.711	0.315	0.272
0.9	(0.035,0.187,0.461,0.929)	1.283	0.715	0.344	0.310
1.0	(0.030,0.186,0.502,0.932)	1.618	0.722	0.377	0.351
1.1	(0.030,0.169,0.512,0.932)	1.958	0.727	0.409	0.389
1.2	(0.028,0.158,0.543,0.942)	2.288	0.737	0.444	0.424
1.3	(0.024,0.152,0.569,0.949)	2.577	0.749	0.476	0.449
1.4	(0.019,0.143,0.598,0.957)	2.788	0.764	0.507	0.455
1.5	(0.015,0.132,0.622,0.961)	2.907	0.778	0.536	0.434
1.6	(0.011,0.127,0.665,0.967)	2.883	0.792	0.563	0.378
1.7	(0.008,0.114,0.690,0.975)	2.674	0.806	0.589	0.288
1.8	(0.004,0.102,0.728,0.985)	2.239	0.814	0.617	0.183
1.9	(0.002,0.087,0.754,0.992)	1.464	0.823	0.664	0.086

Table 3.5.6: Optimal values of p_M subject to (3.5.11) - (3.5.13) which minimise the asymptotic variance of α_{p_M} from an IID process where $\beta = 0.5$. Relative asymptotic efficiencies to ML estimators are included for (i) optimal p_M subject to (3.5.11) - (3.5.13), (ii) optimal p_M subject to (3.5.3) - (3.5.6) and (iii) the standard quantile levels.

Remark 3.5.1 *For α close to 2, the optimal value of p_{M_1} and p_{M_5} approach zero and one respectively. At moderate sample sizes the asymptotic distribution of the empirical quantile estimators at these quantile levels may be quite different to the actual distribution. Consequently the optimal values of p_M at moderate sample sizes may also be quite different from the values calculated using the asymptotic distributions.*

Remark 3.5.2 *Let*

$$p_M^* = \left(p_{M_1}^*, p_{M_2}^*, p_{M_3}^*, p_{M_4}^*, p_{M_5}^*\right) \tag{3.5.14}$$

denote the optimal value of p_M for some α, β under restrictions (3.5.11) – (3.5.13) *. Then the optimal value, q_M^*, of p_M for $\alpha, -\beta$ under restrictions* (3.5.11) – (3.5.13) *is given by*

$$q_M^* = \left(1 - p_{M_5}^*, 1 - p_{M_4}^*, p_{M_3}^*, 1 - p_{M_2}^*, 1 - p_{M_1}^*\right). \tag{3.5.15}$$

Remark 3.5.3 *It is possible to relax the restrictions on p_M by removing* (3.5.13) *. However, results, not included in this thesis, suggest that this relaxation makes: the optimisation process much lengthier; more vulnerable to convergence to a local minima and does not significantly improve the optimal asymptotic variance.*

Remark 3.5.4 *Naturally, it is only possible to calculate estimates using the optimal value for p_M if the true values of α, β and θ are known. This problem can be overcome by making an initial estimate for α, β using the standard value for p_M and then use the initial estimates to make a better choice for p_M and perhaps improve the quality of estimate for α, β. However, an infinite range of choices for p_M requires an infinite number of lookup tables for the evaluation of ψ_1 and ψ_2. In practice, the availability of lookup tables covering a dozen or so choices for p_M would allow improved estimators for many values of α, β and θ.*

A similar optimisation of the asymptotic variances of the quantile-based stable distribution γ parameter estimator is computationally more difficult and not attempted here. However, some optimisation results were achieved under the assumption that α is known and β is known to be zero. Let $\widehat{\gamma}_{p_M}$ denote the estimator for γ using the quantile-based estimator with quantile levels p_M. If α is known and β is known to be zero, then

$$\widehat{\gamma}_{p_M} = \frac{\widehat{\xi}_{p_{M_4}} - \widehat{\xi}_{p_{M_2}}}{\xi^*_{p_{M_4}} - \xi^*_{p_{M_2}}} \tag{3.5.16}$$

α	1.0	1.2	1.4	1.6	1.8	2.0
p_{M_2}	0.250	0.212	0.176	0.141	0.107	0.069
$V_{\widehat{\gamma};p_M}/\gamma^2$	2.467	1.855	1.455	1.171	0.952	0.767

Table 3.5.7: Optimal values of p_{M_2} which minimise the asymptotic variance of $\widehat{\gamma}_{p_M}$ where α is known, β is known to be zero and the sample is IID.

where $\xi^*_{p_{M_j}}$ is the $p_{M_j}th$ quantile of the $S^0_\alpha(0,1,0)$. Note that under these assumptions $\xi^*_{p_{M_j}}$ is also known and the estimator $\widehat{\gamma}_{p_M}$ depends only on $\widehat{\xi}_{p_{M_2}}$ and $\widehat{\xi}_{p_{M_4}}$. Let $V_{\widehat{\gamma}_{p_M}}$ denote the asymptotic variance of $\widehat{\gamma}_{p_M}$. If we further assume that the sample is IID, and that

$$p_{M_2} = 1 - p_{M_4} \tag{3.5.17}$$

then we have the following formula for $V_{\widehat{\gamma};p_M}$

$$V_{\widehat{\gamma};p_M} = \frac{p_{M_2}(2p_{M_2}-1)}{2\xi^2_{p_{M_2}} f^2\left(\xi_{p_{M_2}}\right)} \tag{3.5.18}$$

where f is the density function of the distribution $S^0_\alpha(0,\gamma,0)$. In the special case where $\alpha = 1$, i.e. the Cauchy distribution, there exist the following closed form formulae for $\xi_{p_{M_2}}$ and $f\left(\xi_{p_{M_2}}\right)$

$$\xi_{p_{M_2}} = \gamma \tan\left(\pi\left(p_{M_2} - \frac{1}{2}\right)\right), \tag{3.5.19}$$

$$f\left(\xi_{p_{M_2}}\right) = \frac{1}{\pi\gamma\left(1+\left(\xi_{p_{M_2}}/\gamma\right)^2\right)}. \tag{3.5.20}$$

For $\alpha = 1$, the minimum value of $V_{\widehat{\gamma};p_M}$ is $\gamma^2\pi^2/4$ occuring at $p_{M_2} = 1/4$, which, perhaps by coincidence, was the original choice for p_{M_2} made in McCulloch (1986). Optimal values of p_{M_2} and the associated minimum values of $V_{\widehat{\gamma};p_M}$ for various values of α are listed in Table 3.5.7. Optimal values of p_{M_2} for non-IID, SMA(1) processes are not recorded here but are slightly higher than those listed in Table 3.5.7.

3.6 Estimation of the autocorrelation function of a stable moving average process

For a finite variance process $\{X_t\}$, the autocorrelation function $\rho(h)$ at lag h is defined as

$$\rho(h) = \frac{Cov[X_t, X_{t+h}]}{Var[X_t]}. \tag{3.6.1}$$

However, if $\{X_t\}$ is an SMA(q) process and $\alpha < 2$, i.e. an infinite variance process, then $Var[X_t]$ and possibly $Cov[X_t, X_{t+h}]$ are infinite, and the definition in (3.6.1) is not applicable. An alternative definition for the autocorrelation function $\rho(h)$ at lag h can be made where

$$\rho(h) = \frac{\sum_{j=-\infty}^{\infty} \theta_j \theta_{j+h}}{\sum_{j=-\infty}^{\infty} \theta_j^2}, \tag{3.6.2}$$

which is equivalent to (3.6.1) where $\alpha = 2$.

Definition 3.6.1 ***(Sample Autocorrelation Function)***. *The sample autocorrelation function $\widehat{\rho}(h)$ at lag h from a sample $X_1, \ldots, X_n$ is given by*

$$\widetilde{\rho}(h) = \frac{\sum_{j=1}^{n-h} X_j X_{j+h}}{\sum_{j=1}^{n} X_j^2}. \tag{3.6.3}$$

The mean corrected sample autocorrelation function $\widehat{\rho}(h)$ at lag h from a sample $X_1, \ldots, X_n$ is given by

$$\widehat{\rho}(h) = \frac{\sum_{j=1}^{n-h} \left(X_j - \overline{X}\right)\left(X_{j+h} - \overline{X}\right)}{\sum_{j=1}^{n} \left(X_j - \overline{X}\right)^2}, \tag{3.6.4}$$

where

$$\overline{X} = n^{-1} \sum_{j=1}^{n} X_j \tag{3.6.5}$$

is the sample mean.

The sample autocorrelation function $\widetilde{\rho}(h)$ was shown to be a consistent estimator of $\rho(h)$ in (3.6.2) and its asymptotic distribution was derived in Davis and Resnick (1986).

Theorem 3.6.1 *Let $\{X_t\}$ be an* SMA(q) *process, then for any positive integer h*

$$\left(\frac{n}{\ln(n)}\right)^{1/\alpha} \left(\widetilde{\rho}(1) - \rho(1), \ldots, \widetilde{\rho}(h) - \rho(h)\right)' \xrightarrow{d} (Y_1, \ldots, Y_h)' \tag{3.6.6}$$

where

$$Y_k = \sum_{j=1}^{\infty} \left(\rho\left(k+j\right) + \rho\left(k-j\right) - 2\rho\left(j\right)\rho\left(k\right)\right) S_j / S_0, \quad \textit{for } k = 1, \ldots, h, \tag{3.6.7}$$

$S_0, S_1, \ldots$ *are independent stable random variables such that*

$$S_0 \sim S^0_{\alpha/2}\left(1, C^{-2/\alpha}_{\alpha/2}, 0\right) \tag{3.6.8}$$

$$S_j \sim S^0_{\alpha}\left(0, C^{-1/\alpha}_{\alpha}, 0\right) \tag{3.6.9}$$

and

$$C_\alpha = \begin{cases} \dfrac{(1-\alpha)}{\Gamma\left(2-\alpha\right)\cos\left(\pi\alpha/2\right)} & \textit{if } \alpha \neq 1 \\ \dfrac{2}{\pi} & \textit{if } \alpha = 1 \end{cases} \tag{3.6.10}$$

Corollary 3.6.1 *Let* $\{X_t\}$ *be an* SMA(q) *process where* $1 < \alpha < 2$, *then Theorem 3.6.1 also describes the asymptotic distribution of the mean corrected sample autocorrelation function by replacing* $\widetilde{\rho}$ *with* $\widehat{\rho}$ *in* (3.6.6).

There are no analytic formulae for the distribution of Y_k however, percentiles of the distribution can be calculated numerically or estimated through simulation. In practice, the convergence of $\widehat{\rho}\left(h\right)$ to its asymptotic distribution is very slow. (Adler et al. (1998)) Let $B^{ASY}_{\widehat{\rho}(1)}\left(n, a; \alpha, \beta\right)$ denote the a^{th} percentile of the asymptotic distribution of $\widehat{\rho}\left(1\right)$ from an IID sample of length n and distribution $S_\alpha\left(\beta, 1, 0\right)$. To illustrate the rate of convergence of $\widehat{\rho}\left(1\right)$ to its asymptotic distribution, we calculated $\widehat{\rho}\left(1\right)$ from $10,000$ realisations of IID symmetric stable processes for various values of α and n. For each combination $\left(\alpha, n\right)$, we calculated the percentage of realisations which were outside the interval $\left[B^{ASY}_{\widehat{\rho}(1)}\left(n, 2.5; \alpha, \beta\right), B^{ASY}_{\widehat{\rho}(1)}\left(n, 97.5; \alpha, \beta\right)\right]$, (Table 3.6.1). Note that similar results are reported in Table 4(i) of Adler et al. (1998).

If the finite sample distribution is the same as the asymptotic distribution, then one would expect the values in Table 3.6.1 to be close to 5%. If a value in Table 3.6.1 is less than 5%, it indicates that the spread of the finite sample distribution is smaller than the spread of the asymptotic distribution. The converse applies if the value in Table 3.6.1 is greater than 5%.

The values in Table 3.6.1 move closer to 5% as n increases though, even for $n = 100,000$ it is clear that the finite sample distributions and asymptotic distributions are very different

$\alpha \diagdown n$	100	300	1,000	3,000	10,000	30,000	100,000
1.0	0.00	0.56	1.38	1.72	2.02	2.28	2.92
1.1	0.02	0.59	1.27	1.48	1.96	2.48	2.50
1.2	0.06	0.65	1.35	1.92	2.12	2.18	3.13
1.3	0.19	0.80	1.51	2.09	1.99	2.46	2.71
1.4	0.37	0.86	1.52	1.78	2.12	2.32	2.83
1.5	0.85	1.48	1.73	2.17	2.44	2.74	2.81
1.6	2.13	2.54	2.56	2.84	2.92	3.02	3.14
1.7	5.56	4.76	4.57	4.12	4.47	4.15	4.28
1.8	14.53	11.93	10.05	9.38	8.54	7.85	6.74
1.9	33.23	28.19	26.09	22.71	20.55	19.55	17.02
2.0	4.84	4.69	4.53	4.68	4.86	5.34	4.73

Table 3.6.1: Percentage of $\rho(1)$ estimates from 10,000 realisations of IID stable processes which lie outside the interval $[B^{ASY}_{\widehat{\rho}_1}(n, 2.5; \alpha, \beta), B^{ASY}_{\widehat{\rho}_1}(n, 97.5; \alpha, \beta)]$.

for most values of α. Curiously, for low values of α, we see the percentages in Table 3.6.1 increasing with n towards 5% and for high values of α, we see the the percentages in Table 3.6.1 decreasing with n towards 5%. There appears to be an "optimal" value in the range $(1.7, 1.8)$ where the spread of the finite sample distribution and the asymptotic distribution are fairly close for all n.

In this thesis, we require percentiles of the sample autocovariance function on stable processes of length $n \leq 720$ to use as significance levels for statistical tests. For samples of that size, it is clear from Table 3.6.1 that percentiles derived from the asymptotic distribution are only reliable for a very narrow range of values of α. As an alternative to percentiles from the asymptotic distribution, we can instead use percentiles derived from simulations.

Let $B_{\widehat{\rho}(1)}(n, a; \alpha, \beta)$ denote the a^{th} percentile of the small sample distribution of $\widehat{\rho}(1)$ from an IID sample of length n and distribution $S_\alpha(\beta, 1, 0)$. To estimate values for the $B_{\widehat{\rho}(1)}(n, a; \alpha, \beta)$ percentiles, we calculate $\widehat{\rho}(1)$ from 100,000 realisations of IID stable processes at various values of α and β, (Table 3.6.2). Note that for $\alpha < 1$, the sample autocorrelation function $\widetilde{\rho}(1)$ without mean correction is used instead of $\widehat{\rho}(1)$.

$\alpha \diagdown \beta$	0.0	0.2	0.4	0.6	0.8	1.0
0.8	$\begin{bmatrix} -0.030 \\ 0.030 \end{bmatrix}$	$\begin{bmatrix} -0.028 \\ 0.030 \end{bmatrix}$	$\begin{bmatrix} -0.022 \\ 0.036 \end{bmatrix}$	$\begin{bmatrix} -0.014 \\ 0.045 \end{bmatrix}$	$\begin{bmatrix} -0.004 \\ 0.058 \end{bmatrix}$	$\begin{bmatrix} 0.000 \\ 0.075 \end{bmatrix}$
1.0	$\begin{bmatrix} -0.045 \\ 0.041 \end{bmatrix}$	$\begin{bmatrix} -0.043 \\ 0.042 \end{bmatrix}$	$\begin{bmatrix} -0.041 \\ 0.045 \end{bmatrix}$	$\begin{bmatrix} -0.037 \\ 0.048 \end{bmatrix}$	$\begin{bmatrix} -0.033 \\ 0.053 \end{bmatrix}$	$\begin{bmatrix} -0.029 \\ 0.061 \end{bmatrix}$
1.2	$\begin{bmatrix} -0.055 \\ 0.052 \end{bmatrix}$	$\begin{bmatrix} -0.055 \\ 0.053 \end{bmatrix}$	$\begin{bmatrix} -0.051 \\ 0.056 \end{bmatrix}$	$\begin{bmatrix} -0.048 \\ 0.058 \end{bmatrix}$	$\begin{bmatrix} -0.043 \\ 0.063 \end{bmatrix}$	$\begin{bmatrix} -0.039 \\ 0.071 \end{bmatrix}$
1.4	$\begin{bmatrix} -0.063 \\ 0.060 \end{bmatrix}$	$\begin{bmatrix} -0.062 \\ 0.060 \end{bmatrix}$	$\begin{bmatrix} -0.060 \\ 0.061 \end{bmatrix}$	$\begin{bmatrix} -0.057 \\ 0.065 \end{bmatrix}$	$\begin{bmatrix} -0.053 \\ 0.068 \end{bmatrix}$	$\begin{bmatrix} -0.049 \\ 0.074 \end{bmatrix}$
1.6	$\begin{bmatrix} -0.067 \\ 0.064 \end{bmatrix}$	$\begin{bmatrix} -0.067 \\ 0.065 \end{bmatrix}$	$\begin{bmatrix} -0.065 \\ 0.065 \end{bmatrix}$	$\begin{bmatrix} -0.064 \\ 0.067 \end{bmatrix}$	$\begin{bmatrix} -0.062 \\ 0.070 \end{bmatrix}$	$\begin{bmatrix} -0.059 \\ 0.073 \end{bmatrix}$
1.8	$\begin{bmatrix} -0.071 \\ 0.068 \end{bmatrix}$	$\begin{bmatrix} -0.071 \\ 0.068 \end{bmatrix}$	$\begin{bmatrix} -0.070 \\ 0.069 \end{bmatrix}$	$\begin{bmatrix} -0.070 \\ 0.069 \end{bmatrix}$	$\begin{bmatrix} -0.069 \\ 0.070 \end{bmatrix}$	$\begin{bmatrix} -0.068 \\ 0.071 \end{bmatrix}$
2.0	$\begin{bmatrix} -0.075 \\ 0.072 \end{bmatrix}$	-	-	-	-	-

Table 3.6.2: Estimated values of $B_{\widehat{\rho}_1}(n, 2.5; \alpha, \beta)$ and $B_{\widehat{\rho}_1}(n, 97.5; \alpha, \beta)$ where $n = 720$. Values were estimated from simulations consisting of 100,000 realisations of IID stable processes for each pair (α, β).

Note that for finite samples where $\beta \neq 0$ we observe that the 2.5^{th} and 97.5^{th} simulation percentiles in Table 3.6.2 can have quite different absolute values. Values of $B_{\widehat{\rho}(1)}(n, a; \alpha, \beta)$ are symmetric about 0 in β.

3.7 The Q-Statistic for a stable moving average process

Estimation methods such as SLAD can determine the parameter values of a particular type of model for any given data, but do not determine whether that type of model is appropriate for that data. To assess whether a particular type of model is appropriate for the data, there are a number of what are generically called "goodness of fit" tests. Many goodness of fit tests attempt to verify one of the following hypotheses:

1. That the residuals of the model belong to a particular distribution type (e.g. Gaussian, stable, etc);

2. That the residuals of the model are uncorrelated or independent.

For finite variance processes, a popular test for both uncorrelated residuals is the Q-statistic (Box and Pierce (1970)), which uses the statistic

$$Q_s = n \sum_{h=1}^{s} \widehat{\rho}^2(h), \quad s < n \tag{3.7.1}$$

where $\widehat{\rho}(h)$ is the mean corrected sample autocorrelation function (Definition 3.6.1) at lag h of the residual process $\{\widehat{e}_j\}$. Under the null hypothesis that the residuals $\{\widehat{e}_j\}$ are IID, it can be shown for s sufficiently large, that

$$Q_s \xrightarrow{d} \chi^2_s. \tag{3.7.2}$$

The Q-statistic can also be used for order identification, by identifying the order of the proposed model to be the most parsimonious model which has uncorrelated residuals.

For infinite variance processes we can define a Q-statistic in a similar manner although, the asymptotic distribution given in (3.7.2) no longer applies. Instead we can use the following result from Runde (1997).

Theorem 3.7.1 *Let $X_1, X_2, \ldots$ be a sequence of* IID *random variables which belong to the domain of attraction of a symmetric stable distribution with characteristic exponent α, $1 < \alpha < 2$. Then as $n \to \infty$,*

$$\left(\frac{n}{\ln(n)}\right)^{2/\alpha} \sum_{h=1}^{s} \widehat{\rho}^2(h) \xrightarrow{d} \frac{S_1^2 + \cdots + S_s^2}{S_0^2} \tag{3.7.3}$$

where $S_0, S_1, \ldots$ are independent with distributions given in (3.6.8) *and* (3.6.9).

In Lin and McLeod (2008), it was noted that Theorem 3.7.1 also holds for $0 < \alpha \leq 1$ if we replace in (3.7.3) the mean corrected sample autocorrelation function, $\widehat{\rho}(h)$, with the sample autocorrelation function $\widetilde{\rho}(h)$.

As discussed in Section 3.6, the convergence of autocorrelation estimators to their asymptotic distributions is very slow. Naturally this also affects the convergence of the Q-statistic to the asymptotic distribution as given in Theorem 3.7.1. In this thesis, we want to identify significance levels of the Q-statistic from SMA processes of length $n \leq 720$. As an alternative to significance levels derived from the asymptotic distribution, we can instead use significance levels derived from simulations.

Let $B_{Q_s}(n, a; \alpha, \beta)$ denote the a^{th} percentile of the small sample distribution of Q_s from an IID sample of length n and distribution $S_\alpha(\beta, 1, 0)$. To estimate values for $B_{Q_s}(n, a; \alpha, \beta)$, we calculate Q_s for $s = 20$ from $100,000$ realisations of IID stable processes at various values of α and β, (Table 3.7.1).

For $\alpha \geq 1$ where the mean corrected autocorrelation estimator is used, the simulated value of $B_{Q_s}(720, 95; \alpha, \beta)$ appears to decrease with α and be largely independent of β. For $\alpha < 1$, the simulated value of $B_{Q_s}(720, 95; \alpha, \beta)$ appears to be symmetric in β about 0.

Remark 3.7.1 *In this thesis, we use $B_{Q_s}(n, 95; \alpha, \beta)$ as the 95% significance level for Q-statistic tests on residuals from models for* SMA *processes with sample size n.*

3.8 An alternative method for the order identification of a stable moving average process

In this section, we consider an alternative to the Q-statistic for the order identification of a stable moving average process. For a SMA(q) process, all autocorrelations at lags greater

$\alpha \diagdown \beta$	0.0	0.2	0.4	0.6	0.8	1.0
0.8	39.52	40.49	41.13	43.40	48.63	60.49
1.0	41.35	40.83	41.12	41.07	41.10	41.21
1.2	39.82	39.43	39.73	39.75	39.18	39.19
1.4	36.21	36.49	35.99	36.10	36.07	36.34
1.6	32.81	32.83	32.72	32.57	32.82	32.86
1.8	31.13	30.82	30.91	30.79	30.85	30.82
2.0	31.23	-	-	-	-	-

Table 3.7.1: Estimated values of $B_{Q_s}(n, 95; \alpha, \beta)$ where $n = 720$ and $s = 20$. Values were estimated from simulations consisting of 100,000 realisations of IID stable processes for each pair (α, β).

than q are zero. Thus the order of an SMA(q) process can be identified as the highest lag of the sample autocorrelation function which has a value significantly different from zero. However, there are difficulties involved with such a method for SMA(q) processes in determining whether or not a lag of a sample autocorrelation function is significantly different from zero. We have seen in Section 3.6 evidence of the discrepancies between the asymptotic and finite sample distributions of the sample autocorrelation function. Previous work on this topic includes: Rosenfeld (1976), who used data clipping to reduce the effects of the data heavy tails; Adler et al. (1998), who recommended the use of the asymptotic distribution of $\widehat{\rho}\,(1)$ for a symmetric IID Cauchy process to determine the statistical significance of the sample autocorrelation function and Rosadi (2007), who used the codifference function as an alterative to the autocorrelation function for order identification.

Motivation for the use of the Cauchy asymptotic distribution function to determine the statistical significance of the sample autocorrelation function is found in the simulation results in Adler et al. (1998). In this simulation the order is identified of the SMA(1) process

$$X_t = e_t + \theta_1 e_{t-1} \tag{3.8.1}$$

where $\theta_1 = -0.8$ and $\{e_t\}$ is an IID sequence of symmetric stable random variables with distribution $S_\alpha\,(\beta, 1, 0)$. For a sample size n, define the interval I_{Cauchy} to be

$$I_{Cauchy} = \left[B^{ASY}_{\widehat{\rho}(1)}\,(n, 2.5; 1, 0)\,, B^{ASY}_{\widehat{\rho}(1)}\,(n, 97.5; 1, 0)\right]$$

where $B^{ASY}_{\widehat{\rho}(1)}(n, a; \alpha, \beta)$ denotes the a^{th} percentile of the asymptotic distribution of $\widehat{\rho}(1)$ from a symmetric IID stable process with distribution $S_\alpha(\beta, 1, 0)$. For this simulation, 10,000 samples of length 1,000 are generated, the sample autocorrelation function is calculated at lags 1, 2,... , 10 and the order is determined as follows:

1. If $\widehat{\rho}(1)$ lies inside I_{Cauchy}, then the order is identified as 0. We denote this as a Type L misidentification (identified order is lower than true value).

2. Else if any of $\widehat{\rho}(2), \ldots, \widehat{\rho}(10)$ lie outside I_{Cauchy}, then the order is identified as the maximum lag h for $\widehat{\rho}(h)$ satisfies this condition. We denote this as a Type H misidentification (identified order is higher than true value).

3. Otherwise, the order is identified as 1.

Let P_C denote the percentage of correct identifications, let P_L denote the percentage of Type L misidentifications and P_H the percentage of Type H misidentifications. Naturally,

$$P_C + P_L + P_H = 100\%. \tag{3.8.2}$$

In Table 6 of Adler et al. (1998) it was reported that values of P_C between 83% and 90% were achieved in correctly identifying the sample as being from an SMA(1) process for various values of the stable distribution parameter α. For this simulation, the true value of

$$\rho(1) = -0.488 \tag{3.8.3}$$

and the numerical values of the I_{Cauchy} interval are

$$I_{Cauchy} = [-0.086, 0.086]. \tag{3.8.4}$$

Let P_h denote the probability that $\widehat{\rho}(h)$ lies outside I_{Cauchy}, that is

$$P_h = P\{\widehat{\rho}(h) \in I_{Cauchy}\}, \tag{3.8.5}$$

then $P_1 > 99.9\%$ for all $\alpha \in [1, 2]$ and therefore P_L is less than 0.1%. The probabilities $P_2, \ldots, P_{10}$ are between 98% and 99% depending on α and h and therefore P_H is approximately $\sum_{h=2}^{10} 100\% - P_h$ which is between 10% and 17% as reported in Table 6 of Adler et al. (1998). In fact, for this simulation even better results could have been achieved by replacing I_{Cauchy} with, for example, the interval $[-0.150, 0.150]$ .This change of interval would

not greatly effect P_1 but would reduce the probabilities $P_2, \ldots, P_{10}$ and increase the values of P_C to greater than 96% for all α when running the same simulation.

A more useful test of the methods efficacy, might involve the choice of processes which have a greater probability of a Type L misidentification. Such processes would have lower absolute values of the true autocorrelation function. We show that use of the interval I_{Cauchy} is much less successful when applied to such processes and demonstrate that a more efficient method is to use an interval derived from the finite sample distribution of I_{Finite} defined by

$$I_{Finite} = \left[B_{\widehat{\rho}(1)}\left(n, a_1; \alpha, \beta\right), B_{\widehat{\rho}(1)}\left(n, a_2; \alpha, \beta\right)\right] \tag{3.8.6}$$

for appropriate choices of a_1, a_2 where $B_{\widehat{\rho}(1)}\left(n, a; \alpha, \beta\right)$ denotes the a^{th} percentile of the small sample distribution of $\widehat{\rho}\left(1\right)$ from an IID sample of length n and distribution $S_\alpha\left(\beta, 1, 0\right)$.

As discussed in Section 3.6, the convergence of the distribution of $\widehat{\rho}\left(h\right)$ to its asymptotic distribution is very slow. For moderate sample sizes, significance levels derived from the asymptotic distribution of $\widehat{\rho}\left(h\right)$ do not accurately reflect the behaviour of the finite sample distribution (Table 3.6.1). However, it is possible to use simulations to estimate percentiles for the finite sample distribution of $\widehat{\rho}\left(h\right)$. In Table 3.6.2 estimates of $B_{\widehat{\rho}(1)}\left(720, 2.5; \alpha, \beta\right)$ and $B_{\widehat{\rho}(1)}\left(720, 97.5; \alpha, \beta\right)$ are listed for various values of $\left(\alpha, \beta\right)$.

To demonstrate the efficacy of the I_{Finite} interval in comparison with the I_{Cauchy} interval, we ran a simulation of $10,000$ realisations of various symmetric SMA(1) processes, each of length 720. For each choice of θ_1 and α, the percentages P_C and P_L are recorded in Table 3.8.1 using both the I_{Cauchy} interval and the I_{Finite} interval at various choices for the percentile levels a_1 amd a_2 The results in Table 3.8.1 assume that α and β are known.

The misidentification percentages in Table 3.8.1 using the I_{Cauchy} are generally fairly good where $\theta_1 = 0.15$ and 0.20 but quite poor where $\theta_1 = 0.10$. To help to understand why that is so, consider the case where $\theta_1 = 0.10, \alpha = 1$ and $n = 720$. Then the true value of

$$\rho\left(1\right) = 0.099 \tag{3.8.7}$$

and the numerical values of the I_{Cauchy} interval are

$$I_{Cauchy} = \left[-0.113, 0.113\right]. \tag{3.8.8}$$

Note the difference between the numerical values in (3.8.4) and (3.8.8) is due to the difference in the sample sizes. Recall from Theorem 3.6.1, that the values of $B^{ASY}_{\widehat{\rho}(1)}\left(n, a; \alpha, \beta\right)$ decrease

		I_{Cauchy}	I_{Finite}				
θ_1	α		(i)	(ii)	(iii)	(iv)	(v)
0.10	1.0	9.06	54.65	65.67	75.27	79.66	10.88
		(89.59)	(0.89)	(1.34)	(2.42)	(5.18)	(87.51)
0.10	1.4	20.84	50.50	61.66	67.87	64.75	25.14
		(77.16)	(3.56)	(6.49)	(12.88)	(23.70)	(72.36)
0.10	1.8	31.12	42.03	51.70	56.70	56.46	50.76
		(69.27)	(12.78)	(18.23)	(25.56)	(32.04)	(44.15)
0.15	1.0	88.01	54.69	66.26	76.60	83.03	88.16
		(3.82)	(0.36)	(0.53)	(0.78)	(1.07)	(3.48)
0.15	1.4	84.22	51.88	64.72	76.14	82.33	85.14
		(8.86)	(0.58)	(1.01)	(1.63)	(2.50)	(7.45)
0.15	1.8	81.47	47.28	62.05	73.61	79.90	85.19
		(14.46)	(0.71)	(1.23)	(2.18)	(3.28)	(5.85)
0.20	1.0	90.93	53.75	65.61	76.93	83.61	90.80
		(0.82)	(0.11)	(0.15)	(0.24)	(0.34)	(0.78)
0.20	1.4	90.65	50.90	64.31	76.33	83.46	90.38
		(1.19)	(0.19)	(0.29)	(0.35)	(0.47)	(1.09)
0.20	1.8	95.39	47.32	62.20	74.35	82.38	90.51
		(1.26)	(0.06)	(0.11)	(0.18)	(0.20)	(0.35)

Table 3.8.1: Order identification of various SMA(1) processes from a sample of length 720. Reported in this table are the percentage of correct identifications, P_C and the percentage of Type L misidentifications, P_L, in ( ). Identification is conducted using the I_{Cauchy} interval and the I_{Finite} interval where (a_1, a_2) is (i) $(4.0\%, 96.0\%)$, (ii) $(2.5\%, 97.5\%)$, (iii) $(1.5\%, 98.5\%)$, (iv) $(1.0\%, 99.0\%)$ and (v) $(0.5\%, 99.5\%)$,

in absolute value with the sample size n, at the rate $\left(\frac{\log n}{n}\right)^{1/\alpha}$. Clearly, the probability that $\widehat{\rho}(1)$ lies within I_{Cauchy} and consequently P_L is very high. In this case the numerical values of the I_{Finite} interval are

$$I_{Finite} = [-0.080, 0.074], \quad \text{for } (a_1, a_2) = (1.0\%, 99.0\%). \tag{3.8.9}$$

The reduction in size of the I_{Finite} interval compared to the I_{Cauchy} interval is sufficient to reduce P_L from 89.59% to 5.18%.

For all processes, the misidentification percentages in Table 3.8.1 using I_{Finite} with $(a_1, a_2) = (0.5\%, 99.5\%)$ are similar to those using I_{Cauchy}. The percentage of correct identifications using I_{Finite} with $(a_1, a_2) = (1.0\%, 99.0\%)$ is higher than using I_{Finite} with values of (a_1, a_2) closer to 50%. Where the values of (a_1, a_2) closer to 50%, the percentage of correct identifcations is reduced by an increase in the percentage of Type H misidentifications. The optimal value for the I_{Finite} percentile levels is dependent on θ_1, α, to a lesser extent on β and also on the sample size. However, a choice of $(a_1, a_2) = (1.0\%, 99.0\%)$ appears to work reasonably well for many of the SMA(1) processes tested here.

The correct identification of a SMA(2) process requires that both $\widehat{\rho}(1)$ and $\widehat{\rho}(2)$ lie outside the designated interval, thus increasing the vulnerability to Type L misidentifications. A simulation of 10,000 realisations of various symmetric SMA(2) processes, each of length 720, was used to compare the order identification performance using the I_{Cauchy} interval and the I_{Finite} interval where $(a_1, a_2) = (1.0\%, 99.0\%)$, (Table 3.8.2). The results in Table 3.8.2 assume that α and β are known.

Of the ten SMA(2) processes reported in Table 3.8.2, the value for P_C obtained using the I_{Finite} interval is substantially higher than that obtained using I_{Cauchy} interval for four processes, marginally higher for three processes and marginally lower for three processes. Comparisons are not greatly affected by α. Neither method handles well the SMA(2) process where $(\theta_1, \theta_2) = (-0.1, -0.4)$ and the true autocorrelation function is given by

$$\rho(1) = -0.051, \tag{3.8.10}$$

$$\rho(2) = -0.342. \tag{3.8.11}$$

Remark 3.8.1 *In this thesis, we use the interval*

$$I_{Finite} = \left[B_{\widehat{\rho}(1)}(n, 1.0; \alpha, \beta), B_{\widehat{\rho}(1)}(n, 99.0; \alpha, \beta)\right] \tag{3.8.12}$$

	$\alpha = 1.0$		$\alpha = 1.4$		$\alpha = 1.8$	
(θ_1, θ_2)	I_{Cauchy}	I_{Finite}	I_{Cauchy}	I_{Finite}	I_{Cauchy}	I_{Finite}
$(0.4, 0.2)$	89.71 (2.93)	83.24 (1.01)	87.62 (5.23)	82.00 (1.71)	86.61 (8.46)	76.48 (1.86)
$(-0.4, -0.2)$	89.61 (2.83)	83.72 (1.34)	87.86 (4.71)	83.18 (1.62)	88.41 (7.52)	79.51 (2.08)
$(-0.3, 0.1)$	6.76 (91.93)	76.85 (10.93)	15.51 (82.82)	53.28 (37.77)	24.17 (74.17)	46.48 (42.05)
$(0.3, -0.1)$	6.46 (92.46)	72.41 (16.66)	15.22 (83.13)	45.90 (46.48)	25.67 (73.50)	45.75 (45.22)
$(0.2, -0.3)$	80.43 (13.46)	84.03 (1.57)	68.29 (26.41)	80.84 (4.82)	63.80 (33.55)	76.27 (6.92)
$(-0.2, 0.3)$	90.95 (1.10)	84.08 (0.61)	91.79 (0.90)	83.94 (0.54)	95.18 (0.50)	80.16 (0.17)
$(0.1, 0.4)$	70.19 (24.43)	81.48 (3.77)	56.60 (38.52)	73.16 (13.35)	51.60 (45.70)	61.82 (22.29)
$(-0.1, -0.4)$	0.56 (99.17)	2.21 (96.47)	0.83 (98.97)	2.68 (95.95)	1.06 (98.78)	5.67 (91.72)
$(0.2, 0.1)$	7.83 (90.57)	79.53 (7.95)	18.13 (79.91)	59.92 (30.35)	27.91 (71.09)	53.25 (36.67)
$(-0.2, -0.1)$	7.30 (91.52)	77.29 (11.07)	16.76 (81.51)	53.39 (37.63)	26.50 (72.49)	49.12 (40..87)

Table 3.8.2: Order identification of various SMA(2) processes from a sample of length 720. Reported in this table are the percentage of correct identifications, P_C, and the percentage of Type L misidentifications, P_L, in ( ). Identification is conducted using the I_{Cauchy} interval and the I_{Finite} interval where (a_1, a_2) is $(1.0\%, 99.0\%)$.

for all order identification tests on SMA *processes. We refer to an order identification test using this interval as an extended Adler test.*

3.9 Simulation

In this section we present the results of simulations which demonstrate the use of the methods described in this chapter for the estimation of the parameters of a stable moving average process. For selected set of values α, β, and θ_1 a simulation is run where 2,000 realisations of an SMA(1) process each of length 720 are generated . The parameters $\gamma_e = 2$ and $\delta_e = 1$ are fixed for all simulations. All stable distribution parameter estimations are done using the standard quantile levels,

$$p_M = (0.05, 0.25, 0.50, 0.75, 0.95)\,. \tag{3.9.1}$$

Estimates for the parameters α, β, γ, δ and θ_1 are calculated for each realisation. The mean and variance of these estimates across all realisations of a particular simulation are then compared with the true parameter values and the asymptotic variance of the estimators. The results for α, β, γ, δ and θ_1 are reported respectively in Tables 3.9.1, 3.9.2, 3.9.3, 3.9.4 and 3.9.5.

In each case the mean value of the estimator across all realisations is within one standard deviation of the true parameter value and is generally much closer than that. The normalised variance (i.e. the variance multiplied by the sample size) across all realisations is reasonably close to the asymptotic variance. Note that for $\widehat{\theta}_1$ we are unable to calculate a true asymptotic variance and instead substitute the mean of the estimated asymptotic variance across all realisations.

The normalised variance of $\widehat{\alpha}$ where $\alpha = 1.8$ appears to be slightly less than the asymptotic variance. This is due to the truncation of all $\widehat{\alpha}$ estimates into the range $(0, 2]$. A similar effect is seen with $\widehat{\beta}$ estimates where $\alpha = 1.8$ and $\beta_e = 0.5$. Estimates of $\widehat{\beta}$ where $\alpha = 1.8$ are the least precise. This is to be expected as the asymptotic variance of $\widehat{\beta}$ increases to ∞ as α increases to 2.

For each of the selected SMA(1) processes and for each of the estimators $\widehat{\alpha}$, $\widehat{\beta}$, $\widehat{\gamma}$ and $\widehat{\delta}$ the asymptotic variance of the estimator is higher for $\theta_1 = 0.2$ than for $\theta_1 = 0.0$ and higher still for $\theta_1 = 0.4$. The effect of increases in θ_1 on the asymptotic variance of the estimators

		$\theta_1 = 0.0$		$\theta_1 = 0.2$		$\theta_1 = 0.4$	
α	β	(i)	(ii)	(i)	(ii)	(i)	(ii)
1.2	0.0	1.195 (0.061)	2.682 [2.555]	1.195 (0.063)	2.891 [2.740]	1.194 (0.068)	3.322 [3.200]
1.2	0.2	1.196 (0.062)	2.746 [2.833]	1.199 (0.067)	3.224 [3.040]	1.198 (0.072)	3.702 [3.568]
1.2	0.5	1.202 (0.076)	4.134 [3.975]	1.201 (0.076)	4.204 [4.257]	1.200 (0.084)	5.074 [5.058]
1.5	0.0	1.503 (0.076)	4.105 [3.852]	1.504 (0.079)	4.472 [3.984]	1.502 (0.080)	4.604 [4.348]
1.5	0.2	1.504 (0.078)	4.346 [4.076]	1.506 (0.081)	4.693 [4.217]	1.505 (0.087)	5.466 [4.611]
1.5	0.5	1.505 (0.087)	5.384 [5.207]	1.506 (0.091)	5.907 [5.384]	1.506 (0.094)	6.375 [5.919]
1.8	0.0	1.808 (0.108)	8.389 [9.471]	1.809 (0.107)	8.160 [9.544]	1.809 (0.109)	8.495 [9.783]
1.8	0.2	1.810 (0.109)	8.515 [9.536]	1.808 (0.110)	8.678 [9.611]	1.806 (0.109)	8.614 [9.853]
1.8	0.5	1.809 (0.106)	8.083 [9.902]	1.808 (0.108)	8.353 [9.981]	1.809 (0.110)	8.669 [10.24]

Table 3.9.1: Simulation results for the estimation of α from selected SMA(1) processes. Reported in this table for each process are (i) the mean and standard deviation, in ( ), of $\widehat{\alpha}$ across all realisations and (ii) the variance of $\widehat{\alpha}$ across all realisations multiplied by the sample size T and the true asymptotic variance, in [ ].

$\widehat{\alpha}$, $\widehat{\beta}$, $\widehat{\gamma}$ and $\widehat{\delta}$ appears to decrease as α increases and is more significant for $\widehat{\gamma}$ and $\widehat{\delta}$ than for $\widehat{\alpha}$ and $\widehat{\beta}$. From additional simulation results not included in this thesis, we observe that the asymptotic variance of $\widehat{\gamma}$ appears symmetric in θ_1 about zero however, that does not appear to be the case for $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\delta}$ where more complicated relationships exist between the asymptotic variances and the parameter values.

These simulations provide some confidence that the estimators discussed in this appendix, are an unbiased method for the estimation of stable distribution parameters from an SMA(1) process and that the asymptotic variance provides a reasonable approximation for estimator variance at sample sizes equal to 720.

At present we have no method for calculating the asymptotic covariance of $\widehat{\omega}$ and $\widehat{\theta}$ however, we can use simulations to estimate its value. Each simulation contains 2000 realisations of an SMA(1) process of length 720 where θ_1 takes values from the set $\{-0.9, -0.8, \ldots, 0.9\}$, $\omega = (\alpha, 0, 2, 1)$ and α takes values from the set $\{1.2, 1.5, 1.8\}$. For each realisation of a process, the estimates $\widehat{\omega}$ and $\widehat{\theta}_1$ are calculated. The correlation of $\widehat{\omega}$ and $\widehat{\theta}_1$ is calculated across all realisations of each process, (Figure 3.9.1).

The estimators $\widehat{\beta}$ and $\widehat{\delta}$ appear uncorrelated with $\widehat{\theta}_1$. There appears to be a small correlation between $\widehat{\alpha}$ and $\widehat{\theta}_1$ and a slightly larger correlation between $\widehat{\gamma}$ and $\widehat{\theta}_1$. In both cases the correlation appears to increase with α and be most significant at $\theta_1 \approx \pm 0.4$ where the absolute values of the simulated correlation $\widehat{\gamma}$ and $\widehat{\theta}_1$ approaches 0.2 and between $\widehat{\gamma}$ and $\widehat{\theta}_1$ approaches 0.4. Additional simulations were run with non-zero values of β. The correlations from those simulations were not significantly different to those reported here from simulations with zero values of β and are not reported here.

		$\theta_1 = 0.0$		$\theta_1 = 0.2$		$\theta_1 = 0.4$	
α	β	(i)	(ii)	(i)	(ii)	(i)	(ii)
1.2	0.0	0.000 (0.105)	7.961 [8.684]	0.002 (0.121)	10.46 [10.91]	0.001 (0.131)	12.30 [12.94]
1.2	0.2	0.193 (0.101)	7.408 [7.677]	0.196 (0.113)	9.230 [9.657]	0.194 (0.124)	11.03 [11.44]
1.2	0.5	0.498 (0.082)	4.866 [4.223]	0.495 (0.089)	5.751 [5.199]	0.494 (0.098)	6.852 [6.046]
1.5	0.0	-0.001 (0.127)	11.60 [11.67]	0.001 (0.140)	14.10 [13.27]	0.000 (0.151)	16.44 [15.52]
1.5	0.2	0.202 (0.140)	14.06 [11.16]	0.204 (0.139)	13.88 [12.60]	0.203 (0.151)	16.33 [14.63]
1.5	0.5	0.525 (0.147)	15.44 [11.44]	0.522 (0.152)	16.70 [12.24]	0.520 (0.156)	17.57 [13.45]
1.8	0.0	−0.008 (0.367)	97.20 [53.62]	0.008 (0.369)	97.76 [55.74]	0.000 (0.375)	101.5 [60.59]
1.8	0.2	0.223 (0.367)	97.40 [63.83]	0.202 (0.381)	104.4 [65.85]	0.186 (0.380)	104.1 [70.58]
1.8	0.5	0.488 (0.334)	80.51 [118.5]	0.488 (0.347)	86.49 [120.0]	0.484 (0.353)	89.85 [124.1]

Table 3.9.2: Simulation results for the estimation of β from selected SMA(1) processes. Reported in this table for each process are (i) the mean and standard deviation, in ( ), of $\widehat{\beta}$ across all realisations and (ii) the variance of $\widehat{\beta}$ across all realisations multiplied by the sample size T and the true asymptotic variance, in [ ]. Note that $\beta = \beta_e$ for all processes reported in this table.

		$\theta_1 = 0.0$			$\theta_1 = 0.4$		
α	β	γ	(i)	(ii)	γ	(i)	(ii)
1.2	0.0	2.000	1.991 (0.104)	7.751 [7.983]	2.541	2.523 (0.153)	16.82 [16.32]
1.2	0.2	2.000	1.992 (0.112)	9.000 [8.648]	2.541	2.535 (0.156)	17.58 [17.52]
1.2	0.5	2.000	2.001 (0.123)	10.83 [10.80]	2.541	2.543 (0.171)	21.14 [21.15]
1.5	0.0	2.000	2.000 (0.097)	6.822 [6.553]	2.325	2.322 (0.120)	10.32 [10.45]
1.5	0.2	2.000	1.997 (0.094)	6.290 [6.633]	2.325	2.318 (0.120)	10.37 [10.54]
1.5	0.5	2.000	2.000 (0.098)	6.838 [6.877]	2.325	2.321 (0.125)	11.24 [10.77]
1.8	0.0	2.000	2.000 (0.093)	6.162 [6.272]	2.205	2.204 (0.108)	8.407 [8.482]
1.8	0.2	2.000	2.000 (0.094)	6.417 [6.267]	2.205	2.200 (0.108)	8.378 [8.473]
1.8	0.5	2.000	2.003 (0.093)	6.146 [6.227]	2.205	2.210 (0.111)	8.799 [8.408]

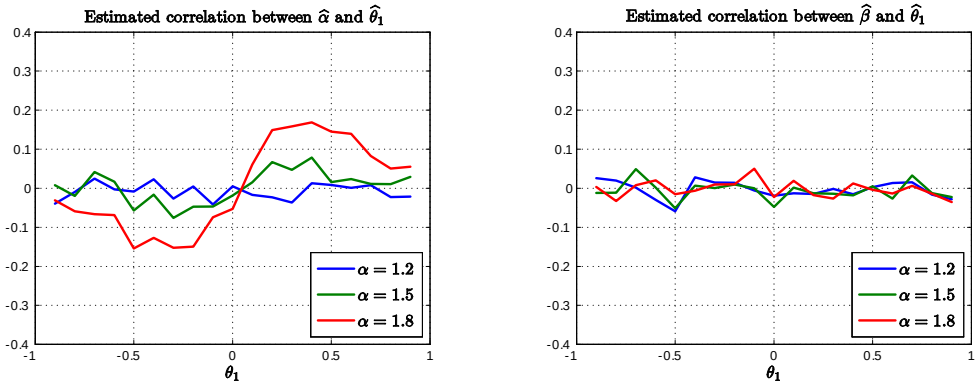
Table 3.9.3: Simulation results for the estimation of γ from selected SMA(1) processes. Reported in this table for each process are (i) the mean and standard deviation, in ( ), of $\widehat{\gamma}$ across all realisations and (ii) the variance of $\widehat{\gamma}$ across all realisations multiplied by the sample size T and the true asymptotic variance, in [ ].

		$\theta_1 = 0.0$			$\theta_1 = 0.4$		
α	β	δ	(i)	(ii)	δ	(i)	(ii)
1.2	0.0	1.000	0.998 (0.130)	12.15 [13.05]	1.400	1.398 (0.194)	27.04 [28.32]
1.2	0.2	1.000	1.003 (0.134)	12.95 [13.19]	1.559	1.558 (0.197)	27.97 [28.62]
1.2	0.5	1.000	1.001 (0.139)	14.00 [14.11]	1.798	1.812 (0.208)	31.15 [30.41]
1.5	0.0	1.000	1.001 (0.143)	14.73 [15.49]	1.400	1.397 (0.200)	28.71 [27.46]
1.5	0.2	1.000	1.005 (0.148)	15.73 [15.61]	1.495	1.500 (0.198)	28.22 [27.68]
1.5	0.5	1.000	0.993 (0.150)	16.27 [16.58]	1.638	1.633 (0.201)	29.09 [29.24]
1.8	0.0	1.000	0.997 (0.160)	18.49 [19.50]	1.400	1.395 (0.202)	29.35 [30.46]
1.8	0.2	1.000	1.010 (0.159)	18.15 [19.61]	1.439	1.454 (0.199)	28.62 [30.59]
1.8	0.5	1.000	1.018 (0.160)	18.35 [20.10]	1.497	1.512 (0.203)	29.80 [31.24]

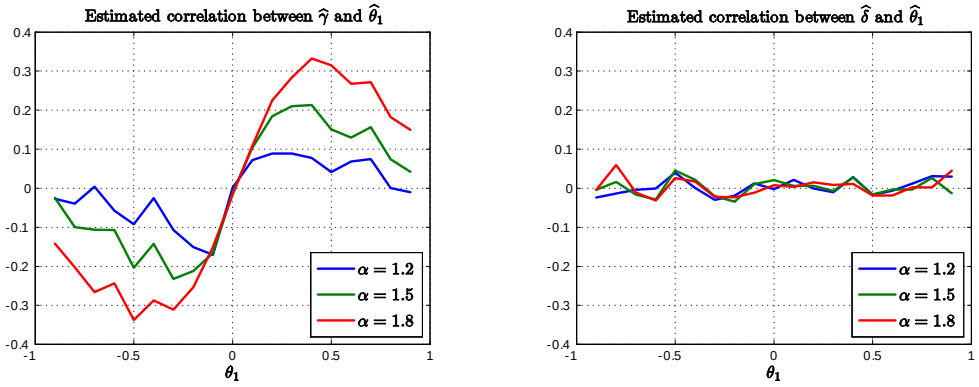
Table 3.9.4: Simulation results for the estimation of δ from selected SMA(1) processes. Reported in this table for each process are (i) the mean and standard deviation, in ( ), of $\widehat{\delta}$ across all realisations and (ii) the variance of $\widehat{\delta}$ across all realisations multiplied by the sample size T and the true asymptotic variance, in [ ].

		$\theta_1 = 0.0$		$\theta_1 = 0.2$		$\theta_1 = 0.4$	
α	β	(i)	(ii)	(i)	(ii)	(i)	(ii)
1.2	0.0	0.000 (0.016)	0.175 [0.164]	0.200 (0.014)	0.145 [0.132]	0.400 (0.012)	0.100 [0.088]
1.2	0.2	0.000 (0.016)	0.181 [0.172]	0.200 (0.014)	0.143 [0.137]	0.400 (0.011)	0.092 [0.090]
1.2	0.5	0.000 (0.017)	0.219 [0.193]	0.200 (0.015)	0.168 [0.153]	0.400 (0.014)	0.132 [0.101]
1.5	0.0	-0.001 (0.028)	0.582 [0.594]	0.200 (0.027)	0.530 [0.530]	0.400 (0.024)	0.401 [0.402]
1.5	0.2	0.000 (0.029)	0.624 [0.596]	0.199 (0.028)	0.551 [0.542]	0.400 (0.024)	0.412 [0.412]
1.5	0.5	0.000 (0.029)	0.607 [0.618]	0.200 (0.027)	0.529 [0.554]	0.400 (0.024)	0.410 [0.419]
1.8	0.0	-0.003 (0.041)	1.226 [1.236]	0.197 (0.040)	1.160 [1.160]	0.398 (0.037)	0.983 [0.962]
1.8	0.2	-0.002 (0.041)	1.202 [1.238]	0.199 (0.040)	1.162 [1.161]	0.399 (0.036)	0.960 [0.965]
1.8	0.5	-0.001 (0.043)	1.304 [1.241]	0.200 (0.041)	1.185 [1.161]	0.400 (0.037)	0.971 [0.966]

Table 3.9.5: Simulation results for the estimation of θ_1 from selected SMA(1) processes. Reported in this table for each process are (i) the mean and standard deviation, in ( ), of $\widehat{\theta}_1$ across all realisations and (ii) the variance of $\widehat{\theta}_1$ across all realisations multiplied by the sample size T and the mean estimated asymptotic variance across all realisations, in [ ].



(a) (b)



(c) (d)

Figure 3.9.1: Estimated correlation of the $\widehat{\omega}$ and $\widehat{\theta}_1$ estimators from selected symmetric SMA(1) processes: (a) between $\widehat{\alpha}$ and $\widehat{\theta}_1$, (b) between $\widehat{\beta}$ and $\widehat{\theta}_1$, (c) between $\widehat{\gamma}$ and $\widehat{\theta}_1$ and (d) between $\widehat{\delta}$ and $\widehat{\theta}_1$.

Chapter 4

Estimation of the log quantile difference of the temporal aggregation an SMA(q) process

4.1 Introduction

In this chapter, three different estimators are proposed for the estimation of the log quantile difference, $\zeta_p^{(r)}$, of the temporal aggregation of an SMA(q) process. These estimators are referred to as the non-parametric, parametric and analytic log quantile difference estimators. Each of these estimators is asymptotically normal. Asymptotic variances are derived for each of these estimators based on the results presented in Chapters 2 and 3.

The parametric LQD estimator, $\widetilde{\zeta}_p^{(r)}$, can be applied to a sample taken from any SMA(q) process. The estimators

$$\widehat{\omega}^{(1)} = \left(\widehat{\alpha}, \widehat{\beta}^{(1)}, \widehat{\gamma}^{(1)}\right)' \tag{4.1.1}$$

and

$$\widehat{\theta} = \left(\widehat{\theta}_1, \ldots, \widehat{\theta}_q\right)' \tag{4.1.2}$$

defined in Section 3.3 and Appendix C.3 respectively are combined in Section 4.2 to provide an estimator

$$\widehat{\omega}^{(r)} = \left(\widehat{\alpha}, \widehat{\beta}^{(r)}, \widehat{\gamma}^{(r)}\right)' \tag{4.1.3}$$

for the stable distribution parameters of the temporal aggregation of an SMA(q) process. In

Section 4.3, the stable distribution parameter estimates $\widehat{\omega}^{(r)}$ are then used to construct the parametric LQD estimator.

The asymptotic variance of parametric LQD estimator, $\widetilde{\zeta}_p^{(r)}$, depends on the asymptotic correlation of the estimators $\widehat{\omega}^{(1)}$ and $\widehat{\theta}$. No formula exists for the asymptotic correlation of the estimators $\widehat{\omega}^{(1)}$ and $\widehat{\theta}$. Simulations reported in Section 3.9 suggest that a modest correlation exists between $\widehat{\theta}$ and both $\widehat{\alpha}$ and $\widehat{\gamma}^{(1)}$ and a negligible correlation between $\widehat{\theta}$ and $\widehat{\beta}^{(1)}$. In Section 4.4, an investigation is conducted into the effect of this correlation on the asymptotic variance of parametric LQD estimator.

In Section 4.5, the non-parametric LQD estimator, $\widehat{\zeta}_p^{(r)}$, is defined as a function of the empirical quantile estimators. The non-parametric LQD estimator can be applied to a sample from any ϕ - mixing process and does not require that the process has a stable distribution. The non-parametric LQD estimator has computational advantages over the parametric LQD estimator however, it suffers from a loss of efficiency at aggregation levels greater than one due to the effective reduction in sample size. At aggregation level equal to one, there are situations where the non-parametric LQD estimator is more efficient that the parametric LQD estimator. In Section 4.6, an investigation is conducted into the relationship between the relative asymptotic efficiency of these estimators and the quantile levels p_M used in the construction of $\widehat{\omega}^{(1)}$.

The analytic LQD estimator, $\check{\zeta}_p^{(r)}$, can be applied to a sample taken from SMA(q) processes which satisfy the conditions for Corollary 2.2.2, i.e. those SMA(q) processes where $\beta^{(r)} = \beta$ and $r \geq q$. Where those conditions are satisfied, Corollary 2.2.2 provides a formula for the analytic LQD estimator in terms of $\widehat{\alpha}, \widehat{\theta}$ and the non-parametric LQD estimator at aggregation level one, $\widehat{\zeta}_p^{(1)}$. Where applicable, the analytic and parametric LQD estimators return the same estimates however, the analytic LQD estimator has minor computational advantages over the parametric LQD estimator. The analytic LQD estimator does not require the calculation of the estimators $\widehat{\beta}^{(1)}$ and $\widehat{\gamma}^{(1)}$ and a formula is provided in Section 4.8 for the asymptotic covariance of $\widehat{\alpha}$ and $\widehat{\zeta}_p^{(1)}$.

The results of simulations are reported in Section 4.9 which demonstrate the use of the methods described in this chapter.

4.2 Estimation of the stable distribution parameters of the temporal aggregation of an SMA(q) process

Let $\{X_t\}$ be the SMA(q) process (Definition 1.2.2),

$$X_t = \sum_{j=0}^{q} \theta_j e_{t-j} \tag{4.2.1}$$

where $\theta_0 = 1$ and $\{e_t\}$ is an IID sequence of stable random variables such that

$$e_t \sim S_\alpha^0 \left(\beta_e, \gamma_e, \delta_e\right) \tag{4.2.2}$$

using the S^0 parameterisation of stable random variables in Nolan (1998). Let θ denote the $q+1$ dimensional vector of moving average parameters

$$\theta = \left(\theta_0, \ldots, \theta_q\right)'. \tag{4.2.3}$$

Let $\left\{S_t^{(r)}\right\}$ denote the temporal aggregation (Definition 1.1.3) of $\{X_t\}$ at aggregation level r

$$S_t^{(r)} = \sum_{i=0}^{r-1} X_{t-i}. \tag{4.2.4}$$

Note that

$$S_t^{(1)} = X_t. \tag{4.2.5}$$

As shown in Theorem 2.2.1, the base process $\{X_t\}$ and aggregated process $\left\{S_t^{(r)}\right\}$ both have a stable distribution. We denote the stable distribution parameters of the aggregated process by

$$S_t^{(r)} \sim S_\alpha^0 \left(\beta^{(r)}, \gamma^{(r)}, \delta^{(r)}\right). \tag{4.2.6}$$

Let $\zeta_p^{(r)}$ denote the log quantile difference (Definition 1.2.1) of the aggregated process, $\left\{S_t^{(r)}\right\}$ at quantile level

$$p = \left(p_1, p_2\right). \tag{4.2.7}$$

In Chapter 2, formulae were derived for the calculation of the log quantile difference of the temporal aggregation of a stable moving average process. In Chapter 3, asymptotically normal estimators for θ and

$$\omega^{(1)} = \left(\alpha, \beta^{(1)}, \gamma^{(1)}\right)' \tag{4.2.8}$$

were defined. For the remainder of this thesis we omit the parameter $\delta^{(1)}$ from the vector of parameters $\omega^{(1)}$ in (4.2.8) since the log quantile difference $\zeta_p^{(r)}$ is independent of $\delta^{(1)}$. In this section we define an estimator for the stable distribution parameters

$$\omega^{(r)} = \left(\alpha, \beta^{(r)}, \gamma^{(r)}\right)' \tag{4.2.9}$$

of the temporal aggregation of an SMA(q) process derived from the estimators for the stable distribution parameters, $\omega^{(1)}$, and an estimator for the moving average parameters, θ. This estimator is based on the results from Theorem 2.2.1. It is convenient to use the symbol $\varkappa$ to denote the vector of parameters

$$\varkappa = \left(\alpha, \beta^{(1)}, \gamma^{(1)}, \theta_1, \ldots, \theta_q\right)'. \tag{4.2.10}$$

Let $C_1^{(r)}(\varkappa), C_2(\varkappa), C_3^{(r)}(\varkappa)$ and $C_4(\varkappa)$ denote the following

$$C_1^{(r)}(\varkappa) = \sum_{j=0}^{r+q-1} \left|c_j^{(r)}\right|^\alpha, \tag{4.2.11}$$

$$C_2(\varkappa) = \sum_{j=0}^{q} |\theta_j|^\alpha, \tag{4.2.12}$$

$$C_3^{(r)}(\varkappa) = \sum_{j=0}^{r+q-1} \operatorname{sign}\left(c_j^{(r)}\right) \left|c_j^{(r)}\right|^\alpha, \tag{4.2.13}$$

$$C_4(\varkappa) = \sum_{j=0}^{q} \operatorname{sign}(\theta_j) |\theta_j|^\alpha \tag{4.2.14}$$

where $c_j^{(r)}$ are the moving average parameters of the aggregated process given in Theorem 2.2.1,

$$c_j^{(r)} = \sum_{k=\max(j-r+1,0)}^{\min(j,q)} \theta_k. \tag{4.2.15}$$

Where appropriate, we may simply refer to $C_1^{(r)}(\varkappa), C_2(\varkappa), C_3^{(r)}(\varkappa)$ and $C_4(\varkappa)$ as $C_1^{(r)}, C_2, C_3^{(r)}$ and C_4 respectively.

Using this notation we can modify equations (2.2.4) and (2.2.5) to obtain

$$\gamma^{(r)} = \left(\frac{C_1^{(r)}}{C_2}\right)^{1/\alpha} \gamma^{(1)} \tag{4.2.16}$$

and if $C_4 \neq 0$

$$\beta^{(r)} = \frac{C_2 C_3^{(r)}}{C_1^{(r)} C_4} \beta^{(1)}. \tag{4.2.17}$$

Note that $C_1^{(r)}$ and C_2 are always strictly positive.

Let $\widehat{\varkappa}$ denote the estimator of $\varkappa$,

$$\widehat{\varkappa} = \left(\widehat{\omega}^{(1)\prime}, \widehat{\theta}\right)' \tag{4.2.18}$$

where

$$\widehat{\omega}^{(1)} = \left(\widehat{\alpha}, \widehat{\beta}^{(1)}, \widehat{\gamma}^{(1)}\right)' \tag{4.2.19}$$

is the quantile-based stable distribution parameter estimator of $\omega^{(1)}$ defined in Section 3.3 and

$$\widehat{\theta} = \left(\widehat{\theta}_1, \ldots, \widehat{\theta}_q\right) \tag{4.2.20}$$

is the SLAD moving average parameter estimator of θ defined in Appendix C.3. Let $\varkappa_0$ denote the true value of the parameter $\varkappa$. The asymptotic distribution of $\widehat{\varkappa}$ is given in the following theorem.

Theorem 4.2.1 *Let $\varkappa_0$ denote the true value of the parameter $\varkappa$. As the the sample size $T \to \infty$ the asymptotic distribution of $\widehat{\varkappa}$ is given by*

$$\sqrt{T}\left(\widehat{\varkappa} - \varkappa_0\right) \xrightarrow{d} N\left(0, V_{\widehat{\varkappa}}\right) \tag{4.2.21}$$

where

$$V_{\widehat{\varkappa}} = \begin{pmatrix} V_{\widehat{\omega}^{(1)}} & V'_{\widehat{\omega}^{(1)},\widehat{\theta}} \\ V_{\widehat{\omega}^{(1)},\widehat{\theta}} & V_{\widehat{\theta}} \end{pmatrix}, \tag{4.2.22}$$

and $V_{\widehat{\omega}^{(1)}}$ is given in (3.3.20) *and $V_{\widehat{\theta}}$ is given in* (C.3.9).

Proof. This theorem follows from the asymptotic distributions of $\widehat{\omega}^{(1)}$ and $\widehat{\theta}$ established in Theorems 3.3.1 and C.3.1 respectively. ■

There is no formula for $V_{\widehat{\omega}^{(1)},\widehat{\theta}}$ though, the simulation results in Section 3.9, suggest that the true values are close to zero. An investigation on the effects of various assumptions for $V_{\widehat{\omega}^{(1)},\widehat{\theta}}$ on the asymptotic variance of log quantile difference estimators is provided in Section 4.4

We are now able to define our estimator $\widehat{\omega}^{(r)}$ of $\omega^{(r)}$.

Definition 4.2.1 ***(Aggregated Stable Distribution Parameter Estimator - $\widehat{\omega}^{(r)}$).*** *If $C_4\left(\widehat{\varkappa}\right) \neq 0$, then the estimator $\widehat{\omega}^{(r)}$ of the stable distribution parameters of the aggregated process $\left\{S_t^{(r)}\right\}$ is given by*

$$\widehat{\omega}^{(r)} = \left(\widehat{\alpha}, \widehat{\beta}^{(r)}, \widehat{\gamma}^{(r)}\right)' \tag{4.2.23}$$

where

$$\widehat{\beta}^{(r)} = \frac{C_2\left(\widehat{\varkappa}\right) C_3^{(r)}\left(\widehat{\varkappa}\right)}{C_1^{(r)}\left(\widehat{\varkappa}\right) C_4\left(\widehat{\varkappa}\right)} \widehat{\beta}^{(1)} \tag{4.2.24}$$

and

$$\widehat{\gamma}^{(r)} = \left(\frac{C_1^{(r)}\left(\widehat{\varkappa}\right)}{C_2\left(\widehat{\varkappa}\right)} \right)^{1/\alpha} \widehat{\gamma}^{(1)}. \tag{4.2.25}$$

In the following theorem, we derive the asymptotic distribution of $\widehat{\omega}^{(r)}$.

Theorem 4.2.2 *Let $\varkappa_0$ denote the true value of the parameters $\varkappa$ and let $\omega_0^{(r)}$ denote the true value of the parameters $\omega^{(r)}$. If*

$$C_3^{(r)}\left(\varkappa_0\right) \neq 0, \tag{4.2.26}$$

$$C_4\left(\varkappa_0\right) \neq 0 \tag{4.2.27}$$

and

$$\begin{aligned} &\text{either } \alpha \geq 1 \\ &\text{or } c_j^{(r)} \neq 0, \quad \text{for } j = 0, \ldots, r+q-1 \text{ and } \theta_j \neq 0, \quad \text{for } j = 1, \ldots, q, \end{aligned} \tag{4.2.28}$$

then the asymptotic distribution of $\widehat{\omega}^{(r)}$ is given by

$$\sqrt{T}\left(\widehat{\omega}^{(r)} - \omega_0^{(r)}\right) \xrightarrow{d} N\left(0, V_{\widehat{\omega}^{(r)}}\right) \tag{4.2.29}$$

where

$$V_{\widehat{\omega}^{(r)}} = D_{\widehat{\omega}^{(r)}} V_{\widehat{\varkappa}} D'_{\widehat{\omega}^{(r)}} \tag{4.2.30}$$

and

$$D_{\widehat{\omega}^{(r)}} = \left[\frac{\partial \widehat{\omega}^{(r)}}{\partial \widehat{\varkappa}_j} \mid_{\widehat{\varkappa}=\varkappa_0} \right]_{j=1:3+q} \tag{4.2.31}$$

and T is the sample size.

Proof. See Appendix A.2. ■

Remark 4.2.1 *The requirement in Theorem 4.2.2 that $C_3^{(r)}\left(\varkappa_0\right) \neq 0$ is a sufficient condition to ensure that the estimator $\widehat{\beta}^{(r)}$ has a non-zero differential at $\widehat{\varkappa} = \varkappa_0$. A necessary condition has not been established. Without this requirement, the estimator $\widehat{\beta}^{(r)}$ still has a differential, but it may be zero. For the application in this thesis, the estimated values of $C_3^{(r)}$ are always positive. For example, it can be shown that $C_3^{(r)}\left(\varkappa_0\right) > 0$, where $r \geq 2$ for all invertible* SMA(2) *processes.*

Remark 4.2.2 *The requirement in Theorem 4.2.2 that* $C_4(\varkappa_0) \neq 0$, *is a necessary condition to ensure that the estimator* $\widehat{\beta}^{(r)}$ *is well defined for* $r > 1$. *If* $C_4(\varkappa_0) = 0$, *then* $\beta^{(1)} = 0$ *even though* β *and* $\beta^{(r)}$ *may not equal zero. For the application in this thesis, the estimated values of* C_4 *are always positive. A graphical illustration of the values of* C_4 *for various invertible* SMA(2) *processes is provided in Figure 4.2.1. For an invertible* SMA(2) *process with* $\alpha \geq 1$ *and* $\theta_1 > -1$, $C_4(\varkappa)$ *is always greater than zero.*

Remark 4.2.3 *The requirement in Theorem 4.2.2 that either* $\alpha \geq 1$ *or* $c_j^{(r)} \neq 0$, *for* $j = 0, \ldots, r+q-1$ *and* $\theta_j \neq 0$, *for* $j = 1, \ldots, q$ *is a necessary condition to ensure that the partial derivatives* $\dfrac{\partial \widehat{\beta}^{(r)}}{\partial \widehat{\theta}_k}$ *and* $\dfrac{\partial \widehat{\gamma}^{(r)}}{\partial \widehat{\theta}_k}$ *to exist. For the application in this thesis, the values of the estimator* $\widehat{\alpha}$ *are greater than one and therefore this requirement is satisfied.*

4.3 Parametric estimation of the log quantile difference of the temporal aggregation of an SMA(q) process

In this section we define an estimator for the log quantile difference of the temporal aggregation of an SMA(q) process using the stable distribution parameter estimators, $\widehat{\omega}^{(r)}$, of the aggregated process defined in Section 4.2. Let $p = (p_1, p_2)$, then for $j = 1, 2$, let $\xi_{p_j}^{(r)}$ denote the $p_j th$ quantile of the stable distribution function $S_\alpha^0\left(\beta^{(r)}, \gamma^{(r)}, \delta^{(r)}\right)$. Define the location adjusted quantile

$$\xi_{p_j}^{*(r)} = \xi_{p_j}^{(r)} - \delta^{(r)}. \tag{4.3.1}$$

Note that $\xi_{p_j}^{*(r)}$ is the $p_j th$ quantile of the stable distribution function $S_\alpha^0\left(\beta^{(r)}, \gamma^{(r)}, 0\right)$, (Lemma E.1.1).

Definition 4.3.1 ***(Parametric Quantile Estimator -*** $\widetilde{\xi}_{p_j}^{*(r)}$***).*** *The estimator* $\widetilde{\xi}_{p_j}^{*(r)}$ *of* $\xi_{p_j}^{*(r)}$ *is defined to be the* $p_j th$ *quantile of the distribution function* $S_{\widehat{\alpha}}^0\left(\widehat{\beta}^{(r)}, \widehat{\gamma}^{(r)}, 0\right)$ *where*

$$\widehat{\omega}^{(r)} = \left(\widehat{\alpha}, \widehat{\beta}^{(r)}, \widehat{\gamma}^{(r)}\right)' \tag{4.3.2}$$

is as defined in Definition 4.2.1.

We use the tilde superscript for $\widetilde{\xi}_{p_j}^{*(r)}$ to distinguish the parametric quantile estimator from the empirical quantile estimator, $\widehat{\xi}_{p_j}$, (Definition 3.2.2). Use of the location adjusted

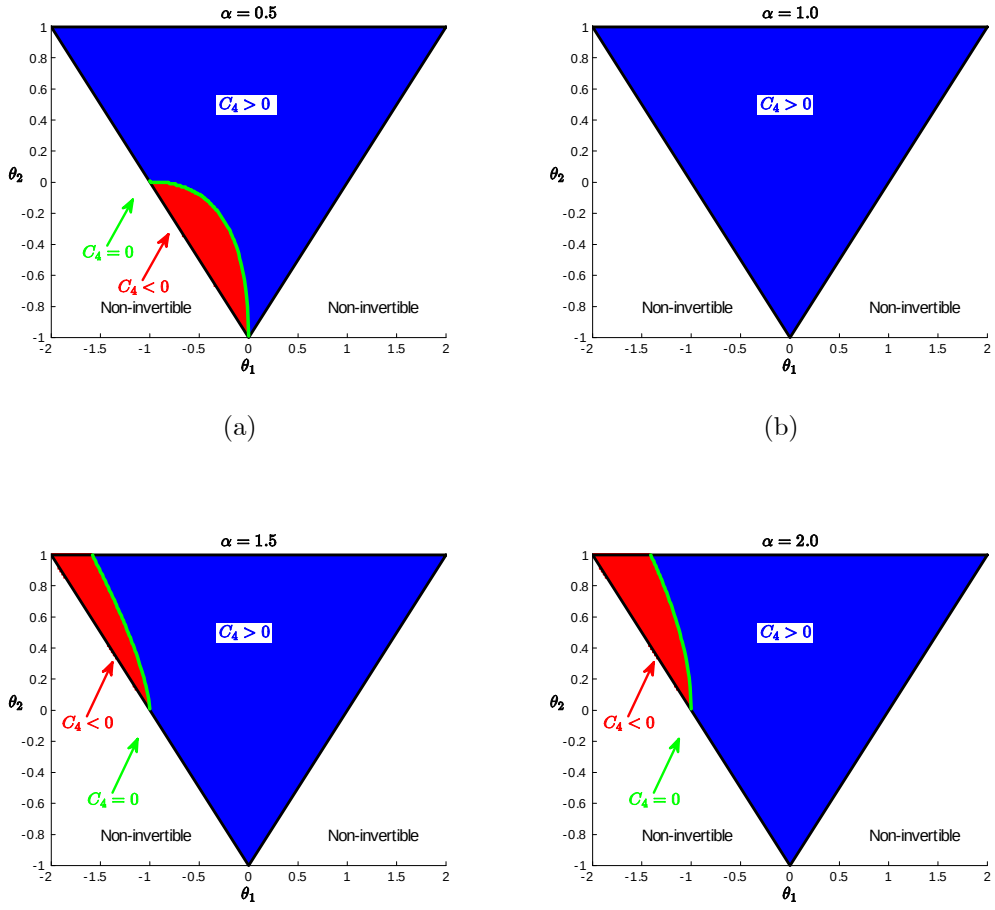


Figure 4.2.1: A graphical display of the categorisation of the invertibility region of the SMA(2) process with respect to the sign of $C_4(\varkappa)$ into positive (blue), zero (green) and negative (red) regions for (a) $\alpha = 0.5$, (b) $\alpha = 1.0$, (c) $\alpha = 1.5$ and (d) $\alpha = 2.0$.

quantile, $\xi_{p_j}^{*(r)}$, rather than the quantile, $\xi_{p_j}^{(r)}$, means that it is not necessary to include estimators for $\delta^{(r)}$ in the vector of stable distribution parameters $\widehat{\omega}^{(r)}$.

Definition 4.3.2 ***(Parametric Log Quantile Difference Estimator -*** $\widetilde{\zeta}_p^{(r)}$***).*** *The estimator* $\widetilde{\zeta}_p^{(r)}$ *of the* LQD *at quantile levels* $p = (p_1, p_2)$, *where* $0 < p_1 < p_2 < 1$, *of the temporal aggregation,* $\left\{S_t^{(r)}\right\}$, *of an* SMA(q) *process* $\{X_t\}$ *is given by*

$$\widetilde{\zeta}_p^{(r)} = \ln\left(\widetilde{\xi}_{p_2}^{*(r)} - \widetilde{\xi}_{p_1}^{*(r)}\right) \tag{4.3.3}$$

where $\widetilde{\xi}_{p_j}^{*(r)}$ *is the parametric quantile estimators (Definition 4.3.1) of* $\left\{S_t^{(r)}\right\}$ *at quantile level* p_j, $j = 1, 2$.

Let $\omega_0^{(r)}$ denote the true value of the parameters $\omega^{(r)}$. It was shown in Theorem 4.2.2, that under certain conditions as the sample size $T \to \infty$

$$\sqrt{T}\left(\widehat{\omega}^{(r)} - \omega_0^{(r)}\right) \xrightarrow{d} N\left(0, V_{\widehat{\omega}^{(r)}}\right). \tag{4.3.4}$$

In the following theorem, we derive the asymptotic distribution of $\widetilde{\zeta}_p^{(r)}$.

Theorem 4.3.1 *If the properties of* $\left\{S_t^{(r)}\right\}$ *are such that the assumptions of Theorem 4.2.2 are satisfied, then as the sample size* $T \to \infty$

$$\sqrt{T}\left(\widetilde{\zeta}_p^{(r)} - \zeta_{p;0}^{(r)}\right) \xrightarrow{d} N\left(0, V_{\widetilde{\zeta}_p^{(r)}}\right) \tag{4.3.5}$$

where

$$V_{\widetilde{\zeta}_p^{(r)}} = D_{\widetilde{\zeta}_p^{(r)}} D_{\widetilde{\xi}_p^{*(r)}} V_{\widehat{\omega}^{(r)}} D'_{\widetilde{\xi}_p^{*(r)}} D'_{\widetilde{\zeta}_p^{(r)}} \tag{4.3.6}$$

and

$$D_{\widetilde{\zeta}_p^{(r)}} = \left[\frac{\partial \widetilde{\zeta}_p^{(r)}}{\partial \widetilde{\xi}_{p_i}^{*(r)}} \Big|_{\widetilde{\xi}_{p_i}^{*(r)} = \xi_{p_i}^{*(r)}}\right]_{i=1,2} \tag{4.3.7}$$

and

$$D_{\widetilde{\xi}_p^{*(r)}} = \left[\frac{\partial \widetilde{\xi}_{p_i}^{*(r)}}{\partial \widehat{\omega}_j^{(r)}} \Big|_{\widehat{\omega}^{(r)} = \omega_0^{(r)}}\right]_{i=1,2;j=1,\ldots,3}. \tag{4.3.8}$$

Proof. See Appendix A.3. ■

No closed form formula exists for the partial derivative matrix $D_{\widetilde{\xi}_p^{*(r)}}$ and consequently it needs to be calculated numerically. To approximate $\dfrac{\partial \widetilde{\xi}_{p_1}^{*(r)}}{\partial \widehat{\alpha}}$, we use the difference equation

$$\frac{\partial \widetilde{\xi}_{p_1}^{*(r)}}{\partial \widehat{\alpha}} \approx \frac{\widetilde{\xi}_{p_1}^{*(r)} \big|_{\widehat{\alpha} = \alpha + \Delta\alpha} - \widetilde{\xi}_{p_1}^{*(r)} \big|_{\widehat{\alpha} = \alpha - \Delta\alpha}}{2\Delta\alpha} \tag{4.3.9}$$

for a suitable choice of $\Delta\alpha$ with similar equations for the other elements of $D_{\widetilde{\xi}_p^{*(r)}}$.

4.4 Effect of the correlation between $\widehat{\omega}^{(1)}$ and $\widehat{\theta}$ on the asymptotic variance of the parametric log quantile difference estimator $\widetilde{\zeta}_p^{(r)}$

Combining the results of Theorems 4.2.2 and 4.3.1, we have under certain conditions that the asymptotic distribution of the parametric log quantile difference estimator, $\widetilde{\zeta}_p^{(r)}$ is given by

$$\sqrt{T}\left(\widetilde{\zeta}_p^{(r)} - \zeta_p^{(r)}\right) \xrightarrow{d} N\left(0, V_{\widetilde{\zeta}_p^{(r)}}\right) \tag{4.4.1}$$

where

$$V_{\widetilde{\zeta}_p^{(r)}} = D_{\widetilde{\zeta}_p^{(r)}} D_{\widetilde{\xi}_p^{*(r)}} D_{\widehat{\omega}^{(r)}} V_{\widehat{\varkappa}} D'_{\widehat{\omega}^{(r)}} D'_{\widetilde{\xi}_p^{*(r)}} D'_{\widetilde{\zeta}_p^{(r)}} \tag{4.4.2}$$

and $V_{\widehat{\varkappa}}$ is the asymptotic covariance matrix of

$$\widehat{\varkappa} = \left(\widehat{\omega}^{(1)\prime}, \widehat{\theta}'\right)' \tag{4.4.3}$$

where

$$\widehat{\omega}^{(1)} = \left(\widehat{\alpha}, \widehat{\beta}^{(1)}, \widehat{\gamma}^{(1)}\right)', \tag{4.4.4}$$

$$\widehat{\theta} = \left(\widehat{\theta}_1, \ldots, \widehat{\theta}_q\right) \tag{4.4.5}$$

and the partial derivative matrices $D_{\widetilde{\zeta}_p^{(r)}}, D_{\widetilde{\xi}_p^{*(r)}}$ and $D_{\widehat{\omega}^{(r)}}$ are defined in (4.3.7), (4.3.8) and (4.2.31) respectively. Let the components of $V_{\widehat{\varkappa}}$ be denoted by

$$V_{\widehat{\varkappa}} = \begin{pmatrix} V_{\widehat{\omega}^{(1)}} & V'_{\widehat{\omega}^{(1)},\widehat{\theta}} \\ V_{\widehat{\omega}^{(1)},\widehat{\theta}} & V_{\widehat{\theta}} \end{pmatrix} \tag{4.4.6}$$

where

$$V_{\widehat{\omega}^{(1)}} = \begin{pmatrix} V_{\widehat{\alpha}} & V_{\widehat{\alpha},\widehat{\beta}^{(1)}} & V_{\widehat{\alpha},\widehat{\gamma}^{(1)}} \\ V_{\widehat{\alpha},\widehat{\beta}^{(1)}} & V_{\widehat{\beta}^{(1)}} & V_{\widehat{\beta}^{(1)},\widehat{\gamma}^{(1)}} \\ V_{\widehat{\alpha},\widehat{\gamma}^{(1)}} & V_{\widehat{\beta}^{(1)}.\widehat{\gamma}^{(1)}} & V_{\widehat{\gamma}^{(1)}} \end{pmatrix} \tag{4.4.7}$$

$$V_{\widehat{\theta}} = \begin{pmatrix} V_{\widehat{\theta}_1} & \cdots & V_{\widehat{\theta}_1,\widehat{\theta}_q} \\ \vdots & \ddots & \vdots \\ V_{\widehat{\theta}_1,\widehat{\theta}_q} & \cdots & V_{\widehat{\theta}_q} \end{pmatrix} \tag{4.4.8}$$

$$V_{\widehat{\omega}^{(1)},\widehat{\theta}} = \begin{pmatrix} V_{\widehat{\alpha},\widehat{\theta}_1} & V_{\widehat{\beta}^{(1)},\widehat{\theta}_1} & V_{\widehat{\gamma}^{(1)},\widehat{\theta}_1} \\ \vdots & \vdots & \vdots \\ V_{\widehat{\alpha},\widehat{\theta}_q} & V_{\widehat{\beta}^{(1)},\widehat{\theta}_q} & V_{\widehat{\gamma}^{(1)},\widehat{\theta}_q} \end{pmatrix}. \tag{4.4.9}$$

Formulae for $V_{\widehat{\omega}^{(1)}}$ and $V_{\widehat{\theta}}$ are given in (3.3.20) and (C.3.9) respectively however, as discussed in Section 4.2, we have no analytic formula for $V_{\widehat{\omega}^{(1)},\widehat{\theta}}$. The simulations, reported in Section 3.9, can be used to estimate the value of $V_{\widehat{\omega}^{(1)},\widehat{\theta}}$. Recall that these simulations consisted of 2,000 realisations of selected SMA(1) processes each of length 720, from each of which the estimators $\widehat{\omega}^{(1)}$ and $\widehat{\theta}$ were calculated. Modest correlations were found between $\widehat{\alpha}$ and $\widehat{\theta}_1$ and between $\widehat{\gamma}^{(1)}$ and $\widehat{\theta}_1$. Negligible correlation was found between $\widehat{\beta}^{(1)}$ and $\widehat{\theta}_1$.

In this section, we investigate the effect of changes to the value of $V_{\widehat{\omega}^{(1)},\widehat{\theta}}$ has on $V_{\widetilde{\zeta}_p^{(r)}}$. More specifically, we investigate changes to the value of $V_{\widetilde{\zeta}_p^{(r)}}$ under each of the following assumptions.

A4.4.1 The asymptotic covariance matrix $V_{\widehat{\omega}^{(1)},\widehat{\theta}}$ is zero, i.e. that $\widehat{\omega}^{(1)}$ and $\widehat{\theta}$ are asymptotically independent. Denote the asymptotic covariance matrix under this assumption by $V^Z_{\widetilde{\zeta}_p^{(r)}}$.

A4.4.2 The asymptotic covariance matrix $V_{\widehat{\omega}^{(1)},\widehat{\theta}}$ is equal to the simulated values reported in Section 3.9. Denote the asymptotic covariance matrix under this assumption by $V^S_{\widetilde{\zeta}_p^{(r)}}$.

Let

$$R^{ZS}_{\widetilde{\zeta}_p^{(r)}} = \frac{V^Z_{\widetilde{\zeta}_p^{(r)}}}{V^S_{\widetilde{\zeta}_p^{(r)}}} \tag{4.4.10}$$

$R^{ZS}_{\widetilde{\zeta}^{(r)}_p}$		θ_1							
α	r	-0.8	-0.6	-0.4	-0.2	0.2	0.4	0.6	0.8
1.2	5	1.002	1.003	1.008	1.015	0.998	1.000	1.000	1.000
1.2	20	1.001	1.003	1.008	1.011	0.999	1.000	1.000	1.000
1.2	120	0.999	1.001	1.006	1.007	0.999	1.000	1.000	1.000
1.2	720	0.998	1.001	1.005	1.004	0.999	1.000	1.000	1.000
1.5	5	1.011	1.032	1.029	1.029	0.997	0.998	0.999	1.000
1.5	20	1.010	1.032	1.019	1.011	1.001	1.001	1.000	1.000
1.5	120	1.003	1.022	1.009	1.001	1.003	1.001	1.000	1.000
1.5	720	0.999	1.015	1.004	0.997	1.003	1.002	1.000	1.000
1.8	5	1.053	1.074	1.002	0.966	1.027	1.000	0.999	0.999
1.8	20	1.036	1.023	0.971	0.941	1.033	1.012	1.004	1.000
1.8	120	0.994	0.992	0.955	0.941	1.028	1.013	1.004	1.000
1.8	720	0.982	0.983	0.956	0.949	1.023	1.011	1.004	1.000

Table 4.4.1: Estimates of $R^{ZS}_{\widetilde{\zeta}^{(r)}_p}$ for a selection of SMA(1) processes at quantile level $p = (0.05, 0.95)$ and selected aggregation levels.

denote the ratio of the asymptotic variances of the parametric LQD estimator under these assumptions.

In Table 4.4.1 we record the values of $R^{ZS}_{\widetilde{\zeta}^{(r)}_p}$ for a selection of SMA(1) processes and aggregation levels. The effect of correlation between $\widehat{\omega}^{(1)}$ and $\widehat{\theta}$ on the asymptotic variance of the log quantile estimator $\widetilde{\zeta}^{(r)}_p$ is less than 6% for all cases reported in Table 4.4.1. Larger effects, up to 10%, were estimated in some cases at low aggregation levels. For $\alpha = 1.8$, the effect is more significant than for lower values of α, due to the higher absolute correlation values between $\widehat{\omega}^{(1)}$ and $\widehat{\theta}$ and also the higher variance of $\widehat{\theta}$. The effect is more significant for negative values of θ_1, than for positive values. Results of simulations not included in Table 4.4.1 suggest that the values of $R^{ZS}_{\widetilde{\zeta}^{(r)}_p}$ are largely unaffected by changes in quantile level p or the innovation skewness parameter β of the SMA(1) process.

We conclude there is a relatively minor effect of the correlation between $\widehat{\omega}^{(1)}$ and $\widehat{\theta}$ on the asymptotic variance of the log quantile estimator $\widetilde{\zeta}^{(r)}_p$. If we only considered SMA(1)

processes, we could use the results in Table 4.4.1 to correct values for $V_{\widetilde{\zeta}_p^{(r)}}$ for the correlation between $\widehat{\omega}^{(1)}$ and $\widehat{\theta}$. However, this is impractical for SMA(q) processes of higher order.

Remark 4.4.1 *On the basis of the results reported in this section, we consider it reasonable to use* $V^Z_{\widetilde{\zeta}_p^{(r)}}$ *as an approximation of* $V_{\widetilde{\zeta}_p^{(r)}}$ *and shall do so for the remainder of this thesis.*

4.5 Non-parametric estimation of the log quantile difference of the temporal aggregation of an SMA(q) process

In this section we define a non-parametric estimator of the log quantile difference of the base process derived from the empirical quantile estimators.

Definition 4.5.1 ***(Non-parametric Log Quantile Difference Estimator -*** $\widehat{\zeta}_p^{(r)}$***).*** *The estimator* $\widehat{\zeta}_p^{(r)}$ *of the* LQD *at quantile levels* $p = (p_1, p_2)$, *where* $0 < p_1 < p_2 < 1$, *of the temporal aggregation,* $\left\{S_t^{(r)}\right\}$, *of an* SMA(q) *process,* $\{X_t\}$, *is given by*

$$\widehat{\zeta}_p^{(r)} = \ln\left(\left|\widehat{\xi}_{p_2}^{(r)} - \widehat{\xi}_{p_1}^{(r)}\right|\right) \tag{4.5.1}$$

where $\widehat{\xi}_{p_j}^{(r)}$ *are the empirical quantile estimators (Definition 3.2.2) of* $\left\{S_t^{(r)}\right\}$ *at quantile level* p_j, $j = 1, 2$.

Let $\xi_p^{(r)} = \left(\xi_{p_1;0}^{(r)}, \xi_{p_2;0}^{(r)}\right)'$ denote the true values of $\left\{S_t^{(r)}\right\}$ at quantile levels p_1 and p_2 respectively and let

$$\zeta_{p;0}^{(1)} = \ln\left(\left|\xi_{p_2;0}^{(r)} - \xi_{p_1;0}^{(r)}\right|\right) \tag{4.5.2}$$

denote the true value of the LQD of $\left\{S_t^{(r)}\right\}$ at quantile levels $p = (p_1, p_2)$. Since $\left\{S_t^{(r)}\right\}$ is also an SMA(q) process, we have from Theorem 3.2.2 that the joint asymptotic distribution of $\widehat{\xi}_p^{(r)} = \left(\widehat{\xi}_{p_1}^{(r)}, \widehat{\xi}_{p_2}^{(r)}\right)'$ is given by

$$\sqrt{T}\left(\widehat{\xi}_p^{(r)} - \xi_p^{(r)}\right) \xrightarrow{d} N\left(0, V_{\widehat{\xi}}\right) \tag{4.5.3}$$

where the formula for $V_{\widehat{\xi}}$ is given in (3.2.19).

In the following theorem we derive the asymptotic distribution of the estimator $\widehat{\zeta}_p^{(r)}$.

Theorem 4.5.1 *The asymptotic distribution of the non-parametric log quantile difference estimator,* $\widehat{\zeta}_p^{(r)}$, *as the sample size* $T \to \infty$ *is given by*

$$\sqrt{T}\left(\widehat{\zeta}_p^{(r)} - \zeta_{p;0}^{(r)}\right) \xrightarrow{d} N\left(0, V_{\widehat{\zeta}}\right) \tag{4.5.4}$$

where

$$V_{\widehat{\zeta}} = D_{\widehat{\zeta}} V_{\widehat{\xi}} D'_{\widehat{\zeta}} \tag{4.5.5}$$

$$D_{\widehat{\zeta}} = \left[\frac{\partial \widehat{\zeta}_p^{(r)}}{\partial \widehat{\xi}_{p_j}^{(r)}} \Big|_{\widehat{\xi}_p^{(r)} = \xi_{p;0}^{(r)}} \right]_{j=1:2}. \tag{4.5.6}$$

Proof. Define the function $g\left(\xi_p^{(r)}\right) : \mathbb{R}^2 \to \mathbb{R}$ by

$$g\left(\xi_p^{(r)}\right) = \ln\left(\left|\xi_{p_2}^{(r)} - \xi_{p_1}^{(r)}\right|\right). \tag{4.5.7}$$

From the Definition (1.2.1) it is clear that

$$g\left(\xi_{p;0}^{(r)}\right) = \zeta_{p;0}^{(r)}. \tag{4.5.8}$$

To prove the theorem, it remains to be shown that g has a non-zero differential at $\widehat{\xi}_p^{(r)} = \xi_{p;0}^{(r)}$, (Theorem B.2.2). To show that g has a non-zero differential at $\widehat{\xi}_p^{(r)} = \xi_{p;0}^{(r)}$ it is sufficient to show that the partial derivatives exist at $\widehat{\xi}_p^{(r)} = \xi_{p;0}^{(r)}$, are continuous at $\widehat{\xi}_p^{(r)} = \xi_{p;0}^{(r)}$ and that at least one of the partial derivatives is non-zero at $\widehat{\xi}_p^{(r)} = \xi_{p;0}^{(r)}$.

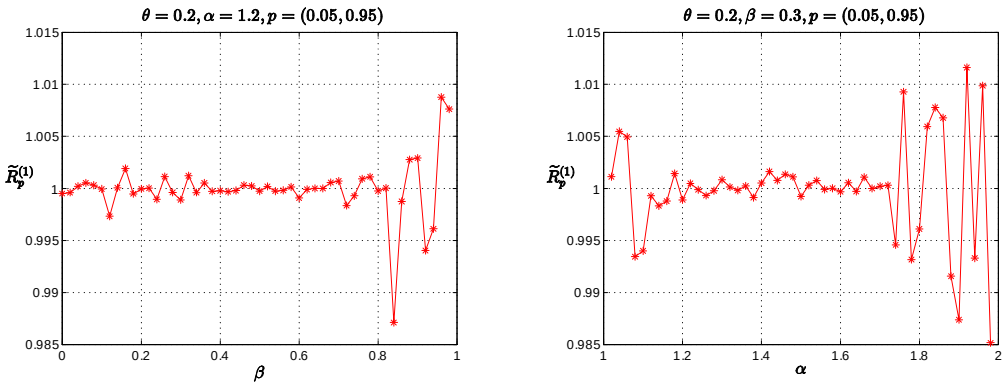
The partial derivatives in $D_{\widehat{\zeta}}$ are given by

$$D_{\widehat{\zeta}} = \left[\frac{-1}{\xi_{p_2;0}^{(r)} - \xi_{p_1;0}^{(r)}}, \frac{1}{\xi_{p_2;0}^{(r)} - \xi_{p_1;0}^{(r)}} \right]. \tag{4.5.9}$$

Given that the density of $\left\{S_t^{(r)}\right\}$ is positive and continuous in the neighbourhoods of $\xi_{p_1;0}^{(1)}$ and $\xi_{p_2;0}^{(1)}$, and by assumption that $0 < p_1, p_2 < 1$, it follows that the partial derivatives $D_{\widehat{\zeta}}$ exist and are continuous and non-zero at $\widehat{\xi}_p^{(r)} = \xi_{p;0}^{(r)}$. This completes the proof of the theorem. ■

Remark 4.5.1 *Although Theorem 4.5.1 assumes that* $\left\{S_t^{(r)}\right\}$ *is the temporal aggregation of an* SMA(q) *process, a similar result is possible for other* ϕ *- mixing processes with non-stable distributions, as long as* $\left\{S_t^{(r)}\right\}$ *has a positive density at* $\xi_{p_1}^{(r)}$ *and* $\xi_{p_2}^{(r)}$.

Remark 4.5.2 *The non-parametric log quantile difference estimator quickly loses efficiency as the aggregation level increases. A comparison of the relative asymptotic efficiency of the non-parametric and parametric* LQD *estimators at aggregation level one is provided in Section 4.6.*



(a) (b)

Figure 4.6.1: Relative asymptotic efficiency of the parametric versus the non-parametric log quantile difference estimators for selected SMA(1) processes at quantile level $p = (0.05, 0.95)$.

4.6 Relative asymptotic efficiency of non-parametric and parametric log quantile difference estimators of an SMA(q) process.

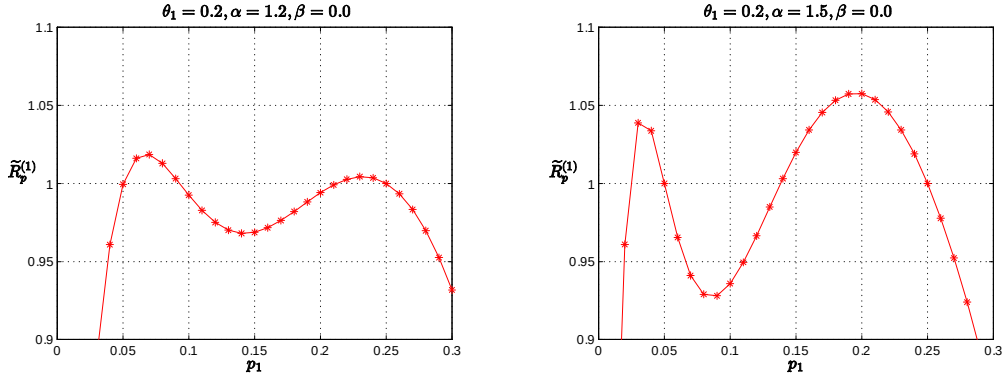
In this section we consider the relative asymptotic efficiency of the non-parametric log quantile difference estimator, $\widehat{\zeta}_p^{(1)}$ and the parametric log quantile difference estimator $\widetilde{\zeta}_p^{(1)}$ as described in earlier sections of this chapter. Let

$$\widetilde{R}_p^{(1)} = V_{\widetilde{\zeta}}/V_{\widehat{\zeta}} \tag{4.6.1}$$

denote the relative asymptotic efficiency of $\widetilde{\zeta}_p^{(1)}$ compared to $\widehat{\zeta}_p^{(1)}$ where $V_{\widetilde{\zeta}}$ and $V_{\widehat{\zeta}}$ are the asymptotic variances given in (4.3.6) and (4.5.5) respectively. Although obscured in the notation, it is worth noting that $\widetilde{R}_p^{(1)}$ may depend on θ and $\omega^{(1)}$ as well as p.

The relative asymptotic efficiencies shown in Figure 4.6.1, are all in the range $[0.985, 1.015]$ and most are in the range $[0.995, 1.005]$. The differences from 1 may be attributabed simply to the numerical approximations involved, especially in calculating $D_{\widehat{\omega}}$. These numerical approximations have a greater impact on the calculation of $D_{\widehat{\omega}}$ where α is close to 2 and $|\beta|$ is close to 1. Calculations of $\widetilde{R}_p^{(1)}$ for other values of θ, α and β produced similar results to those shown in Figure 4.6.1.

Recall that the estimator $\widehat{\omega}^{(1)}$ takes as its inputs, the empirical quantile estimators of



(a) (b)

Figure 4.6.2: Relative asymptotic efficiency of parametric versus non-parametric log quantile difference estimators for $p = (p_1, 1 - p_1)$.

$\{X_t\}$ at the quantile levels

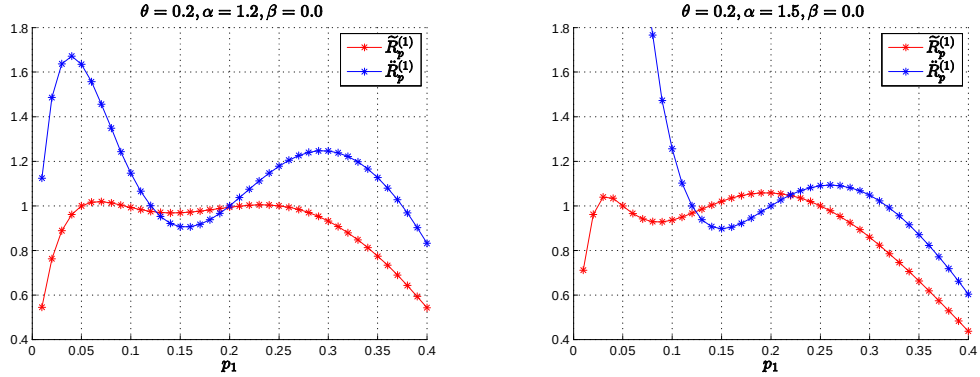
$$p_M = (0.05, 0.25, 0.50, 0.75, 0.95)' \tag{4.6.2}$$

and that the non-parametric log quantile difference estimator takes as its inputs, the empirical quantile estimators of $\{X_t\}$ at $p = (p_1, p_2)'$. Where the quantile levels of p are included in the quantile levels of p_M, one might expect that $\widetilde{R}_p^{(1)} \geq 1$ since the all information available non-parametric estimator is also available to the parametric estimator. That $\widetilde{R}_p^{(1)} \approx 1$ suggests that the empirical quantile estimators at the other quantile levels do not contribute significantly to the parametric log quantile difference estimator.

Where the quantile levels of p are not included in the quantile levels of p_M, we see that there can be significant deviations from 1 in $\widetilde{R}_p^{(1)}$ (Figure 4.6.2). Where $p_1 = 0.05$ and $p_1 = 0.25$, we still have $\widetilde{R}_p^{(1)} \approx 1$. However, there are values of p_1 for which $\widetilde{R}_p^{(1)} < 1$, i.e. the parametric estimator is more efficient, and other values of p_1 for which $\widetilde{R}_p^{(1)} > 1$, i.e. the non-parametric estimator is more efficient.

As seen in Section 3.5, more efficient estimators for $\omega^{(1)}$ are available through an appropriate choice of the quantile levels p_M other than the standard levels. Similarly, given some choice of p it may be possible through an appropriate choice of p_M to define a more efficient parametric estimator of $\zeta_p^{(1)}$.

Suppose we wish to estimate $\zeta_p^{(1)}$ where $p = (0.15, 0.85)$. A more efficient choice of p_M



(a) (b)

Figure 4.6.3: Comparison of relative asymptotic efficiency of log quantile difference estimators $\widetilde{\zeta}_p^{(1)}$ and $\ddot{\zeta}_p^{(1)}$ for $p = (p_1, 1-p_1)$.

for this task might be

$$p_M = (0.12, 0.20, 0.50, 0.80, 0.88)' \tag{4.6.3}$$

which provides a pair of empirical quantile estimators at quantile levels close to each of the quantile levels 0.15 and 0.85. Let $\ddot{\omega}^{(1)}$ denote the stable distribution parameter estimator calculated using quantile levels for p_M given in (4.6.3) . Let $\ddot{\zeta}_p^{(1)}$ denote the corresponding LQD estimator and $\ddot{R}_p^{(1)}$ the corresponding relative asymptotic efficiency compared with the non-parametric estimator.

The relative asymptotic efficiencies $\widetilde{R}_p^{(1)}$ and $\ddot{R}_p^{(1)}$ are shown in Figure 4.6.3 for the same processes used in Figure 4.6.2. For most values of p_1, the original non-parametric estimator $\widetilde{\zeta}_p^{(1)}$ is superior to the new non-parametric estimator $\ddot{\zeta}_p^{(1)}$. However, the new non-parametric estimator $\ddot{\zeta}_p^{(1)}$ does produce a superior estimate at $p = (0.15, 0.85)'$.

α	p	$\widetilde{R}_p^{(1)}$	$\ddot{R}_p^{(1)}$
1.2	$(0.15, 0.85)'$	0.9687	0.9063
1.5	$(0.15, 0.85)'$	1.0199	0.8974

If the quantile levels in p are included in the quantile levels of p_M, it is unclear from the calculated values of $\widetilde{R}_p^{(1)}$ given in this section whether or not $\widetilde{R}_p^{(1)}$ is exactly equal to one or whether it is just very close. We note that $\ddot{R}_p^{(1)}$ is very close to one at $p = (0.12, 0.88)'$ and $(0.20, 0.80)'$.

Let us consider the proposition that $\widetilde{R}_p^{(1)}$ is equal to one whenever the quantile levels in p are included in the quantile levels in p_M. Let ξ_M denote the quantiles of $\{X_t\}$ at the quantile levels p_M and let $V_{\widetilde{\xi}_M}, V_{\widehat{\xi}_M}$ denote respectively the asymptotic covariance matrices of the parametric quantile estimator $\widetilde{\xi}_M$ and the non-parametric quantile estimator $\widehat{\xi}_M$. Now if $\widetilde{R}_p^{(1)}$ is equal to one, then $V_{\widetilde{\zeta}} = V_{\widehat{\zeta}}$ and $V_{\widetilde{\xi}_M} = V_{\widehat{\xi}_M}$ and it would follow from (4.3.6) and (3.3.20) that

$$V_{\widetilde{\xi}_M} = D_{\widetilde{\xi}} D_{\widehat{\omega}} V_{\widehat{\xi}_M} D'_{\widehat{\omega}} D'_{\widetilde{\xi}} = V_{\widehat{\xi}_M} \tag{4.6.4}$$

and therefore that

$$D_{\widetilde{\xi}} D_{\widehat{\omega}} L_{\widehat{\xi}_M} L'_{\widehat{\xi}_M} D'_{\widehat{\omega}} D'_{\widetilde{\xi}} = L_{\widehat{\xi}_M} L'_{\widehat{\xi}_M} \tag{4.6.5}$$

and

$$D_{\widetilde{\xi}} D_{\widehat{\omega}} = I_4 \tag{4.6.6}$$

where $L_{\widehat{\xi}_M}$ is the lower Cholesky decomposition of $V_{\widehat{\xi}_M}$ and I_4 is the 4×4 identity matrix. Whilst $D_{\widetilde{\xi}}$ and $D_{\widehat{\omega}}$ are not square matrices and therefore do not have proper inverses we can construct a left-sided inverse of $D_{\widetilde{\xi}}$ given by

$$D^{-1}_{\widetilde{\xi};left} = \left(D'_{\widetilde{\xi}} D_{\widetilde{\xi}} \right)^{-1} D'_{\widetilde{\xi}} \tag{4.6.7}$$

with which we can multiply the left sides of both parts of (4.6.6) to get

$$D_{\widehat{\omega}} = D^{-1}_{\widetilde{\xi};left}. \tag{4.6.8}$$

Now if (4.6.8) were true, it would be quite useful since $D_{\widehat{\omega}}$ is difficult to accurately calculate (Section 3.3) whereas $D_{\widetilde{\xi}}$ is relatively easy to accurately calculate (Section 4.3). For $\theta_1 = 0.2, \alpha = 1.5$ and $\beta = 0.0$, we have calculated $D_{\widehat{\omega}}$ and $D^{-1}_{\widetilde{\xi};left}$ with the following results

$$D_{\widehat{\omega}} = \begin{pmatrix} -0.179 & 0.564 & 0.000 & -0.564 & 0.179 \\ 0.323 & 0.000 & -0.646 & 0.000 & 0.323 \\ -0.006 & 0.536 & 0.000 & -0.536 & 0.006 \\ -0.089 & 0.000 & 1.179 & -0.000 & -0.089 \end{pmatrix}, \tag{4.6.9}$$

and

$$D^{-1}_{\widetilde{\xi};left} = \begin{pmatrix} -0.179 & 0.564 & 0.000 & -0.564 & 0.179 \\ 0.345 & -0.208 & -0.275 & -0.208 & 0.345 \\ -0.006 & 0.536 & 0.000 & -0.536 & 0.006 \\ -0.133 & 0.400 & 0.465 & 0.400 & -0.133 \end{pmatrix}. \tag{4.6.10}$$

The first and third rows of $D_{\widehat{\omega}}$ and $D^{-1}_{\widetilde{\xi};left}$, representing respectively the partial derivatives of $\widehat{\alpha}$ and $\widehat{\gamma}^{(1)}$ with respect to $\widehat{\xi}_M$ are identical to three decimal places. However, the second and fourth rows of $D_{\widehat{\omega}}$ and $D^{-1}_{\widetilde{\xi};left}$, representing respectively the partial derivatives of $\widehat{\beta}^{(1)}$ and $\widehat{\delta}^{(1)}$ with respect to $\widehat{\xi}_M$ are quite different. The zero entries in second and fourth rows of $D_{\widehat{\omega}}$ indicate that the estimates for $\widehat{\beta}^{(1)}$ and $\widehat{\delta}^{(1)}$ are not affected by the non-parametric quantile estimates $\widehat{\xi}_{0.25}$ and $\widehat{\xi}_{0.75}$. The corresponding non-zero entries in $D^{-1}_{\widetilde{\xi};left}$ indicate that parametric quantile estimates $\widetilde{\xi}_{0.25}$ and $\widetilde{\xi}_{0.75}$ are not affected by $\widehat{\beta}^{(1)}$ and $\widehat{\delta}^{(1)}$.

That $D_{\widehat{\omega}}$ and $D^{-1}_{\widetilde{\xi};left}$ in (4.6.9) and (4.6.10) respectively are not identical suggests that, at least in this case, our proposition that $\widetilde{R}^{(1)}_p$ is equal to one whenever the quantile levels in p are also in p_M is incorrect.

4.7 Analytic estimation of the log quantile difference of the temporal aggregation of an SMA(q) process

In this section we define an analytic estimator, $\check{\zeta}^{(r)}_p$, for the log quantile difference of the temporal aggregation of an SMA(q) process. Unlike the parametric log quantile difference estimator $\widetilde{\zeta}^{(r)}_p$ (Definition 4.3.2) which is valid for all SMA(q) processes, the analytic estimator, $\check{\zeta}^{(r)}_p$, is valid only where the base process satisfies the assumptions of Corollary 2.2.2. We show that this estimator is asymptotically normal and derive a formula for its asymptotic distribution. The use of $\check{\zeta}^{(r)}_p$ instead of $\widetilde{\zeta}^{(r)}_p$ has minor computational advantages.

Let $\widehat{\alpha}$ be an estimator of the stable distribution parameter α of the base process $\{X_t\}$. Let $\widehat{\theta} = \left(\widehat{\theta}_1, \ldots, \widehat{\theta}_q\right)'$ be estimators of the moving average parameters $\theta = (\theta_1, \ldots, \theta_q)'$ of $\{X_t\}$ and let $\widehat{\zeta}^{(1)}_p$ (Definition 4.5.1) be the non-parametric estimator of the log quantile difference of $\{X_t\}$ at quantile levels $p = (p_1, p_2)$. Note that we could use the parametric estimator $\widetilde{\zeta}^{(1)}_p$ instead of the non-parametric estimator $\widehat{\zeta}^{(1)}_p$ however, doing so mitigates the computational advantages of the analytic estimator $\check{\zeta}^{(r)}_p$.

To simplify the notation we define the $2+q$ dimensional vector of parameters

$$\chi = \left(\alpha, \zeta^{(1)}_p, \theta'\right)'. \tag{4.7.1}$$

We use similar definitions for $C^{(r)}_1(\chi)$ and $C_2(\chi)$ to those used in Section 4.2,

$$C^{(r)}_1(\chi) = \sum_{j=0}^{r+q-1} \left|c^{(r)}_j\right|^{\alpha} \tag{4.7.2}$$

and

$$C_2(\chi) = \sum_{j=0}^{q} |\theta_j|^{\alpha} \tag{4.7.3}$$

where $c_j^{(r)}$ are the moving average parameters of the aggregated process given in Theorem 2.2.1,

$$c_j^{(r)} = \sum_{k=\max(j-r+1,0)}^{\min(j,q)} \theta_k. \tag{4.7.4}$$

Where appropriate, we may simply refer to $C_1^{(r)}(\varkappa)$ and $C_2(\varkappa)$ as $C_1^{(r)}$ and C_2 respectively.

Let $\widehat{\chi}$ denote the estimator of χ

$$\widehat{\chi} = \left(\widehat{\alpha}, \widehat{\zeta}_p^{(1)}, \widehat{\theta}'\right)' \tag{4.7.5}$$

where $\widehat{\alpha}$ is the quantile-based estimator of α defined in Section 3.3, $\widehat{\zeta}_p^{(1)}$ is the non-parametric log quantile difference estimator of $\zeta_p^{(1)}$ defined in Section 4.5 and $\widehat{\theta}$ is the SLAD moving average parameter estimator of θ defined in Appendix C.3. The asymptotic distribution of $\widehat{\chi}$ is given in the following theorem.

Theorem 4.7.1 *Let χ_0 denote the true value of the parameters χ. As the sample size $T \to \infty$ the asymptotic distribution of $\widehat{\chi}$ is given by*

$$\sqrt{T}(\widehat{\chi} - \chi) \xrightarrow{d} N\left(0, V_{\widehat{\chi}}\right) \tag{4.7.6}$$

where

$$V_{\widehat{\chi}} = \begin{pmatrix} V_{\widehat{\alpha}} & V_{\widehat{\alpha},\widehat{\zeta}_p^{(1)}} & V_{\widehat{\alpha},\widehat{\theta}} \\ V_{\widehat{\alpha},\widehat{\zeta}_p^{(1)}} & V_{\widehat{\zeta}_p^{(1)}} & V_{\widehat{\zeta}_p^{(1)},\widehat{\theta}} \\ V_{\widehat{\alpha},\widehat{\theta}} & V_{\widehat{\zeta}_p^{(1)},\widehat{\theta}} & V_{\widehat{\theta}} \end{pmatrix} \tag{4.7.7}$$

and $V_{\widehat{\alpha}}$ is the asymptotic variance of $\widehat{\alpha}$ included in (3.3.20), *$V_{\widehat{\zeta}_p^{(1)}}$ is given in* (4.5.5) *and $V_{\widehat{\theta}}$ given in* (C.3.9).

Proof. This theorem follows from the asymptotic distributions of $\widehat{\alpha}, \widehat{\theta}$ and $\widehat{\zeta}_p^{(1)}$ established in Theorems 3.3.1, C.3.1 and 4.5.1 respectively. ■

As with the calculation of $V_{\widehat{\omega}^{(1)},\widehat{\theta}}$ discussed in Section 4.4, we have no formulae for the asymptotic covariances $V_{\widehat{\alpha},\widehat{\theta}}$ and $V_{\widehat{\zeta}_p^{(1)},\widehat{\theta}}$. However, we do have a formula for $V_{\widehat{\alpha},\widehat{\zeta}_p^{(1)}}$. Since the estimators $\widehat{\alpha}$ and $\widehat{\zeta}_p^{(1)}$ are both derived from quantile estimates of the base process $\{X_t\}$ it

is not surprising that there is an non-zero asymptotic correlation between them. A method for the calculation of $V_{\widehat{\alpha},\widehat{\zeta}_p^{(1)}}$ is described in Section 4.8. The results of simulations used to estimate $V_{\widehat{\alpha},\widehat{\theta}}$ and $V_{\widehat{\zeta}_p^{(1)},\widehat{\theta}}$ are reported in Section 4.9.

We are now able to define the analytic log quantile difference estimator $\check{\zeta}_p^{(r)}$ of $\zeta_p^{(r)}$.

Definition 4.7.1 ***(Analytic Log Quantile Difference Estimator -*** $\check{\zeta}_p^{(r)}$***).*** *We define the estimator,* $\check{\zeta}_p^{(r)}$*, for the log quantile difference of the temporal aggregation* $\left\{S_t^{(r)}\right\}$ *of an* SMA(q) *process* $\{X_t\}$ *at aggregation level* r *and quantile levels* $p = (p_1, p_2)'$ *to be*

$$\check{\zeta}_p^{(r)} = \widehat{\alpha}^{-1} \ln\left(\frac{C_1^{(r)}(\widehat{\chi})}{C_2(\widehat{\chi})}\right) + \widehat{\zeta}_p^{(1)}. \tag{4.7.8}$$

In the following theorem we derive the asymptotic distribution of $\check{\zeta}_p^{(r)}$.

Theorem 4.7.2 *Let* χ_0 *denote the true value of the parameters* χ *and let* $\zeta_{p;0}^{(r)}$ *denote the true value of the log quantile difference* $\zeta_p^{(r)}$*. If*

$$\begin{aligned} \textit{either } \beta &= 0 \\ \textit{or } \theta_j &\geq 0 \quad \textit{for } i = 0, \ldots, q \end{aligned} \tag{4.7.9}$$

and

$$\begin{aligned} \textit{either } \alpha &\geq 1 \\ \textit{or } c_j^{(r)} &\neq 0, \quad \textit{for } j = 0, \ldots, r+q-1 \textit{ and } \theta_j \neq 0, \quad \textit{for } j = 1, \ldots, q, \end{aligned} \tag{4.7.10}$$

then the asymptotic distribution of $\check{\zeta}_p^{(r)}$ *is given by*

$$\sqrt{T}\left(\check{\zeta}_p^{(r)} - \zeta_{p;0}^{(r)}\right) \xrightarrow{d} N\left(0, V_{\check{\zeta}_p^{(r)}}\right) \tag{4.7.11}$$

where

$$V_{\check{\zeta}_p^{(r)}} = D_{\check{\zeta}_p^{(r)}} V_{\widehat{\chi}} D'_{\check{\zeta}_p^{(r)}} \tag{4.7.12}$$

and

$$D_{\check{\zeta}_p^{(r)}} = \left[\frac{\partial \check{\zeta}_p^{(r)}}{\partial \widehat{\chi}_j} \mid_{\widehat{\chi}=\chi_0}\right]_{j=1:2+q}. \tag{4.7.13}$$

Proof. See Appendix A.4. ■

Remark 4.7.1 *The assumption in Theorem 4.7.2 that either $\beta = 0$ or $\theta_j \geq 0$ for $i = 0, \ldots, q$ is necessary to ensure that $\check{\zeta}_p^{(r)}$ is an unbiased estimator.*

Remark 4.7.2 *The assumption in Theorem 4.7.2 that either $\alpha \geq 1$ or $c_j^{(r)} \neq 0$, for $j = 0, \ldots, r+q-1$ and $\theta_j \neq 0$, for $j = 1, \ldots, q$ is necessary to ensure that the partial derivative $\dfrac{\partial \check{\zeta}_p^{(r)}}{\partial \widehat{\theta}_k}$ exist.*

Remark 4.7.3 *It is possible for the assumptions in Theorem 4.7.2 to be satisfied and the assumptions in Theorem 4.2.2 are not satisfied. For example, if $\beta = 0$, $C_3 = 0$ and $\alpha \geq 1$. Under such circumstances, an asymptotic distribution for the analytic log quantile difference estimator can be calculated, but an asymptotic distribution for the parametric log quantile difference estimator cannot be calculated.*

There are computational advantages of using the analytic LQD estimator $\check{\zeta}_p^{(r)}$ where possible rather than parametric LQD estimator $\widetilde{\zeta}_p^{(r)}$. These advantages are largely related to the calculation of respective asymptotic variances. Recall the formula for the asymptotic variance of the parametric LQD estimator $\widetilde{\zeta}_p^{(r)}$

$$V_{\widetilde{\zeta}_p^{(r)}} = D_{\widetilde{\zeta}_p^{(r)}} D_{\widetilde{\xi}_p^{(r)}} D_{\widehat{\omega}^{(r)}} V_{\widehat{\varkappa}} D'_{\widehat{\omega}^{(r)}} D'_{\widetilde{\xi}_p^{(r)}} D'_{\widetilde{\zeta}_p^{(r)}}, \tag{4.7.14}$$

and for the analytic LQD estimator $\check{\zeta}_p^{(r)}$

$$V_{\check{\zeta}_p^{(r)}} = D_{\check{\zeta}_p^{(r)}} V_{\widehat{\chi}} D'_{\check{\zeta}_p^{(r)}}. \tag{4.7.15}$$

1. Calculation of $V_{\widehat{\varkappa}}$ requires simulated estimates for the asymptotic covariances $V_{\widehat{\alpha},\widehat{\theta}_1}$, $V_{\widehat{\beta}^{(1)},\widehat{\theta}_1}$ and $V_{\widehat{\gamma}^{(1)},\widehat{\theta}_1}$ and numerical approximations for the asymptotic covariances $V_{\widehat{\alpha}}$, $V_{\widehat{\beta}^{(1)}}$, $V_{\widehat{\gamma}^{(1)}}$, $V_{\widehat{\alpha},\widehat{\beta}^{(1)}}$, $V_{\widehat{\alpha},\widehat{\gamma}^{(1)}}$, $V_{\widehat{\beta}^{(1)},\widehat{\gamma}^{(1)}}$ and $V_{\widehat{\theta}_1}$.

2. Calculation of $V_{\widehat{\chi}}$ requires simulated estimates for the asymptotic covariances $V_{\widehat{\alpha},\widehat{\theta}_1}$ and $V_{\widehat{\zeta}_p^{(1)},\widehat{\theta}_1}$, use of formulae for $V_{\widehat{\alpha},\widehat{\zeta}_p^{(1)}}$ and $V_{\widehat{\zeta}_p^{(1)}}$ and numerical approximations for the asymptotic covariances $V_{\widehat{\alpha}}$ and $V_{\widehat{\theta}_1}$.

3. The partial differential matrices $D_{\widehat{\omega}^{(r)}}$ and $D_{\check{\zeta}_p^{(r)}}$ are calculated using similar formulae, though the formulae for $D_{\check{\zeta}_p^{(r)}}$ are somewhat simpler.

4. The numerical approximations for $D_{\widetilde{\xi}_p^{(r)}}$ are not required for the analytic estimator.

4.8 Joint asymptotic distribution of $\widehat{\alpha}$ and $\widehat{\zeta}_p^{(1)}$

The estimators $\widehat{\alpha}$ and $\widehat{\zeta}_p^{(1)}$ are both asymptotic normal estimators derived from quantile estimates of the base process $\{X_t\}$. Thus, we can use the multivariate delta theorem (Theorem B.2.3) to derive a joint asymptotic distribution for $\widehat{\alpha}$ and $\widehat{\zeta}_p^{(1)}$. We begin with the case where the quantile levels

$$p = (p_1, p_2)' \tag{4.8.1}$$

of the log quantile difference $\zeta_p^{(1)}$ are not included in the quantile levels

$$p_M = (p_{M;1}, p_{M;2}, p_{M;3}, p_{M;4}, p_{M;5}) \tag{4.8.2}$$

used for the quantile-based estimation of α. Let

$$p_0 = \left(p', p_M'\right)' \tag{4.8.3}$$

denote the following vector of the union of quantile levels p and p_M, let

$$\xi_{p_0} = \left(\xi_{p_1}, \xi_{p_2}, \xi_{p_{M;1}}, \xi_{p_{M;2}}, \xi_{p_{M;3}}, \xi_{p_{M;4}}, \xi_{p_{M;5}}\right)' \tag{4.8.4}$$

denote the vector of quantiles of $\{X_t\}$ at quantile levels p_0 and let $\widehat{\xi}_{p_0}$ denote the associated vector of empirical quantile estimators from the base process $\{X_t\}$. The following theorem derives an expression joint asymptotic distribution of $\widehat{\alpha}$ and $\widehat{\zeta}_p^{(1)}$.

Theorem 4.8.1 *Given a sample of size T from the base process $\{X_t\}$, the estimators $\widehat{\alpha}$ and $\widehat{\zeta}_p^{(1)}$ satisfy*

$$\sqrt{T}\left(\begin{pmatrix} \widehat{\alpha} \\ \widehat{\zeta}_p^{(1)} \end{pmatrix} - \begin{pmatrix} \alpha \\ \zeta_p^{(1)} \end{pmatrix}\right) \xrightarrow{d} N\left(0, D'\Sigma_{p_0}D\right) \tag{4.8.5}$$

where

$$D = \begin{pmatrix} 0 & 0 & \frac{\partial\widehat{\alpha}}{\partial\widehat{\xi}_{p_{M;1}}} & \frac{\partial\widehat{\alpha}}{\partial\widehat{\xi}_{p_{M;2}}} & \frac{\partial\widehat{\alpha}}{\partial\widehat{\xi}_{p_{M;3}}} & \frac{\partial\widehat{\alpha}}{\partial\widehat{\xi}_{p_{M;4}}} & \frac{\partial\widehat{\alpha}}{\partial\widehat{\xi}_{p_{M;5}}} \\ \frac{-1}{\widehat{\xi}_{p_2} - \widehat{\xi}_{p_1}} & \frac{1}{\widehat{\xi}_{p_2} - \widehat{\xi}_{p_1}} & 0 & 0 & 0 & 0 & 0 \end{pmatrix}' |_{\widehat{\xi}_{p_0} = \xi_{p_0}}, \tag{4.8.6}$$

and Σ_{p_0} is the asymptotic covariance matrix of the empirical quantile estimators $\widehat{\xi}_{p_0}$ given in Theorem 3.2.1.

Proof. The proof of this theorem involves an application of the multivariate delta theorem (Theorem B.2.3). Define the function g as

$$g\left(\widehat{\xi}_{p_0}\right) = \begin{pmatrix} \widehat{\alpha} \\ \widehat{\zeta}_p^{(1)} \end{pmatrix}. \tag{4.8.7}$$

Since $\widehat{\alpha}$ does not depend on $\widehat{\xi}_{p_2}, \widehat{\xi}_{p_1}$, the partial derivatives of $\widehat{\alpha}$ with respect to these estimators is 0. Similarly the partial derivatives of $\widehat{\zeta}_p^{(1)}$ with respect to $\widehat{\xi}_{p_{M;1}}$, $\widehat{\xi}_{p_{M;2}}$, $\widehat{\xi}_{p_{M;3}}$, $\widehat{\xi}_{p_{M;4}}$ and $\widehat{\xi}_{p_{M;5}}$ are also all 0. Thus D in (4.8.6) represents the matrix of partial derivatives of $g\left(\widehat{\xi}_{p_0}\right)$ with respect to each of the components evaluated at ξ_{p_0} and an application of Theorem B.2.3 proves this theorem. ■

In the proof of Theorem 4.8.1, we assume that

$$p_1, p_2 \notin \{p_{M;1}, p_{M;2}, p_{M;3}, p_{M;4}, p_{M;5}\}. \tag{4.8.8}$$

If say, $p_1 = p_{M;2}$, then we can achieve a similar result by defining

$$\xi_{p_0} = \left(\xi_{p_2}, \xi_{p_{M;1}}, \xi_{p_{M;2}}, \xi_{p_{M;3}}, \xi_{p_{M;4}}, \xi_{p_{M;5}}\right)' \tag{4.8.9}$$

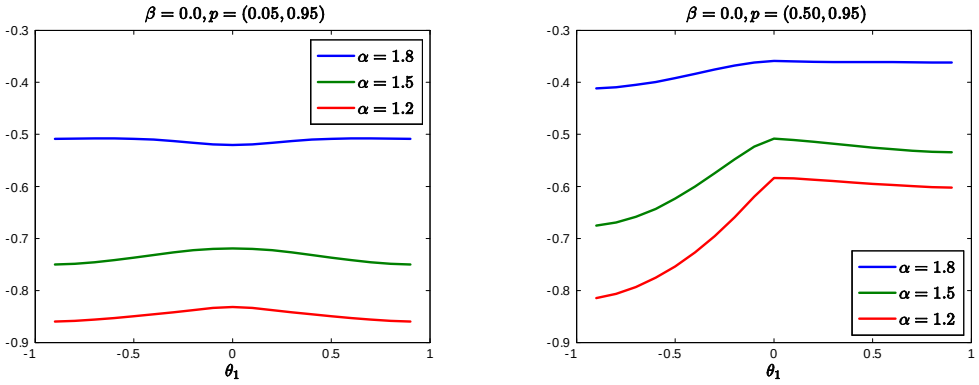
and

$$D = \begin{pmatrix} 0 & \frac{\partial\widehat{\alpha}}{\partial\widehat{\xi}_{p_{M;1}}} & \frac{\partial\widehat{\alpha}}{\partial\widehat{\xi}_{p_{M;2}}} & \frac{\partial\widehat{\alpha}}{\partial\widehat{\xi}_{p_{M;3}}} & \frac{\partial\widehat{\alpha}}{\partial\widehat{\xi}_{p_{M;4}}} & \frac{\partial\widehat{\alpha}}{\partial\widehat{\xi}_{p_{M;5}}} \\ \frac{1}{\widehat{\xi}_{p_2} - \widehat{\xi}_{p_1}} & 0 & \frac{-1}{\widehat{\xi}_{p_2} - \widehat{\xi}_{p_1}} & 0 & 0 & 0 \end{pmatrix}' |_{\widehat{\xi}_{p_0} = \xi_{p_0}}. \tag{4.8.10}$$

Similar adjustments can be made should both p_1 and p_2 be elements of p_M.

To illustrate the result of Theorem 4.8.1, we plot in Figure 4.8.1 the asymptotic correlation between $\widehat{\alpha}$ and $\widehat{\zeta}_p^{(1)}$ where $\{X_t\}$ are selected symmetric SMA(1) processes. A significant negative correlation between $\widehat{\alpha}$ and $\widehat{\zeta}_p^{(1)}$ can be seen for each of the selected values of α and θ_1. This correlation appears to become more significant as α decreases and to be dependent on the quantile levels p.

From some additional calculations not shown in Figure 4.8.1, it can be seen that, for quantile levels $p = (0.05, 0.95)'$ and $\alpha \approx 1.73$, the asymptotic correlation is almost independent of θ_1. Recall that in Section 3.6, we saw that the value of $\alpha \approx 1.73$ is also where the asymptotic and finite sample distributions of the $\widehat{\rho}(1)$ estimator most closely aligned. An interesting area of further research may be to investigate whether this is merely a coincidence or whether this is the result of some inherent property of the stable distribution.



(a) (b)

Figure 4.8.1: Asymptotic correlation between $\widehat{\alpha}$ and $\widehat{\zeta}_p^{(1)}$ for (a) $p = (0.05, 0.95)$ and (b) $p = (0.50, 0.95)$ where $\{X_t\}$ is a symmetric SMA(1) process with moving average parameter θ_1.

4.9 Simulation

In this section, the results of simulations are presented to demonstrate the use of the methods described in this chapter for the estimation of the log quantile difference of the temporal aggregation of an SMA(q) process. For each simulation 1,000 realisations, each of length 720, of an SMA(1) process are generated for selected values of α, β and θ_1. For each realisation of an SMA(1) process, estimates for $\zeta_p^{(r)}$ are calculated and the results compared with the true value. The parameters $\gamma = 2$ and $\delta = 1$ are fixed for all simulations. We refer to the product of the sample size and the variance of an estimator across all realisations of an SMA(q) process as the normalised variance of the estimator.

The first round of simulations demonstrates the use of the parametric log quantile difference estimator, $\widetilde{\zeta}_p^{(r)}$, assuming that θ_1 is known, (Table 4.9.1). The assumption that θ_1 is known means that the components $V_{\widehat{\theta}}$ and $V_{\widehat{\omega}^{(1)},\widehat{\theta}}$ of $V_{\widehat{\varkappa}}$ in (4.4.3) are zero. Thus we can calculate $V_{\widehat{\varkappa}}$ and consequently $V_{\widetilde{\zeta}_p^{(r)}}$ without any further assumptions about $V_{\widehat{\theta}}$ and $V_{\widehat{\omega}^{(1)},\widehat{\theta}}$. Note that estimates for different aggregation levels are taken from the same process realisations.

For each process and aggregation level, the mean of the estimator $\widetilde{\zeta}_p^{(r)}$ across all realisations of the process is within one standard deviation of the true value. For each process

$p=(0.05,0.95)'$			$\theta_1=-0.4$			$\theta_1=0.4$		
α	β	r	$\zeta_p^{(r)}$	(i)	(ii)	$\zeta_p^{(r)}$	(i)	(ii)
1.2	0.0	20	4.905	4.917 (0.217)	33.97 [33.24]	5.689	5.704 (0.224)	36.01 [33.15]
1.2	0.0	720	7.834	7.857 (0.380)	104.18 [101.69]	8.680	8.707 (0.394)	111.62 [101.68]
1.2	0.5	20	4.918	4.922 (0.223)	35.94 [37.48]	5.704	5.720 (0.257)	47.53 [46.07]
1.2	0.5	720	7.849	7.859 (0.392)	110.71 [116.25]	8.695	8.725 (0.464)	154.87 [150.24]
1.5	0.0	20	4.043	4.049 (0.150)	16.24 [15.67]	4.828	4.824 (0.156)	17.45 [15.65]
1.5	0.0	720	6.379	6.391 (0.272)	53.29 [51.24]	7.225	7.217 (0.283)	57.59 [51.24]
1.5	0.5	20	4.055	4.056 (0.154)	17.07 [17.04]	4.841	4.850 (0.169)	20.43 [19.87]
1.5	0.5	720	6.393	6.395 (0.281)	56.81 [56.74]	7.238	7.253 (0.310)	69.28 [67.54]
1.8	0.0	20	3.511	3.512 (0.114)	8.95 [10.61]	4.296	4.286 (0.111)	9.06 [10.64]
1.8	0.0	720	5.450	5.451 (0.230)	36.60 [43.87]	6.296	6.273 (0.225)	36.85 [43.87]
1.8	0.5	20	3.515	3.515 (0.114)	9.47 [10.86]	4.300	4.301 (0.116)	9.56 [11.16]
1.8	0.5	720	5.455	5.453 (0.231)	38.12 [44.83]	6.300	6.301 (0.235)	38.63 [45.98]

Table 4.9.1: Parametric estimation of the log quantile difference $\zeta_p^{(r)}$ for $p=(0.05,0.95)'$ at selected aggregation levels from samples of selected SMA(1) processes assuming that θ_1 is known. Included in this table are (i) the mean and standard deviation, in ( ), of the estimates across all realisations and (ii) the standardised variance, i.e. the variance multiplied by the sample size T, across all realisations and the asymptotic variance, in [ ].

and aggregation level, the normalised variance of the estimator $\widetilde{\zeta}_p^{(r)}$ across all realisations is within 20% of the asymptotic variance. There does not appear to be any obvious bias in either the mean or normalised variance of $\widetilde{\zeta}_p^{(r)}$.

The second round of simulations demonstrates the use of the parametric log quantile difference estimator, $\widetilde{\zeta}_p^{(r)}$ where θ_1 needs to be estimated, (Table 4.9.2). We have no method for calculating the true values of $V_{\widehat{\theta}}$ and hence also $V_{\widetilde{\zeta}_p^{(r)}}$. However, we can estimate $V_{\widetilde{\zeta}_p^{(r)}}$ under the assumption that all elements of $V_{\widehat{\omega}^{(1)},\widehat{\theta}}$ are zero (Section 4.4) and compare those estimates to the normalised variance of $\widetilde{\zeta}_p^{(r)}$. The same samples used in the first round of simulations are also used in the second round of simulations.

For each process and aggregation level, the mean of the estimator $\widetilde{\zeta}_p^{(r)}$ across all realisations of the process is within one standard deviation of the true value. For each process and aggregation level, the normalised variance of the estimator $\widetilde{\zeta}_p^{(r)}$ across all realisations is within 20% of the mean estimated asymptotic variance. There does not appear to be any obvious bias in either the mean or normalised variance of $\widehat{\zeta}_p^{(r)}$.

The variances of $\widetilde{\zeta}_p^{(r)}$ in Table 4.9.1 where θ_1 is assumed known are generally higher than the variances of $\widetilde{\zeta}_p^{(r)}$ in Table 4.9.2 where θ_1 needs to be estimated. This difference is more significant where $\theta_1 = -0.4$ and $\alpha = 1.8$ but surprisingly small for other processes. This is partially explained by the increase in V_θ as α increases.

The third round of simulations compares the use parametric log quantile difference estimator $\widetilde{\zeta}_p^{(r)}$ and the non-parametric log quantile difference estimator at aggregation level $r = 1$. All parametric estimators in these simulations are calculated using

$$p_M = (0.05, 0.25, 0.50, 0.75, 0.95)'. \tag{4.9.1}$$

Included in Tables 4.9.3, 4.9.4 and 4.9.5 are the simulation results at quantile levels $p = (0.05, 0.95)'$, $p = (0.10, 0.90)'$ and $p = (0.20, 0.80)'$ respectively.

For each simulation, the mean of the log quantile difference estimates across all realisations of a process is within one standard deviation of the true value. The standardised variance of the log quantile difference estimators are within 15% of their asymptotic variance.

In Table 4.9.3 for quantile level $p = (0.05, 0.95)'$ where both elements of p are included in p_M the asymptotic variance of the non-parametric and parametric estimators are almost

$p=(0.05,0.95)'$			$\theta_1=-0.4$			$\theta_1=0.4$		
α	β	r	$\zeta_p^{(r)}$	(i)	(ii)	$\zeta_p^{(r)}$	(i)	(ii)
1.2	0.0	20	4.905	4.916 (0.218)	34.18 [34.43]	5.689	5.703 (0.224)	36.02 [35.01]
1.2	0.0	720	7.834	7.856 (0.381)	104.43 [104.95]	8.680	8.707 (0.394)	111.64 [108.72]
1.2	0.5	20	4.918	4.923 (0.224)	36.03 [38.21]	5.704	5.721 (0.257)	47.51 [47.53]
1.2	0.5	720	7.849	7.860 (0.392)	110.85 [117.84]	8.695	8.725 (0.464)	154.83 [155.23]
1.5	0.0	20	4.043	4.046 (0.155)	17.35 [17.63]	4.828	4.824 (0.156)	17.51 [16.22]
1.5	0.0	720	6.379	6.387 (0.276)	54.96 [55.07]	7.225	7.217 (0.283)	57.64 [53.67]
1.5	0.5	20	4.055	4.055 (0.159)	18.27 [18.97]	4.841	4.850 (0.169)	20.45 [20.67]
1.5	0.5	720	6.393	6.393 (0.286)	58.93 [60.40]	7.238	7.253 (0.310)	69.24 [70.45]
1.8	0.0	20	3.511	3.507 (0.130)	12.14 [13.44]	4.296	4.285 (0.113)	9.12 [10.55]
1.8	0.0	720	5.450	5.444 (0.242)	42.19 [46.75]	6.296	6.273 (0.226)	36.65 [43.23]
1.8	0.5	20	3.515	3.509 (0.134)	12.97 [13.53]	4.300	4.301 (0.115)	9.59 [11.20]
1.8	0.5	720	5.455	5.446 (0.249)	44.74 [46.98]	6.300	6.301 (0.231)	38.41 [45.30]

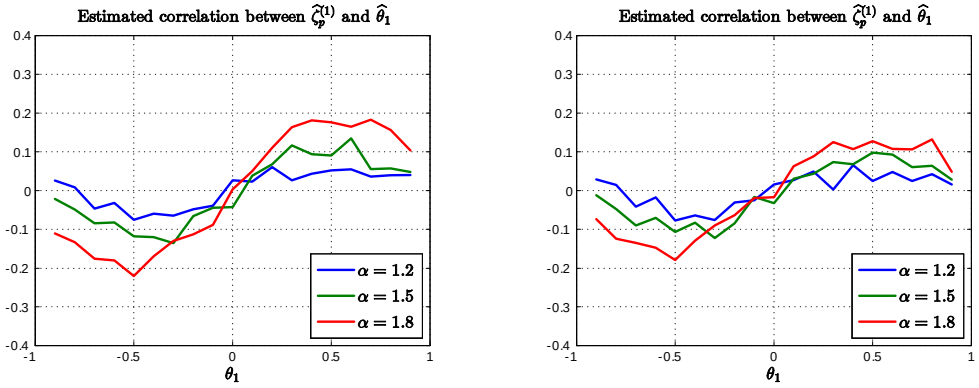
Table 4.9.2: Parametric estimation of the log quantile difference $\zeta_p^{(r)}$ for $p=(0.05,0.95)'$ at selected aggregation levels from samples of selected SMA(1) processes where θ_1 needs to be estimated. Included in this table are (i) the mean and standard deviation, in ( ), of the estimates across all realisations and (ii) the standardised variance, i.e. the variance multiplied by the sample size T, across all realisations and the estimated asymptotic variance, in [ ].

$p = (0.05, 0.95)'$					$\widehat{\zeta}_p^{(1)}$		$\widetilde{\zeta}_p^{(1)}$	
θ_1	α	β	$\zeta_p^{(1)}$	$\widetilde{R}_p^{(1)}$	(i)	(ii)	(i)	(ii)
0.2	1.2	0.0	2.974	1.000	2.980 (0.103)	7.672 [7.307]	2.980 (0.103)	7.678 [7.304]
0.2	1.2	0.5	2.988	1.000	2.990 (0.112)	8.973 [8.626]	2.990 (0.112)	8.974 [8.624]
0.2	1.5	0.0	2.559	1.000	2.560 (0.066)	3.175 [3.129]	2.560 (0.066)	3.175 [3.129]
0.2	1.5	0.5	2.573	1.000	2.575 (0.070)	3.534 [3.500]	2.575 (0.070)	3.534 [3.499]
0.4	1.2	0.0	3.100	1.000	3.113 (0.111)	8.866 [8.645]	3.113 (0.111)	8.874 [8.640]
0.4	1.2	0.5	3.115	1.000	3.125 (0.122)	10.713 [10.215]	3.125 (0.122)	10.715 [10.213]
0.4	1.5	0.0	2.652	1.000	2.655 (0.072)	3.700 [3.692]	2.655 (0.072)	3.700 [3.692]
0.4	1.5	0.5	2.666	1.000	2.668 (0.074)	3.937 [4.142]	2.668 (0.074)	3.938 [4.141]

Table 4.9.3: Simulation results for the non-parametric $\widehat{\zeta}_p^{(1)}$ and parametric $\widetilde{\zeta}_p^{(1)}$ log quantile difference estimators at $p = (0.05, 0.95)'$ from samples of selected SMA(1) processes. Included in this table are (i) the mean and standard deviation, in ( ), of the estimates across all realisations and (ii) the standardised variance and the asymptotic variance, in [ ].

$p=(0.10,0.90)'$					$\widehat{\zeta}_p^{(1)}$		$\widetilde{\zeta}_p^{(1)}$	
θ_1	α	β	$\zeta_p^{(1)}$	$\widetilde{R}_p^{(1)}$	(i)	(ii)	(i)	(ii)
0.2	1.2	0.0	2.407	0.993	2.407 (0.070)	3.485 [3.516]	2.410 (0.072)	3.690 [3.489]
0.2	1.2	0.5	2.451	0.937	2.450 (0.079)	4.465 [4.193]	2.451 (0.075)	4.046 [3.927]
0.2	1.5	0.0	2.167	0.936	2.166 (0.049)	1.735 [1.713]	2.168 (0.046)	1.527 [1.603]
0.2	1.5	0.5	2.189	0.938	2.189 (0.052)	1.930 [1.904]	2.189 (0.050)	1.829 [1.786]
0.4	1.2	0.0	2.534	0.995	2.539 (0.075)	4.069 [4.108]	2.542 (0.074)	3.973 [4.088]
0.4	1.2	0.5	2.578	0.934	2.582 (0.083)	4.925 [4.922]	2.585 (0.082)	4.817 [4.595]
0.4	1.5	0.0	2.260	0.950	2.260 (0.051)	1.890 [1.987]	2.262 (0.050)	1.824 [1.887]
0.4	1.5	0.5	2.282	0.952	2.282 (0.054)	2.100 [2.222]	2.283 (0.053)	2.013 [2.115]

Table 4.9.4: Simulation results for the non-parametric $\widehat{\zeta}_p^{(1)}$ and parametric $\widetilde{\zeta}_p^{(1)}$ log quantile difference estimators at $p=(0.10,0.90)'$ from samples of selected SMA(1) processes. Included in this table are (i) the mean and standard deviation, in ( ), of the estimates across all realisations and (ii) the standardised variance and the asymptotic variance, in [ ].



(a) (b)

Figure 4.9.1: Estimated correlation of the $\widehat{\zeta}_p^{(1)}$ and $\widehat{\theta}_1$ estimators from selected symmetric SMA(1) processes: (a) for $p = (0.05, 0.95)$ and (b) for $p = (0.50, 0.95)$.

identical for each process. In Tables 4.9.4 for quantile level $p = (0.10, 0.90)'$ and 4.9.5 for quantile level $p = (0.20, 0.80)'$ both elements of p are not included in p_M. Examples where the parametric estimator is 5-7% more efficient than the non-parametric estimator are included in Table 4.9.4. Examples where the parametric estimator is 4-6% less efficient than the non-parametric estimator are included in Table 4.9.5.

At present we have no method for calculating the asymptotic covariance of $\widehat{\zeta}_p^{(1)}$ and $\widehat{\theta}$ however, we can use simulations to estimate its value. Each simulation contains 2000 realisations of an SMA(1) process of length 720 where θ_1 takes values from the set $\{-0.9, -0.8, \ldots, 0.9\}$, $\omega^{(1)} = (\alpha, 0, 1, 0)$ and α takes values from the set $\{1.2, 1.5, 1.8\}$. For each realisation of a process, the estimates $\widehat{\zeta}_p^{(1)}$ and $\widehat{\theta}_1$ are calculated. The correlation of $\widehat{\zeta}_p^{(1)}$ and $\widehat{\theta}_1$ is calculated across all realisations of each process, (Figure 4.9.1).

The estimated correlation between $\widehat{\zeta}_p^{(1)}$ and $\widehat{\theta}_1$ increases with α and is most significant at $\theta_1 \approx \pm 0.5$. The estimated correlation is greater for $p = (0.05, 0.95)$ than for $p = (0.50, 0.95)$. For most processes the absolute value of the correlation is less than 0.2.

$p=(0.20,0.80)'$					$\widehat{\zeta}_p^{(1)}$		$\widetilde{\zeta}_p^{(1)}$	
θ_1	α	β	$\zeta_p^{(1)}$	$\widetilde{R}_p^{(1)}$	(i)	(ii)	(i)	(ii)
0.2	1.2	0.0	1.756	0.994	1.755 (0.056)	2.281 [2.207]	1.754 (0.057)	2.321 [2.194]
0.2	1.2	0.5	1.828	0.993	1.825 (0.061)	2.701 [2.641]	1.826 (0.061)	2.653 [2.622]
0.2	1.5	0.0	1.654	1.058	1.655 (0.043)	1.303 [1.502]	1.656 (0.045)	1.448 [1,589]
0.2	1.5	0.5	1.679	1.049	1.678 (0.047)	1.554 [1.602]	1.678 (0.049)	1.713 [1.681]
0.4	1.2	0.0	1.883	0.988	1.886 (0.057)	2.362 [2.514]	1.886 (0.057)	2.356 [2.484]
0.4	1.2	0.5	1.955	0.984	1.956 (0.066)	3.110 [3.049]	1.958 (0.065)	3.040 [3.001]
0.4	1.5	0.0	1.747	1.049	1.748 (0.047)	1.607 [1.690]	1.748 (0.049)	1.702 [1.777]
0.4	1.5	0.5	1.772	1.044	1.771 (0.050)	1.803 [1.815]	1.772 (0.051)	1.888 [1.895]

Table 4.9.5: Simulation results for the non-parametric $\widehat{\zeta}_p^{(1)}$ and parametric $\widetilde{\zeta}_p^{(1)}$ log quantile difference estimators at $p=(0.20,0.80)'$ from samples of selected SMA(1) processes. Included in this table are (i) the mean and standard deviation, in ( ), of the estimates across all realisations and (ii) the standardised variance and the asymptotic variance, in [ ].

Chapter 5

Measurement of realised volatility using log quantile differences

5.1 Introduction

Let O_d and C_d denote the opening and closing price respectively of a financial market asset on day d. Let

$$X_d = \ln\left(C_d/O_d\right) \tag{5.1.1}$$

denote the daily log return of the asset on day d. Elsewhere, the daily log return is used to refer to the log of the ratio of closing prices on consecutive days. For markets which trade throughout the day, these definitions are equivalent.

1) *Finite Variance Case*

The daily log returns $\{X_d\}$ of a financial market asset are modelled in the stochastic volatility (SV) model by the following equations

$$X_d = \sigma_d \varepsilon_d \tag{5.1.2}$$

and

$$\ln \sigma_d \sim \text{AR}\left(k\right) \tag{5.1.3}$$

where σ_d^2 is the conditional variance, $\{\varepsilon_d\}$ is a sequence of IID $N\left(0,1\right)$ random variables assumed to be independent of $\{\ln \sigma_d\}$ and k is the autoregressive order of the $\{\ln \sigma_d\}$ process,

(see for example Tsay (2010)). The sequence of random variables $\{X_d\}$ in (5.1.2) is said to be conditionally normal given σ_d, which we denote by

$$(X_d \mid \sigma_d) \sim N\left(0, \sigma_d^2\right). \tag{5.1.4}$$

Note that the generalised autoregressive conditional heteroscedastic (GARCH) model also has a conditionally normal distribution for X_d, but uses a different equation to model the conditional variance, σ_d^2. A comparison of two popular methods for estimation of the SV model, generalised method of moments (GMM) and quasi maximum likelihood (QML) can by found in Anderson and Sorensen (1997).

In the SV model, the sequences of random variables $\{\sigma_d\}$ and $\{\varepsilon_d\}$ are not observable, only the sequence, $\{X_d\}$, of daily log returns is observable. Taking the logarithm the absolute values of both sides of (5.1.2), gives

$$\ln|X_d| = \ln\sigma_d + \ln|\varepsilon_d|. \tag{5.1.5}$$

It can be shown that the variance of $\ln|\varepsilon_d|$ is given by

$$V\left[\ln|\varepsilon_d|\right] = \pi^2/8. \tag{5.1.6}$$

Now suppose that the intraday log returns $\{X_{d;t}\}$ are from the invertible MA(1) process,

$$X_{d;t} = e_{d;t} + \theta_1 e_{d;t-1} \tag{5.1.7}$$

where $\{e_{d;t}\}$ is an IID sequence of zero-mean finite variance random variables. Let

$$S_{d;t}^{(r)} = \sum_{i=0}^{r-1} X_{d;t-j} \tag{5.1.8}$$

denote the temporal aggregation of $\{X_{d;t}\}$ at aggregation level r. Let $\left(\sigma_d^{(r)}\right)^2$ denote the variance of $\left\{S_{d;t}^{(r)}\right\}$. If there are T intraday log returns in a day, then the following relationships exist between the daily log return and the temporal aggregation of the intraday log returns

$$X_d = S_{d;T}^{(T)}, \tag{5.1.9}$$

$$\sigma_d = \sigma_d^{(T)} \tag{5.1.10}$$

It can be shown that is given by

$$\left(\sigma_d^{(r)}\right)^2 = r\lambda_0 + 2\left(r-1\right)\lambda_1 \tag{5.1.11}$$

where λ_0, λ_1 are the autocovariance functions of $\{X_{d;t}\}$ at lags $0, 1$ respectively. Thus, we define an estimator $\ln \widehat{\sigma}_d^{(r)}$ of $\ln \sigma_d^{(r)}$ by

$$\ln \widehat{\sigma}_d^{(r)} = \frac{1}{2} \ln \left(r\widehat{\lambda}_0 + 2\left(r-1\right)\widehat{\lambda}_1 \right) \tag{5.1.12}$$

where $\widehat{\lambda}_0, \widehat{\lambda}_1$ are the sample autocovariance functions of $\{X_{d;t}\}$ at lags $0, 1$ respectively.

The following theorem shows that $\ln \widehat{\sigma}_d^{(r)}$ is an asymptotically normal estimator of $\ln \sigma_d^{(r)}$. Note that the asymptotic distribution in this theorem assumes that the aggregation level r is fixed as the sample n increases to ∞.

Theorem 5.1.1 *For a given d, let $\{X_{d;t}\}$ be the* MA(1) *process defined in* (5.1.11) *and $\varepsilon_{d;t} \sim N\left(0, \sigma^2_{\varepsilon;d}\right)$. Let λ_0, λ_1 denote the autocovariance function of $\{X_{d;t}\}$ at lags 0,1 respectively. Then as the sample size $n \to \infty$*

$$\sqrt{n}\left(\ln \widehat{\sigma}_d^{(r)} - \ln \sigma_d^{(r)} \right) \xrightarrow{d} N\left(0, D^{(r)} V_\lambda D^{(r)\prime}\right) \tag{5.1.13}$$

where

$$D^{(r)} = \frac{1}{r\lambda_0 + 2\left(r-1\right)\lambda_1} \left(r/2, r-1\right) \tag{5.1.14}$$

and

$$V_\lambda = \begin{pmatrix} 2\lambda_0^2 + 4\lambda_1^2 & 4\lambda_0\lambda_1 \\ 4\lambda_0\lambda_1 & \lambda_0^2 + 3\lambda_1^2 \end{pmatrix}. \tag{5.1.15}$$

Proof. See Appendix A.5. ■

Remark 5.1.1 *The limit as $r \to \infty$ of $D^{(r)} V_\lambda D^{(r)\prime}$ is given by*

$$\lim_{r\to\infty} D^{(r)} V_\lambda D^{(r)\prime} = 1 + \frac{\lambda_0^2}{2\left(\lambda_0 + 2\lambda_1\right)^2}. \tag{5.1.16}$$

Thus, for large r the variance of $\ln \widehat{\sigma}_d^{(r)}$ as an estimator of $\ln \sigma_d^{(r)}$ is $O\left(n^{-1}\right)$

The estimator $\widehat{\sigma}_d^{(r)}$ is sometimes referred to as the realised volatility, (French et al. (1987)). Realised volatility was also investigated in Anderson et al. (2001a) and Anderson et al. (2001b) where an allowance was included for the intraday log returns to have a non-zero mean.

The realised volatility is an estimate of the unobservable conditional standard deviation σ_d and can be used to estimate the parameters of the SV model assuming that the MA(1)

model for the intraday log returns is valid. For assets which have a large number of intraday returns, the realised volatility estimates of the conditional standard deviation σ_d should be more precise than those derived from the daily returns in (5.1.5) .

Inherent in the use of the SV model, is that the daily log returns and the intraday log returns are conditionally normal. Conditionally normal daily returns imply that the intraday returns are if not conditionally normal then at least conditionally of finite variance. Given the availability of intraday returns of sufficiently high frequency it is possible to test the conditional finite variance hypothesis on each day's intraday returns, see for example Chapter 5 of Dacorogna et al. (2001) which estimates the tail index of foreign exchange data over various time intervals.

2) *Stable Distribution Case*

As a generalisation of the SV model, which can accomodate both the success and failure of the conditional finite variance hypothesis, we propose the stable stochastic variance (SSV) model. Construction of the SSV model requires replacement of the IID sequence of normal random variables in (5.1.2) with an IID sequence of stable random variables. The SSV model equations are

$$X_d = \Gamma_d \varepsilon_d \tag{5.1.17}$$

and

$$\ln \Gamma_d \sim \text{AR}\,(k) \tag{5.1.18}$$

where $\{\varepsilon_d\}$ is a sequence of IID $S_\alpha\,(0,1,0)$ random variables assumed to be independent of $\{\ln \Gamma_d\}$ and k is the autoregressive order of the $\{\ln \Gamma_d\}$ process. The sequence of random variables $\{X_d\}$ in (5.1.17) is said to be conditionally stable given Γ_d, which we denote by

$$(X_d \mid \Gamma_d) \sim S^0_\alpha\,(0, \Gamma_d, 0)\,. \tag{5.1.19}$$

In this thesis, we do not address the properties of the unconditional distribution of $\{X_d\}$ in the SSV model.

Remark 5.1.2 *Where* $\alpha = 2$, *the* SSV *model in* (5.1.17) *and* (5.1.18) *is identical to the* SV *model in* (5.1.2) *and* (5.1.3) .

As in the SV model, the sequences of random variables $\{\Gamma_d\}$ and $\{\varepsilon_d\}$ are not observable, only the sequence, $\{X_d\}$, of daily log returns is observable. Taking the logarithm the absolute values of both sides of (5.1.17), gives

$$\ln|X_d| = \ln\Gamma_d + \ln|\varepsilon_d|. \tag{5.1.20}$$

Although stable distributions other than the Gaussian do not have a variance, the logarithm of the absolute value of a stable distribution can have a variance. In the case where $\alpha = 1$, it can be shown that the variance of $\ln|\varepsilon_d|$ is given by

$$V[\ln|\varepsilon_d|] = \pi^2/4. \tag{5.1.21}$$

The GMM and QML methods available for the estimation of the SV model are not immediately available for the estimation of the SSV model. Indeed most of the moments commonly used by the GMM would not exist for the SSV model where $\alpha < 2$. In this thesis, we do not attempt to adapt these methods for the estimation of the SSV model, but instead concentrate on using the realised volatility as a basis for estimation. However, where $\alpha < 2$ realised volatility as defined for the SV model does not exist for the SSV model. Instead, we create a new definition for realised volatility based on log quantile differences.

As with the SV model, we assume that the T intraday log returns are recorded each day over uniform time intervals. Suppose that the intraday log returns for each day are from an SMA(1) process, i.e.

$$X_{d;t} = e_{d;t} + \theta_1 e_{d;t-1} \tag{5.1.22}$$

where $\{e_{d;t}\}$ is an IID sequence of $S^0_\alpha(0,\gamma_d,0)$ random variables. Again we use $S^{(r)}_{d;t}$ to denote the temporal aggregation of $\{X_{d;t}\}$ at aggregation level r, that is

$$S^{(r)}_{d;t} = \sum_{i=0}^{r-1} X_{d;t-j} \tag{5.1.23}$$

From Theorem 2.2.1 we have that

$$S^{(r)}_{d;t} \sim S^0_\alpha\left(0,\gamma^{(r)}_d,0\right) \tag{5.1.24}$$

where

$$\ln\gamma^{(r)}_d = \alpha^{-1}\ln\left(\sum_{j=0}^{r+q-1}\left|c^{(r)}_j\right|^\alpha\right) + \ln\gamma_d \tag{5.1.25}$$

and

$$c_j^{(r)} = \begin{cases} 1 & j = 0 \\ 1+\theta_1 & j = 1, \ldots, r-1 \\ \theta_1 & j = r \end{cases} . \tag{5.1.26}$$

If there are T intraday log returns in a day, then the following relationships exist between the daily log return and the temporal aggregation of the intraday log returns

$$X_d = S_{d;T}^{(T)}, \tag{5.1.27}$$

$$\Gamma_d = \gamma_d^{(T)}. \tag{5.1.28}$$

Let $\zeta_{p;d}^{(r)}$ and $\zeta_{p;d}$ denote the LQD of $\left\{S_{d;t}^{(r)}\right\}$ and $\{e_{d;t}\}$ respectively at quantile levels p. Since $\{e_{d;t}\}$ is symmetric, we have from Theorem 2.2.2 that

$$\zeta_{p;d}^{(r)} = \ln\left(\gamma_d^{(r)}/\gamma_d\right) + \zeta_{p;d} \tag{5.1.29}$$

and in addition at aggregation level $r = T$ that

$$\zeta_{p;d}^{(T)} = \ln\left(\Gamma_d/\gamma_d\right) + \zeta_{p;d}. \tag{5.1.30}$$

Under the assumptions of the ssv model, the random variables $\zeta_{p;d}^{(T)}, \zeta_{p;d}, \ln\Gamma_d$ and $\ln\gamma_d$ are all separated by a constant. Let ζ_p^* denote the constant value of the LQD of the distribution $S_\alpha^0(0,1,0)$ at quantile levels p. Then from Lemma E.1.2, we have

$$\zeta_{p;d} = \ln\gamma_d + \zeta_p^* \tag{5.1.31}$$

and also

$$\zeta_{p;d}^{(T)} = \ln\Gamma_d + \zeta_p^*. \tag{5.1.32}$$

Rearranging (5.1.25) shows that $\ln\Gamma_d$ and $\ln\gamma_d$ are also separated by a constant. Although either $\ln\Gamma_d$ or $\zeta_{p;d}^{(T)}$ could be used as a definition for realised volatility, our preference is for $\zeta_{p;d}^{(T)}$ as it more directly relates to value at risk statistics.

For estimation purposes we assume that all of the intraday log return processes are independent of each other. In practice, it is often beneficial to fit an ARMA time series model to $\left\{\zeta_{p;d}^{(T)}\right\}$ rather than just an AR(k) as suggested in (5.1.18).

In Section 5.2, we adapt the aggregated LQD estimators discussed in Chapter 4 so that they can be used for estimation within the SSV model. In Section 5.3, methods are described for the estimation of the parameters of the $\left\{\zeta_{p;d}^{(T)}\right\}$ process in the presence of measurement error. In Section 5.4, simulation results are presented.

5.2 Estimation of realised volatility in a stable stochastic volatility model

In Section 5.1, it was suggested for the SSV model that $\zeta_{p;d}^{(T)}$, the LQD of the temporal aggregation of the SMA intraday log return process, be used as a measure of the realised volatility of the daily log returns on day d. Various methods for the estimation of $\zeta_{p;d}^{(r)}$ were proposed in Chapter 4, including the parametric LQD estimator, $\widetilde{\zeta}_{p;d}^{(r)}$, (Definition 4.3.2). The parametric LQD estimator was defined as a function of the parameters $\varkappa = \left(\alpha, \beta^{(1)}, \gamma_d^{(1)}, \theta_1\right)'$ of the base process, in this case the intraday log return process. However, the SSV model places additional constraints on the parameters $\varkappa$ which were not considered in the definition of parametric LQD estimator,

1. the distribution of each of the intraday log returns processes is symmetric i.e. $\beta^{(1)}$ is zero and

2. the α and θ_1 parameters are the same for each day d.

Let $\widehat{\varkappa}_d = \left(\widehat{\alpha}_d, \widehat{\beta}_d^{(1)}, \widehat{\gamma}_d^{(1)}, \widehat{\theta}_{1;d}\right)'$ denote the vector of estimators for the parameters of the intraday log return process $\{X_{d;t}\}$ calculated using the quantile-based stable distribution parameter estimators described in Section 3.3 and SLAD estimator in Appendix C.3. As discussed in Remark 3.3.2, we adapt the quantile-based stable distribution parameter estimators to allow for a known symmetric distribution, so that $\widehat{\beta}_d^{(1)} = 0$ for all d. Suppose we have a sample which contains D days of intraday log returns, then we define a new vector of estimators $\overline{\varkappa}_d = \left(\overline{\alpha}, \overline{\beta}, \widehat{\gamma}_d^{(1)}, \overline{\theta}_1\right)'$ which satisfies the constraints of the SSV model by

$$\overline{\alpha} = D^{-1} \sum_{d=1}^{D} \widehat{\alpha}_d, \tag{5.2.1}$$

$$\overline{\beta} = 0, \tag{5.2.2}$$

$$\overline{\gamma}_d^{(1)} = \frac{\widehat{\xi}_{p_{M_4};d} - \widehat{\xi}_{p_{M_2};d}}{\overline{\xi}^*_{p_{M_4}} - \overline{\xi}^*_{p_{M_2}}}, \tag{5.2.3}$$

$$\overline{\theta}_1 = D^{-1} \sum_{d=1}^{D} \widehat{\theta}_{1;d} \tag{5.2.4}$$

where

$$p_M = (p_{M_1}, p_{M_2}, p_{M_3}, p_{M_4}, p_{M_5}) \tag{5.2.5}$$

are the chosen quantile levels for the quantile-based estimation of $\widehat{\alpha}_d, \widehat{\beta}_d^{(1)}, \widehat{\gamma}_d^{(1)}$, $\widehat{\xi}_{p_{M_j};d}$ is the empirical quantile estimator of the $p_{M_j}th$ quantile from the intraday log returns sample on day d and $\overline{\xi}^*_{p_{M_j}}$ is the $p_{M_j}th$ quantile of the distribution $S^0_{\overline{\alpha}}(0,1,0)$.

As discussed in Section 4.2, as the sample size $T \to \infty$, the asymptotic properties of $\widehat{\varkappa}_d$ are

$$\sqrt{T}\left(\widehat{\varkappa}_d - \varkappa\right) \xrightarrow{d} N\left(0, V_{\widehat{\varkappa}_d}\right) \tag{5.2.6}$$

where

$$V_{\widehat{\varkappa}_d} = \begin{pmatrix} V_{\widehat{\omega}_d^{(1)}} & V_{\widehat{\omega}_d^{(1)},\widehat{\theta}_{1;d}} \\ V_{\widehat{\omega}_d^{(1)},\widehat{\theta}_{1;d}} & V_{\widehat{\theta}_{1;d}} \end{pmatrix}, \tag{5.2.7}$$

$$V_{\widehat{\omega}_d^{(1)}} = D_{\widehat{\omega}_d^{(1)}} V_{\widehat{\xi}_{p_M}} D_{\widehat{\omega}_d^{(1)}}, \tag{5.2.8}$$

$$D_{\widehat{\omega}_d^{(1)}} = \left(\frac{\partial \widehat{\omega}_{i;d}^{(1)}}{\partial \widehat{\xi}_{p_{M_j};d}} \Big|_{\widehat{\xi}_{p_{M_j};d} = \xi_{p_{M_j};d}} \right)_{i=1,\ldots,3, j=1,\ldots,5}, \tag{5.2.9}$$

$$\widehat{\omega}_d^{(1)} = \left(\widehat{\alpha}_d, \widehat{\beta}_d^{(1)}, \widehat{\gamma}_d^{(1)}\right), \tag{5.2.10}$$

$V_{\widehat{\theta}_{1;d}}$ is given in C.3.9 and $V_{\widehat{\xi}_{p_M;d}}$ is given in (3.2.10). There is no formula for $V_{\widehat{\omega}^{(1)},\widehat{\theta}_1}$ however, the simulation results in Section 4.4 suggest that we can assume $V_{\widehat{\omega}^{(1)},\widehat{\theta}_1} = 0$ with little effect on approximations of the asymptotic variance of $\zeta_{p;d}^{(T)}$.

Following a similar logic to the above, we get

$$\sqrt{T}\left(\overline{\varkappa}_d - \varkappa\right) \xrightarrow{d} N\left(0, V_{\overline{\varkappa}_d}\right) \tag{5.2.11}$$

where

$$V_{\overline{\varkappa}_d} = \begin{pmatrix} V_{\overline{\omega}_d^{(1)}} & V_{\overline{\omega}^{(1)},\overline{\theta}_1} \\ V_{\overline{\omega}_d^{(1)},\overline{\theta}_1} & V_{\overline{\theta}_1} \end{pmatrix}, \tag{5.2.12}$$

$$V_{\overline{\omega}^{(1)}} = D_{\overline{\omega}_d^{(1)}} V_{\widehat{\xi}_{p_M;d}} D_{\overline{\omega}_d^{(1)}}, \tag{5.2.13}$$

$$D_{\overline{\omega}_d^{(1)}} = \left(\frac{\partial \overline{\omega}_{d;i}^{(1)}}{\partial \widehat{\xi}_{p_{M_j};d}} \Big|_{\widehat{\xi}_{p_{M_j};d} = \xi_{p_{M_j};d}} \right)_{i=1,\ldots,3, j=1,\ldots,5} \tag{5.2.14}$$

and

$$\overline{\omega}_d^{(1)} = \left(\overline{\alpha}, \overline{\beta}^{(1)}, \overline{\gamma}_d^{(1)}\right). \tag{5.2.15}$$

Note that the same sample quantiles, $\widehat{\xi}_{p_M;d}$, are used in the calculation of both $\widehat{\omega}_d^{(1)}$ and $\overline{\omega}_d^{(1)}$. By assumption, the intraday log return processes are independent of each other. Therefore

the sequences of estimators $\left\{\widehat{\theta}_{1;d}\right\}$ and $\{\widehat{\alpha}_d\}$ are also independent and we have

$$V_{\overline{\theta}_1} = D^{-1}V_{\widehat{\theta}_{1;d}}, \tag{5.2.16}$$

$$\frac{\partial\overline{\alpha}}{\partial\widehat{\xi}_{p_{M_j};d}} = D^{-1/2}\frac{\partial\widehat{\alpha}_d}{\partial\widehat{\xi}_{p_{M_j};d}}. \tag{5.2.17}$$

As discussed in Section 3.3, no general formulae exist for the partial derivatives in (5.2.14). For the moderately large values of D used in this thesis, it is useful to approximate $\dfrac{\partial\overline{\gamma}_d^{(1)}}{\partial\widehat{\xi}_{p_{M_j};d}}$ by the values it would take were α known, i.e.

$$\frac{\partial\overline{\gamma}_d^{(1)}}{\partial\widehat{\xi}_{p_{M_j};d}} \approx \begin{cases} 0 & j = 1,3,5 \\ \dfrac{-1}{\overline{\xi}^*_{p_{M_4}} - \overline{\xi}^*_{p_{M_2}}} & j = 2 \\ \dfrac{1}{\overline{\xi}^*_{p_{M_4}} - \overline{\xi}^*_{p_{M_2}}} & j = 4 \end{cases}. \tag{5.2.18}$$

To calculate an estimator $\overline{\zeta}^{(r)}_{p;d}$ for the log quantile difference $\zeta^{(r)}_{p;d}$ from the estimator $\overline{\omega}^{(1)}_d$, we adopt the same approach described in Sections 4.2 and 4.3 to calculate the parametric LQD estimator $\widetilde{\zeta}^{(r)}_{p;d}$ from $\widehat{\omega}^{(1)}_d$. We refer to $\overline{\zeta}^{(r)}_{p;d}$ as the mean symmetric parametric LQD estimator. The estimators $\overline{\omega}^{(r)}_d$ and $\overline{\xi}^{*(r)}_{p_j;d}$ are defined analogously to the aggregated stable distribution parameter estimator, $\widehat{\omega}^{(r)}_d$ (Definition 4.2.1), and the parametric quantile estimator, $\widetilde{\xi}^{*(r)}_{p_j;d}$ (Definition 4.3.1), respectively.

The following theorem proves that the asymptotic variance of the $\overline{\zeta}^{(r)}_{p;d}$ does not depend on $\gamma^{(1)}_d$. This result affects the available choices of estimation method for the parameters of the ARMA$(1,1)$ model for $\left\{\overline{\zeta}^{(T)}_{p;d}\right\}$, (see Section 5.3).

Theorem 5.2.1 *The asymptotic variance, $V_{\overline{\zeta}^{(r)}_{p;d}}$, of the mean symmetric parametric* LQD *estimator, $\overline{\zeta}^{(r)}_{p;d}$, is the same for all values of $\gamma^{(1)}_d > 0$.*

Proof. See Appendix A.6. ■

5.3 Estimation of log quantile difference processes in the presence of measurement error

In practice, data from the true log quantile difference process $\left\{\zeta^{(r)}_{p;d}\right\}$ is not observable. Instead, we can calculate an estimated log quantile difference process $\left\{\overline{\zeta}^{(r)}_{p;d}\right\}$ using the

methods described in Section 5.2. The relationship between the true and estimated log quantile difference processes is given by

$$\overline{\zeta}_{p;d}^{(r)} = \zeta_{p;d}^{(r)} + u_{p;d}^{(r)} \tag{5.3.1}$$

where the asymptotic distribution of $u_{p;d}^{(r)}$ is derived in Section 5.2. If the true process, $\left\{\zeta_{p;d}^{(r)}\right\}$, is an ARMA process, then using data from the estimated process, $\left\{\overline{\zeta}_{p;d}^{(r)}\right\}$, can introduce a bias into the parameter estimates unless an allowance is made for the presence of measurement error.

To illustrate the effect of measurement error on ARMA process parameter estimation, let us examine a simple case in more detail. Suppose $\{X_d\}$ is the unobservable stationary AR(1) process defined by

$$X_d = \phi_1 X_{d-1} + u_d \tag{5.3.2}$$

where $\{u_d\}$ is IID with zero mean and variance $\sigma_u^2 > 0$ and Y_d is the observable process

$$Y_d = X_d + w_d \tag{5.3.3}$$

where the measurement error $\{w_d\}$ is IID with zero mean and variance $\sigma_w^2 > 0$. We assume that $\{X_d\}$ and $\{w_d\}$ are uncorrelated. Let $\lambda_X(h)$ and $\lambda_Y(h)$ denote the autocovariance functions at lag h of the processes $\{X_d\}$ and $\{Y_d\}$ respectively. Multiplying (5.3.2) by itself at lag h and taking expectations we get

$$\begin{aligned} E[Y_d Y_{d-h}] &= E[X_d X_{d-h}] + E[X_d w_{d-h}] + E[w_d X_{d-h}] + E[w_d w_{d-h}] \\ &= E[X_d X_{d-h}] + E[w_d w_{d-h}]. \end{aligned} \tag{5.3.4}$$

Expressing (5.3.4) in terms of autocovariance functions gives

$$\lambda_Y(h) = \begin{cases} \lambda_X(h) + \sigma_w^2 & \text{for } h = 0 \\ \lambda_X(h) & \text{for } h \neq 0 \end{cases}. \tag{5.3.5}$$

The following relationships exist between the model parameters and the autocovariance function of the unobservable process $\{X_d\}$,

$$\phi_1 = \frac{\lambda_X(1)}{\lambda_X(0)}, \tag{5.3.6}$$

$$\sigma_u^2 = \lambda_X(0) - \frac{\lambda_X^2(1)}{\lambda_X(0)}. \tag{5.3.7}$$

Substitution of the sample autocovariance function, $\widehat{\lambda}_X(h)$, of $\{X_d\}$ into (5.3.6) and (5.3.7) would allow the calculation of estimators for ϕ_1 and σ_u^2. However, since $\{X_d\}$ is unobservable we are unable to calculate $\widehat{\lambda}_X(h)$ directly. The naive substitution of the sample autocovariance function, $\widehat{\lambda}_Y(h)$, of the observable process $\{Y_d\}$ into (5.3.6) and (5.3.7) would introduce a bias into the estimators for ϕ_1 and σ_u^2 due to the fact that $\lambda_Y(0) \neq \lambda_X(0)$.

The properties of the sums of uncorrelated ARMA processes are well known and can be used in the estimation of ARMA processes affected by measurement error, (Theorem C.2.1). In particular it can be shown that the sum of uncorrelated ARMA processes is also an ARMA process. However, for a given ARMA(p,q) process $\{Y_d\}$ and ARMA(r,s) process $\{w_d\}$ there may not exist an ARMA(k,m) process $\{X_d\}$ such that $Y_d = X_d + w_d$ even though the process orders satisfy Theorem C.2.1. See Granger and Morris (1976) for examples of this, including the simple case where $\{X_d\}$ is AR(1), $\{w_d\}$ is IID and $\{Y_d\}$ is ARMA$(1,1)$.

Thus, if the unobserved true process is an ARMA process and the unobserved measurement error is also an ARMA process, not correlated with the true process, then the observed estimated process is also an ARMA process. The ARMA parameters of the observed estimated process can be estimated using standard methods. It then remains to determine estimates for the ARMA parameters of the unobserved processes from the estimates for the ARMA parameters of the observed process. A method was proposed in Pagano (1974) which covers the case where the true process is an AR process and the measurement error is IID. This was extended by Miazaki and Dorea (1993) to allow the noise to be a MA process. The restricted maximum likelihood (RML) method proposed by Lee and Shin (1997), allows both the true process and the measurement error to be ARMA processes.

Assume that the order of the measurement error process $\{w_d\}$ is known and an appropriate selection is made for the orders of the true process $\{X_d\}$ and the estimated process $\{Y_d\}$. The RML method uses Newton-Raphson optimisation to identify the maximum likelihood model for $\{Y_d\}$ subject to the restriction that the corresponding model for $\{X_d\}$ has the selected order. We refer to this model for $\{Y_d\}$ as the restricted model. The maximum likelihood model for $\{Y_d\}$ not subject to any restriction on the order of the corresponding model for $\{X_d\}$ is referred to as the unrestricted model. In the case where $\{X_d\}$ is AR(1), $\{w_d\}$ is IID and $\{Y_d\}$ is ARMA$(1,1)$, the restricted and unrestricted models for $\{Y_d\}$ are unlikely to be the same. However, in the case where $\{X_d\}$ is ARMA$(1,1)$, $\{w_d\}$ is IID and

$\{Y_d\}$ is ARMA$(1,1)$, the restricted and unrestricted models for $\{Y_d\}$ usually are the same. Moreover, we can obtain formulae for the parameters of the model for $\{X_d\}$ in terms of the parameters of the model for $\{Y_d\}$. We use these formulae extensively in Chapter 6.

Suppose that the true process $\{X_d\}$ is ARMA$(1,1)$ with the following model equation

$$X_d - \mu = \phi_1 (X_{d-1} - \mu) + u_d + \psi_1 u_{d-1} \tag{5.3.8}$$

where $\{u_d\}$ is an IID sequence of $N\left(0, \sigma_u^2\right)$ random variables. Further suppose that the measurement error $\{w_d\}$ of the estimated process $\{Y_d\}$ be an IID sequence of $N\left(0, \sigma_w^2\right)$ random variables where σ_w^2 is assumed known, so that

$$Y_d = X_d + w_d. \tag{5.3.9}$$

We assume that $\{X_d\}$ and $\{w_d\}$ are uncorrelated. Combining (5.3.8) and (5.3.9) gives

$$Y_d - \mu = \phi_1 (Y_{d-1} - \mu) + u_d + \psi_1 u_{d-1} + w_d - \phi_1 w_{d-1} \tag{5.3.10}$$

where the last four terms on the RHS of (5.3.10) form an MA(1) process. Let $\{v_d\}$ be the MA(1) process such that

$$v_d + \eta_1 v_{d-1} = u_d + \psi_1 u_{d-1} + w_d - \phi_1 w_{d-1} \tag{5.3.11}$$

where $\{v_d\}$ is an IID sequence of $N\left(0, \sigma_v^2\right)$ random variables. The parameters of the $\{v_d\}$ process are determined by the requirement that both sides of (5.3.11) have the same autocovariance function at all lags. The ARMA$(1,1)$ model equation for $\{Y_d\}$ becomes

$$Y_d - \mu = \phi_1 (Y_{d-1} - \mu) + v_d + \eta_1 v_{d-1}. \tag{5.3.12}$$

The parameters μ and ϕ_1 are common to the ARMA$(1,1)$ models of both $\{X_d\}$ and $\{Y_d\}$. To obtain formulae for the remaining parameters ψ_1 and σ_u^2 of the model for $\{X_d\}$ in terms of the parameters of the model for $\{Y_d\}$, we equate the autocovariances at lags 0 and 1 of both sides of (5.3.11)

$$\begin{aligned} &\text{Lag 0:} \quad \left(1 + \eta_1^2\right) \sigma_v^2 = \left(1 + \psi_1^2\right) \sigma_u^2 + \left(1 + \phi_1^2\right) \sigma_w^2, \\ &\text{Lag 1:} \quad \eta_1 \sigma_v^2 = \psi_1 \sigma_u^2 - \phi_1 \sigma_w^2. \end{aligned} \tag{5.3.13}$$

Solving (5.3.13) for ψ_1 and σ_u^2 gives

$$\psi_1 = \begin{cases} \dfrac{\kappa_0/\kappa_1 + \sqrt{(\kappa_0/\kappa_1)^2 - 4}}{2} & \text{if } \kappa_1 < 0 \\ 0 & \text{if } \kappa_1 = 0 \\ \dfrac{\kappa_0/\kappa_1 - \sqrt{(\kappa_0/\kappa_1)^2 - 4}}{2} & \text{if } \kappa_1 > 0 \end{cases} \tag{5.3.14}$$

$$\sigma_u^2 = \frac{\kappa_0}{1+\psi_1^2} \tag{5.3.15}$$

where

$$\kappa_0 = \left(1+\eta_1^2\right)\sigma_v^2 - \left(1+\phi_1^2\right)\sigma_w^2, \tag{5.3.16}$$

$$\kappa_1 = \eta_1\sigma_v^2 + \phi_1\sigma_w^2. \tag{5.3.17}$$

Substitution of the maximum likelihood estimators $\widehat{\Theta} = \left(\widehat{\mu}, \widehat{\phi}_1, \widehat{\eta}_1, \widehat{\sigma}_v^2\right)'$ for the parameters $\Theta = \left(\mu, \phi_1, \eta_1, \sigma_v^2\right)'$ of the observable process $\{Y_d\}$ into the equations (5.3.14) - (5.3.17) defines a set of estimators $\widehat{\Omega} = \left(\widehat{\mu}, \widehat{\phi}_1, \widehat{\psi}_1, \widehat{\sigma}_u^2\right)'$ of the parameters $\Omega = \left(\mu, \phi_1, \psi_1, \sigma_u^2\right)'$ of the unobservable process $\{X_d\}$. We call these estimators $\widehat{\Omega}$, the Corrected Maximum Likelihood (CML) estimators. The asymptotic covariance of the CML estimators is given in the following theorem.

Theorem 5.3.1 *Let $\{X_d\}, \{u_d\}, \{Y_d\}$ and $\{w_d\}$ be the processes described in (5.3.8) and (5.3.9). Let Ω_0 denote the true values of the parameters in the model for $\{X_d\}$. If $\sigma_w^2, \sigma_v^2 > 0$ and $\phi_1 \neq -\eta_1$, then as the sample size $T \to \infty$,*

$$\sqrt{T}\left(\widehat{\Omega} - \Omega_0\right) \xrightarrow{d} N\left(0, DVD'\right) \tag{5.3.18}$$

where

$$V = \begin{pmatrix} \sigma_v^2\dfrac{(1+\eta_1)^2}{(1-\phi_1)^2} & 0 & 0 \\ 0 & \begin{pmatrix} \left(1-\phi_1^2\right)^{-1} & \left(1+\phi_1\eta_1\right)^{-1} \\ \left(1+\phi_1\eta_1\right)^{-1} & \left(1-\eta_1^2\right)^{-1} \end{pmatrix}^{-1} & 0 \\ 0 & 0 & 2\left(\sigma_v^2\right)^2 \end{pmatrix}, \tag{5.3.19}$$

$$D = \left(\frac{\partial\Omega_i}{\partial\Theta_j}\,|_{\widehat{\Omega}=\Omega_0}\right)_{i,j=1,\ldots,4} \tag{5.3.20}$$

Proof. See Appendix A.7. ■

Remark 5.3.1 *The motivation for the use of* CML *rather than* RML *estimators is purely computational. For* CML *estimation, any standard* ARMA *estimation software can be used to calculate estimates for the parameters of the observed process, after which straightforward calculations are used to calculate estimates for the parameters of the unobserved process. For* RML *estimation, the need for restricted optimisation makes the use of standard* ARMA *estimation software more difficult.*

Variations of Theorem 5.3.1 can be written where the processes have a different order however, closed form formulae such as those in equations (5.3.14) - (5.3.17) may not exist. Instead numerical values for the ARMA model parameters of $\{Y_d\}$ can be calculated using the Wilson Factorisation Algorithm. (Wilson (1969), see also Appendix C.2)

Remark 5.3.2 *The* CML *method is only appropriate where maximum likelihood estimators of the parameters of* $\{Z_d\}$ *can be transformed into an* ARMA *process for* $\{Y_d\}$ *which exists and is of the correct order. In general, the* CML *method is not appropriate where the true process is* AR(1)*, the measurement error is* IID *and the observed process is* ARMA$(1,1)$*.In this case the* CML *method will usually result in an* ARMA$(1,1)$ *model for the true process rather than an* AR(1) *model as required. In this case we can instead use the* RML *estimation method.*

Several other methods have been proposed for the estimation of time series models in the presence of measurement error. The modified Yule-Walker method (Walker (1960)) applies to AR time series models with heteroscedastic measurement error of unknown variance. The Staudenmayer-Buonaccorsi method (Staudenmayer and Buonaccorsi (2005)) applies to AR time series models with heteroscedastic measurement error of known or at least estimable variance. Generalised method of moments estimation (Melino and Turnbull (1990)) is commonly used for the estimation of SV models where the measurement error has a known log normal distribution. State space models (e.g. Durbin and Koopman (2001)) provide solutions to many measurement error time series estimation problems. A comparison of some of these methods can be found in Buonaccorsi (2010).

5.4 Simulation

In this section the results of simulations are presented which demonstrate the estimation of the parameters of an SSV model using CML on the log quantile difference realised volatility. Each days trading is divided into T equally spaced periods and the intraday log returns $\{X_{d;t}\}$, $t = 1, \ldots, T$ on day d follow the SMA(1) process

$$X_{d;t} = e_{d,t} + \theta_1 e_{d,t-1} \tag{5.4.1}$$

where $\{e_{d,t}\}$ is a sequence of symmetric IID random variables with distribution

$$e_{d;t} \sim S^0_\alpha (0, \gamma_d, 0) . \tag{5.4.2}$$

The daily log returns $\{X_d\}$ are the temporal aggregation of the intraday returns

$$X_d = \sum_{t=1}^{T} X_{d;t} \tag{5.4.3}$$

and are conditionally stable,

$$(X_d \mid \mathcal{F}_{d-1}) \sim S^0_\alpha \left(0, \gamma_d^{(T)}, 0\right), \tag{5.4.4}$$

and $\mathcal{F}_{d-1}$ denotes the sigma field of all information available at the end of day $d-1$. From Theorem 2.2.1,

$$\gamma_d^{(T)} = \left(\sum_{j=0}^{T} \left| c_j^{(T)} \right|^\alpha \right)^{1/\alpha} \gamma_d, \tag{5.4.5}$$

where

$$c_j^{(T)} = \begin{cases} 1 & j = 0 \\ 1 + \theta_1 & j = 1, \ldots, T-1 \\ \theta_1 & j = T \end{cases} . \tag{5.4.6}$$

Let $\zeta_{p;d}^{(T)}$ denote the log quantile difference of the daily log returns, $\{X_d\}$, at quantile level $p = (p_1, p_2)$. The model equation of the ARMA(1,1) process for $\left\{\zeta_{p;d}^{(T)}\right\}$ is

$$\zeta_{p;d}^{(T)} - \mu = \phi_1 \left(\zeta_{p;d-1}^{(T)} - \mu \right) + u_d + \psi_1 u_{d-1} \tag{5.4.7}$$

where $\{u_d\}$ is an IID sequence of $N\left(0, \sigma_u^2\right)$ random variables. The following parameters

were chosen for the $\left\{\zeta_{p;d}^{(T)}\right\}$ process

$$\mu = -2, \tag{5.4.8}$$

$$\phi_1 = 0.9, \tag{5.4.9}$$

$$\psi_1 = -0.5, \tag{5.4.10}$$

$$\sigma_u^2 = 0.01. \tag{5.4.11}$$

For each round of this simulation a single realisation of the process $\left\{\zeta_{p;d}^{(T)}\right\}$ is generated of length $D = 500$. In addition, D realisations of the process $\{X_{d;t}\}$ are generated each of length T. The objective of this simulation is to attempt to estimate the parameters μ, ϕ_1, ψ_1 and σ_v^2 of the unobserved daily volatility process $\left\{\zeta_{p;d}^{(T)}\right\}$ from the observed intraday log returns $\{X_{d;t}\}$.

The mean symmetric parametric log quantile difference estimator $\overline{\zeta}_{p;d}^{(r)}$ is used to estimate $\zeta_{p;d}^{(r)}$ at $r = T$. Although the asymptotic properties of $\overline{\zeta}_{p;d}^{(r)}$, established in Section 5.2, assume that the aggregation level r is fixed as the sample size $T \to \infty$, nevertheless, for these simulations, the asymptotic variance still provides a reasonable estimate for the variance of the estimator $\overline{\zeta}_{p;d}^{(T)}$. Let

$$\overline{\zeta}_{p;d}^{(T)} = \zeta_{p;d}^{(T)} + w_d \tag{5.4.12}$$

where w_d is the measurement error with variance σ_w^2. From Theorem 5.2.1, we know that $V_{\overline{\zeta}_{p;d}^{(r)}}$, the asymptotic variance of $\overline{\zeta}_{p;d}^{(r)}$, has the same value for all values of $\gamma_d > 0$. For the purposes of this simulation, we approximate the σ_w^2 by

$$\sigma_w^2 \approx T^{-1} V_{\overline{\zeta}_p^{(T)}} \tag{5.4.13}$$

where $V_{\overline{\zeta}_p^{(T)}}$ is the asymptotic variance of $\overline{\zeta}_{p;d}^{(r)}$ at $r = d$.

The quantile levels chosen for stable distribution parameter estimation in this simulation are

$$p_M = (0.040, 0.200, 0.500, 0.800, 0.960). \tag{5.4.14}$$

For the $\overline{\zeta}_{p;d}^{(r)}$ estimator, α and θ_1, being the mean of D independent estimates, are very precisely estimated and the size of the asymptotic variance $V_{\overline{\zeta}_p^{(r)}}$ depends largely on the precision of the $\gamma_d^{(1)}$ estimates. From Table 3.5.7, it can be seen that the choice of $p_{M_2} = 0.200$ provides better estimates of $\gamma_d^{(1)}$ than the standard choice of $p_{M_2} = 0.250$ where

$\alpha > 1.2$ is known, β is known to be zero and the sample is IID. From Table 3.5.1, it can be seen that the choice of $p_{M_1} = 0.040$ provides better estimators of α than the standard choice of $p_{M_1} = 0.050$ where $\alpha > 1.2, \beta$ is zero and the sample is IID. Similar statements can be made where θ_1 is non-zero and say, less than 0.5. In choosing the quantile levels p_M, one needs to be wary that:

1. due to a slower convergence of the quantile estimators to their asymptotic distribution, the choice of quantile levels too close to zero and one may result in estimators with a small sample variance much greater than the asymptotic variance,

2. a choice of p_{M_2} which is too close to p_{M_1} tends to increase the asymptotic variance of the estimator of α.

Rather than attempt to find an optimal choice for p_M for each simulation, we use the value of p_M in (5.4.14) for all simulations. For each simulation, the choice of p_M in (5.4.14) has a lower value for σ_u^2 than the standard quantile levels.

Results of these simulations are recorded in Tables 5.4.1 and 5.4.2. Generally, the mean of each parameter is within one standard deviation of the true parameter value. The measurement error variance , σ_w^2, is naturally higher by a factor of four at $T = 180$ than at $T = 720$. The value of σ_w^2, is also higher for $\alpha = 1.4$ than for $\alpha = 1.7$ and higher for $\theta_1 = 0.4$, than for $\theta_1 = 0.2$ or -0.2 which have approximately the same measurement error. The higher the measurement error variance the higher the variance in the parameter estimates. The highest value for the measurement error variance in these simulations is $\sigma_w^2 = 0.0106$ which is slightly higher than $\sigma_u^2 = 0.0100$.

The variance of the ϕ_1 parameter is much higher than the asymptotic value. This is largely due to a few outliers which have values up to four standard deviations less than the true value.

			μ		ϕ_1		ψ_1	
α	θ_1	T	(i)	(ii)	(i)	(ii)	(i)	(ii)
1.4	-0.2	180	-1.997 (0.030)	0.435 [0.259]	0.880 (0.037)	0.683 [0.514]	-0.475 (0.093)	4.345 [3.569]
1.4	-0.2	720	-2.001 (0.024)	0.293 [0.252]	0.880 (0.037)	0.685 [0.400]	-0.479 (0.063)	1.952 [1.905]
1.4	0.2	180	-1.983 (0.030)	0.452 [0.259]	0.876 (0.041)	0.832 [0.514]	-0.468 (0.090)	4.079 [3.566]
1.4	0.2	720	-1.998 (0.025)	0.304 [0.252]	0.880 (0.038)	0.707 [0.400]	-0.475 (0.063)	1.964 [1.905]
1.4	0.4	180	-1.980 (0.030)	0.443 [0.261]	0.876 (0.040)	0.783 [0.536]	-0.466 (0.089)	3.939 [3.932]
1.4	0.4	720	-1.997 (0.025)	0.304 [0.253]	0.879 (0.039)	0.747 [0.406]	-0.473 (0.064)	2.018 [1.981]
1.7	-0.2	180	-2.020 (0.027)	0.361 [0.257]	0.881 (0.036)	0.654 [0.470]	-0.498 (0.074)	2.709 [2.882]
1.7	-0.2	720	-2.016 (0.024)	0.294 [0.252]	0.880 (0.037)	0.697 [0.388]	-0.485 (0.060)	1.812 [1.751]
1.7	0.2	180	-2.014 (0.026)	0.346 [0.257]	0.879 (0.039)	0.743 [0.470]	-0.492 (0.081)	3.289 [2.877]
1.7	0.2	720	-2.014 (0.024)	0.281 [0.252]	0.880 (0.038)	0.711 [0.388]	-0.482 (0.061)	1.874 [1.749]
1.7	0.4	180	-2.009 (0.026)	0.347 [0.257]	0.878 (0.039)	0.761 [0.484]	-0.489 (0.081)	3.284 [3.089]
1.7	0.4	720	-2.013 (0.024)	0.291 [0.252]	0.881 (0.040)	0.787 [0.392]	-0.483 (0.062)	1.947 [1.796]

Table 5.4.1: Simulation results for the CML estimators of the log quantile difference process parameters μ, ϕ_1 and ψ_1. Included in this table are (i) the mean and standard deviation, in ( ), of the estimates across all realisations and (ii) the standardised variance and the asymptotic variance, in [ ].

			σ_w^2		σ_u^2	
α	θ_1	T	(i)	(ii)	(ii)	(iii)
1.4	−0.2	180	0.0092	0.0092 (9.5×10^{-5})	0.0100 (0.0014)	1.0×10^{-3} $[9.3\times10^{-4}]$
1.4	−0.2	720	0.0023	0.0023 (1.2×10^{-5})	0.0099 (0.0008)	3.6×10^{-4} $[3.3\times10^{-4}]$
1.4	0.2	180	0.0092	0.0092 (1.0×10^{-4})	0.0100 (0.0014)	1.0×10^{-3} $[9.3\times10^{-4}]$
1.4	0.2	720	0.0023	0.0023 (1.2×10^{-5})	0.0098 (0.0008)	3.3×10^{-4} $[3.3\times10^{-4}]$
1.4	0.4	180	0.0105	0.0106 (1.3×10^{-4})	0.0099 (0.0015)	1.2×10^{-3} $[1.1\times10^{-3}]$
1.4	0.4	720	0.0027	0.0027 (1.5×10^{-5})	0.0097 (0.0008)	3.3×10^{-4} $[3.5\times10^{-4}]$
1.7	−0.2	180	0.0065	0.0064 (5.8×10^{-5})	0.0108 (0.0012)	7.5×10^{-4} $[6.6\times10^{-4}]$
1.7	−0.2	720	0.0016	0.0016 (8.1×10^{-6})	0.0102 (0.0008)	3.3×10^{-4} $[2.9\times10^{-4}]$
1.7	0.2	180	0.0065	0.0064 (5.5×10^{-5})	0.0107 (0.0013)	8.4×10^{-4} $[6.6\times10^{-4}]$
1.7	0.2	720	0.0016	0.0016 (8.1×10^{-6})	0.0100 (0.0008)	3.2×10^{-4} $[2.9\times10^{-4}]$
1.7	0.4	180	0.0073	0.0073 (7.0×10^{-5})	0.0106 (0.0014)	9.3×10^{-4} $[7.4\times10^{-4}]$
1.7	0.4	720	0.0018	0.0018 (9.4×10^{-6})	0.0100 (0.0008)	3.5×10^{-4} $[3.0\times10^{-4}]$

Table 5.4.2: Simulation results for the measurement error σ_w^2 and the CML estimator of the log quantile difference process parameter σ_u^2. Included in this table are (i) the true asymptotic measurement error, (ii) the mean and standard deviation, in ( ), of the estimates across all realisations and (iii) the standardised variance and the asymptotic variance, in [ ].

Chapter 6

Application

6.1 Introduction

In this chapter we apply the models developed in previous chapters to actual financial market data. In particular we attempt to create time series models for sequences of intraday returns and show how these models may provide an insight into the volatility of daily returns. The financial market asset used in this chapter is the ASX200 index of the Australian Stock Exchange (ASX). The ASX200 index reflects the combined market capitalisation of the leading 200 companies listed on the ASX. The ASX200 index data used in this chapter has been collected over an observation period from 2 Jan 2009 to 31 December 2010, a total of $D = 507$ trading days (Figure 6.1.1). The data was provided by Thomson Reuters (https://tickhistory.thomsonreuters.com/TickHistory/login.jsp).

Trading hours for the ASX are 10am to 4pm, Monday to Friday, except for public holidays when the ASX is closed and for Christmas Eve and Near Years Eve when trading hours are from 10am to 2pm. Data for the ASX200 index is available at 30 second intervals during trading hours, resulting in 721 prices and 720 intraday log returns on a normal trading day. An example of a single day's intraday log returns over 30 second intervals is plotted in Figure 6.1.2.

Figure 6.1.1: Daily closing price of ASX200 index from 2 Jan 2009 to 31 Dec 2010.

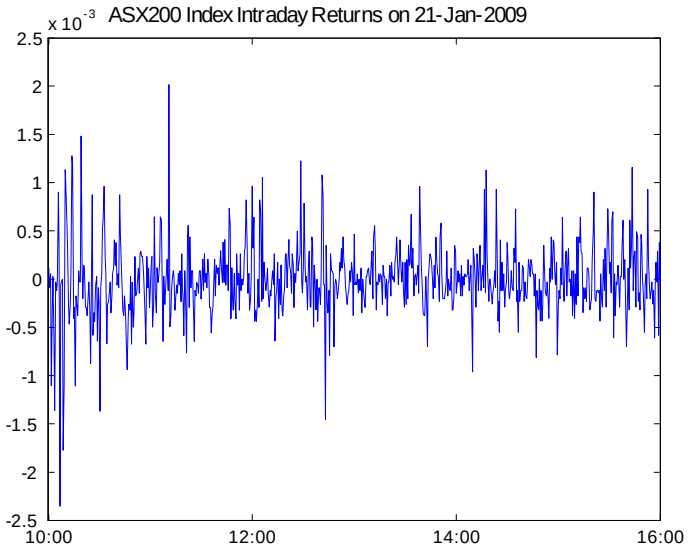


Figure 6.1.2: Intraday returns over 30 second intervals of ASX200 index on 21 Jan 2009.

6.2 Data cleaning

The data is provided by Thomson Reuters in CSV file format. Each line in the file includes date, time and price fields in addition to other fields which we do not use. Approximately every 30 seconds, there is a pair of records in the CSV file. Both records in the pair always contain the same price data. Data included in these files for times outside normal trading hours is not used in this analysis. Cleaning of the data was required to address the following issues:

1. Periods of missing data,

2. Periods of frozen data,

3. Discreteness of data.

Where data is missing it is usually not for more than one consecutive 30 second interval. If the missing data was in the middle of the day, then we used linear interpolation on the adjacent data to calculate an estimate for the missing data. If the missing data was at the beginning or end of the day, or the gap too long for any sort of interpolation, then it was ignored and the base process sample for that day is considered to be shorter than usual.

Where the ASX200 index has the same value, correct to one decimal place, for a number of consecutive 30 second intervals, it is said a period of frozen data exists. Whilst it is possible that periods of frozen data can occur legitimately, the volatility of the ASX200 index is sufficient, that such periods are unlikely to last for more than a few consecutive intervals.

Somewhat arbitrarily, we decided that any period of frozen data longer than 2 minutes, be considered to be caused by some error at the ASX and not a true reflection of the price of the ASX200 index. Over the two year period 2009-2010, five periods of frozen data longer than two minutes were found. Each of these periods were removed from the base process sample for that day.

After data cleaning, the length of the base process samples for the two year period 2009-2010 were

Base Process Sample Length	425	452	480	588	685	693	694	712	719	720
No. of Trading Days	1	1	2	1	1	1	1	1	15	483

The final data cleaning task is related to the discreteness of the data. The ASX200 index is reported correct to one decimal place, thus if the price of the ASX200 index were say 5000, then the smallest possible absolute value of a non-zero return is

$$\log\left(\frac{5000.1}{5000}\right) = 2 \times 10^{-5} \tag{6.2.1}$$

and all returns are an approximately a multiple of (6.2.1) .

Over the two year period 2009-2010, 5.9% of the 30 second interval returns were zero and 11.5% were due to a move in the ASX200 index price of ± 0.1. Returns due to a price move of ± 0.1 are "fuzzied" by changes to the ASX200 index price, but the ASX200 index price has no such effect on zero returns. Consequently, often the median of the base process is zero and there is a large number of of elements of the base process with exactly the same value as the median. This has little effect on the estimation of the stable distribution parameters but can have a significant effect on the estimation of their asymptotic variance, often resulting in negative variance estimators.

To avoid this problem we "fuzzify" the zero returns through the addition of a small amount of noise to all zero returns. The distribution of this noise is $U\left[-10^{-7}, 10^{-7}\right]$. The addition of this noise has no effect on the order identification of the base processes, but for some base processes had an effect on the moving average parameter estimates.

6.3 Estimation of SSV model parameters

In this section we describe the results of fitting SSV models to ASX200 index intraday log returns over the two year period 2009-2010 containing $D = 507$ trading days. The highest frequency ASX200 index intraday log returns available are reported every $\Delta t = 30$ seconds with results in a maximum of 720 intraday log returns each full trading day. Let $\{X_{d;t}\}$ denote the intraday log return on day d and time t. Only trading days are included in this index, therefore $d = 1$ refers to Friday 2nd Jan 2009, day $d = 2$ refers to Monday 5th Jan

2009 and so on up to $d = 507$ which refers to Friday 31st December 2010. The time index $t = 1$ refers to the period ending at 10:00:30, $t = 2$ refers to the period ending at 10:01:00 and so on up to $t = 720$ which refers to the period ending at 16:00:00.

For this analysis, SSV models are not fitted only on the highest frequency intraday log returns, but also on lower frequency returns. Non-overlapping temporal aggregation can convert of the 30 second intraday log returns into lower frequency returns. Let

$$S_{d;t}^{(r)} = \sum_{j=0}^{r-1} X_{d;t-j}, \quad t = r\tau, \tau = 1, \ldots . T_r \tag{6.3.1}$$

denote the non-overlapping temporal aggregation of $X_{d;t}$. Clearly

$$X_{d;t} = S_{d;t}^{(1)}. \tag{6.3.2}$$

The frequency of the intraday log returns used for fitting SSV models may be identified interchangeably by either $\Delta t, r$ or $\max(T_r)$ as listed below:

Δt (sec)	30	60	90	120	150	180	240	300
r	1	2	3	4	5	6	8	10
$\max(T_r)$	720	360	240	180	144	120	90	72

. (6.3.3)

The observable base processes for SSV model estimation are $\left\{S_{d;t}^{(r)}\right\}$ for $d = 1, \ldots, D$.

Let $\zeta_{p;d}^{(r)}$ denote the LQD of $\left\{S_{d;t}^{(r)}\right\}$ at aggregation level r and quantile level $p = (p_1, p_2)$. Recall that an SSV model makes the following assumptions about the nature of the intraday log returns.

A6.3.1 That the intraday log returns follow an SMA(q) process and that the parameters $\theta_1, \ldots, \theta_q$ do not change with d.

A6.3.2 That the distribution of the intraday log returns is symmetric.

A6.3.3 That the stability parameter α does not change with d.

A6.3.4 That $\left\{\zeta_{p;d}^{(T_r)}\right\}$ is an ARMA process indexed by day d.

The first step in the fitting of an SSV model involves the estimation of $\zeta_{p;d}^{(T_r)}$ for $d = 1, \ldots, D$ at some choice of quantile level $p = (p_1, p_2)$ from the corresponding base process $\left\{S_{d;t}^{(r)}\right\}$. The LQD estimator used in this analysis is the mean symmetric parametric LQD

estimator, $\overline{\zeta}_{p;d}^{(T_r)}$, described in Section 5.2. The steps involved in calculating $\overline{\zeta}_{p;d}^{(T_r)}$ are listed below.

S6.3.1 Calculate the estimators $\widehat{\alpha}_d$ and $\widehat{\theta}_{1;d}$ for each d from the observed intraday log returns $\left\{S_{d;t}^{(r)}\right\}$.

S6.3.2 Calculate the means $\overline{\alpha}$ and $\overline{\theta}_1$ of the $\widehat{\alpha}_d$ and $\widehat{\theta}_{1;d}$ estimators respectively.

S6.3.3 Calculate the estimator $\overline{\gamma}_d^{(1)}$ from $\overline{\alpha}$ using (5.2.3).

S6.3.4 Calculate the estimator $\overline{\zeta}_{p;d}^{(T_r)}$ for each d from $\overline{\alpha}, \overline{\gamma}_d^{(1)}$ and $\overline{\theta}_1$.

S6.3.5 Estimate the asymptotic variance $V_{\overline{\zeta}_p^{(T_r)}}$ of $\overline{\zeta}_{p;d}^{(T_r)}$. Note that $V_{\overline{\zeta}_p^{(T_r)}}$ has the same value for each d.

S6.3.6 Use $V_{\overline{\zeta}_p^{(T_r)}}$ to calculate an approximation for the measurement error variance

$$\sigma_w^2 \approx (T_r)^{-1} V_{\overline{\zeta}_p^{(T_r)}}. \tag{6.3.4}$$

Once individual estimates for $\zeta_{p;d}^{(T_r)}$ and an approximation for σ_w^2 have been calculated, it remains to estimate the parameters of the SSV model. Order identification of an ARMA(k,m) model is done by finding the values of k, m which minimise the AICC statistic, (see Section 9.2, Brockwell and Davis (1991)). For each r, and all $k, m \leq 2$, the AICC statistic was calculated on the observable process $\left\{\overline{\zeta}_{p;d}^{(T_r)}\right\}$ and either an ARMA$(1,1)$ or an ARMA$(2,2)$ model was identified as optimal. For this analysis, we assume that $\left\{\overline{\zeta}_{p;d}^{(T_r)}\right\}$ is an ARMA$(1,1)$ process and fit to the unobservable process $\left\{\zeta_{p;d}^{(T_r)}\right\}$ both an ARMA$(1,1)$ model using the CML method as well as an AR(1) model using the RML method.

In this chapter, we use three different sets of quantile levels for the quantile-based estimation of stable distribution parameters. We hereafter refer to those quantile levels as follows:

$$p_{M_A} = (0.040, 0.200, 0.500, 0.800, 0.960), \tag{6.3.5}$$

$$p_{M_B} = (0.022, 0.237, 0.500, 0.763, 0.978) \tag{6.3.6}$$

and the standard quantile levels as

$$p_{M_C} = (0.050, 0.250, 0.500, 0.750, 0.950) \tag{6.3.7}$$

The precision of $\overline{\zeta}_{p;d}^{(T_r)}$ estimates depends largely on the precision of the $\overline{\gamma}_d^{(1)}$ estimates and less so on the precision of $\overline{\alpha}$ and $\overline{\theta}_1$ which being the mean of D independent estimates are already very precisely estimated. The quantile levels p_{M_A} were chosen because they provide better estimates of $\overline{\gamma}_d^{(1)}$ and $\overline{\alpha}$ than the standard quantile levels, p_{M_C}, from symmetric IID samples where $\alpha > 1.2$, (See Tables 3.5.1 and 3.5.7). We do not claim that the quantile levels p_{M_A} are optimal, but merely superior to p_{M_B} and p_{M_C} for the purposes of this analysis. The quantile levels p_{M_B} are the optimal quantile levels for estimation of α from symmetric IID samples where $\alpha = 1.6$, (See Table 3.5.1).

Estimates for $\overline{\alpha}$ and $\overline{\theta}_1$ and the approximations for σ_w^2 are listed in Table 6.3.1. Whilst the estimators using p_{M_B} provide more accurate estimates of α than those using p_{M_A} and p_{M_C}, it is the estimators using p_{M_A} which provide the more accurate estimates of $\zeta_{p;d}^{(T_r)}$ and therefore the lower measurement error variance σ_w^2. Estimates of θ_1 are unaffected by the choice of quantile levels for stable distribution parameter estimation.

If the ASX200 index intraday log returns were truly from an SSV model then we would expect $\overline{\alpha}$ to be the same for all intervals Δt and $\overline{\theta}_1$ to decrease in absolute value as Δt increases. The box plots of $\widehat{\alpha}_d$ and $\widehat{\theta}_{1;d}$ in Figures 6.3.1 and 6.3.2 respectively illustrate these relationships. The relationship between $\overline{\alpha}$ and $\overline{\theta}_1$ and Δt is as expected where Δt is greater than 120 seconds, but contrary to expectations, both $\overline{\alpha}$ and $\overline{\theta}_1$ appear to increase with Δt where Δt is less than 120 seconds. This may be partially explained by the discontinuous nature of the intraday log return distributions at low interval lengths (See Section 6.2).

If the intraday log returns have a stable distribution, it is expected that estimates of α obtained using each of the quantile level choices p_{M_A}, p_{M_B} and p_{M_C} would be similar. Again, the difference between the estimates of α is generally greater at low values of Δt than high values of Δt. Typically, an analysis of variance (ANOVA) test might be used to test for any differences between the $\overline{\alpha}$ estimators. However, in this case the results would be compromised by the dependence between the $\overline{\alpha}$ estimators.

Estimates for ARMA$(1,1)$ SSV model parameters $\widehat{\mu}, \widehat{\phi}_1, \widehat{\psi}_1$ and $\widehat{\sigma}_u^2$ are listed in Table 6.3.2 and estimates for AR(1) SSV model parameters $\widehat{\mu}, \widehat{\phi}_1$ and $\widehat{\sigma}_u^2$ are listed in Table 6.3.3. All SSV model parameter estimates were calculated using LQD estimators calculated from stable distribution estimators calculated using quantile levels p_{M_A}.

If the ASX200 index intraday log returns were truly from an SSV model then we would

Δt (sec)	$\overline{\alpha}$			$\overline{\theta}_1$	σ_w^2		
	(i)	(ii)	(iii)		(i)	(ii)	(iii)
30	1.588 (0.003)	1.620 (0.003)	1.578 (0.004)	-0.022 (0.001)	0.0016	0.0018	0.0018
60	1.572 (0.005)	1.617 (0.004)	1.543 (0.005)	0.031 (0.002)	0.0033	0.0037	0.0037
90	1.601 (0.006)	1.638 (0.005)	1.566 (0.006)	0.047 (0.003)	0.0049	0.0055	0.0056
120	1.637 (0.007)	1.652 (0.006)	1.604 (0.007)	0.050 (0.003)	0.0063	0.0073	0.0073
150	1.653 (0.008)	1.669 (0.006)	1.624 (0.008)	0.055 (0.004)	0.0078	0.0091	0.0090
180	1.691 (0.009)	1.700 (0.007)	1.658 (0.010)	0.043 (0.004)	0.0091	0.0107	0.0106
240	1.687 (0.010)	1.660 (0.008)	1.673 (0.012)	0.047 (0.005)	0.0122	0.0147	0.0141
300	1.683 (0.011)	1.665 (0.009)	1.688 (0.013)	0.020 (0.005)	0.0153	0.0182	0.0175

Table 6.3.1: Estimates for $\overline{\alpha}$, $\overline{\theta}_1$ and the measurement error variance σ_w^2 from the ASX200 index intraday log return data between Jan 2009 and Dec 2010 at selected return interval lengths, Δt. Estimates for $\overline{\alpha}$ and σ_w^2 were calculated using quantile-based stable distribution parameter estimators at each of the quantile levels (i) p_{M_A}, (ii) p_{M_B} and (iii) p_{M_C}. Estimates for σ_w^2 were calculated for $\overline{\zeta}_p^{(T_r)}$ at $p = (0.25, 0.75)$. Standard errors of estimates are in ( ).

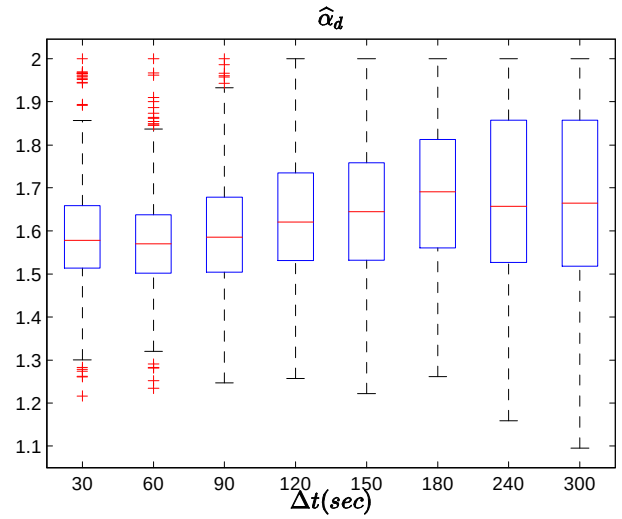


Figure 6.3.1: Box plots of estimates for $\widehat{\alpha}_d$ calculated using quantile levels p_{M_A}.

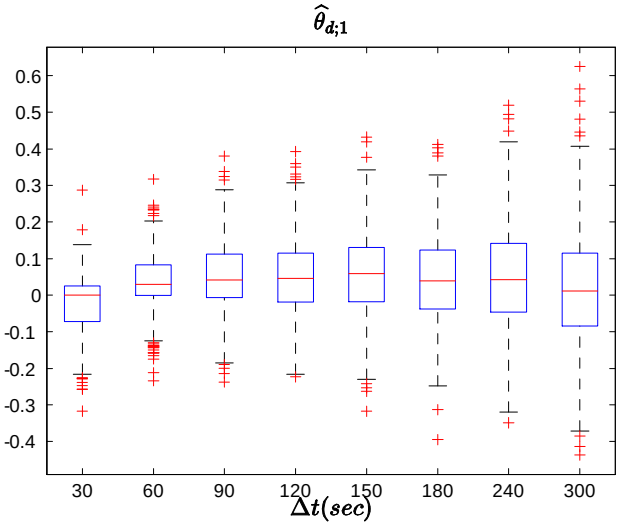


Figure 6.3.2: Box plots of estimates for $\widehat{\theta}_{d;1}$.

Δt (sec)	$\widehat{\mu}$	$\widehat{\phi}_1$	$\widehat{\psi}_1$	$\widehat{\sigma}_u^2$
30	-4.359	0.975	-0.608	0.016
	(0.101)	(0.011)	(0.040)	(0.001)
60	-4.328	0.969	-0.584	0.026
	(0.112)	(0.013)	(0.042)	(0.002)
90	-4.380	0.969	-0.563	0.031
	(0.122)	(0.013)	(0.045)	(0.002)
120	-4.449	0.968	-0.549	0.034
	(0.123)	(0.013)	(0.045)	(0.003)
150	-4.463	0.965	-0.501	0.032
	(0.118)	(0.014)	(0.048)	(0.003)
180	-4.547	0.969	-0.563	0.035
	(0.123)	(0.012)	(0.046)	(0.003)
240	-4.538	0.964	-0.528	0.035
	(0.113)	(0.013)	(0.047)	(0.003)
300	-4.558	0.960	-0.507	0.035
	(0.105)	(0.014)	(0.055)	(0.004)

Table 6.3.2: Parameters of the ARMA$(1,1)$,SSV model on ASX200 index intraday log return data between Jan 2009 and Dec 2010. The ARMA$(1,1)$SSV model parameter estimates were calculated using LQD estimators calculated from stable distribution parameter estimators calculated using quantile levels p_{M_A}. Standard errors of parameters estimates are in ( ).

expect that the each of the ARMA$(1,1)$ and AR(1) SSV model parameters would be the same for all intervals Δt. A graphical display of these parameters and their standard errors is provided in Figure 6.3.3 for the ARMA$(1,1)$ SSV model and Figure 6.3.4 for the AR(1) SSV model. As was the case with the estimators $\overline{\alpha}$ and $\overline{\theta}_1$, any deviation from the equality of the SSV model parameter estimates across Δt values appears more significant at low values of Δt.

A similar analysis on the same data was done in Barker (2014). However, in Barker (2014) the distribution of the intraday log return processes was not assumed to be symmetrical and the parameters α and $\beta^{(1)}$ were allowed to take different values on each day.

Δt (sec)	$\widehat{\mu}$	$\widehat{\phi}_1$	$\widehat{\sigma}_u^2$
30	-4.297	0.831	0.018
	(0.030)	(0.022)	(0.001)
60	-4.266	0.822	0.029
	(0.051)	(0.032)	(0.003)
90	-4.315	0.837	0.033
	(0.072)	(0.037)	(0.004)
120	-4.379	0.848	0.036
	(0.092)	(0.042)	(0.005)
150	-4.399	0.871	0.032
	(0.114)	(0.044)	(0.005)
180	-4.473	0.854	0.037
	(0.117)	(0.051)	(0.006)
240	-4.469	0.871	0.033
	(0.145)	(0.057)	(0.007)
300	-4.490	0.878	0.028
	(0.167)	(0.063)	(0.008)

Table 6.3.3: Parameters of the AR(1), SSV model on ASX200 index intraday log return data between Jan 2009 and Dec 2010. The AR(1)SSV model parameter estimates were calculated using LQD estimators calculated from stable distribution parameter estimators calculated using quantile levels p_{M_A}. Standard errors of parameters estimates are in ( ).

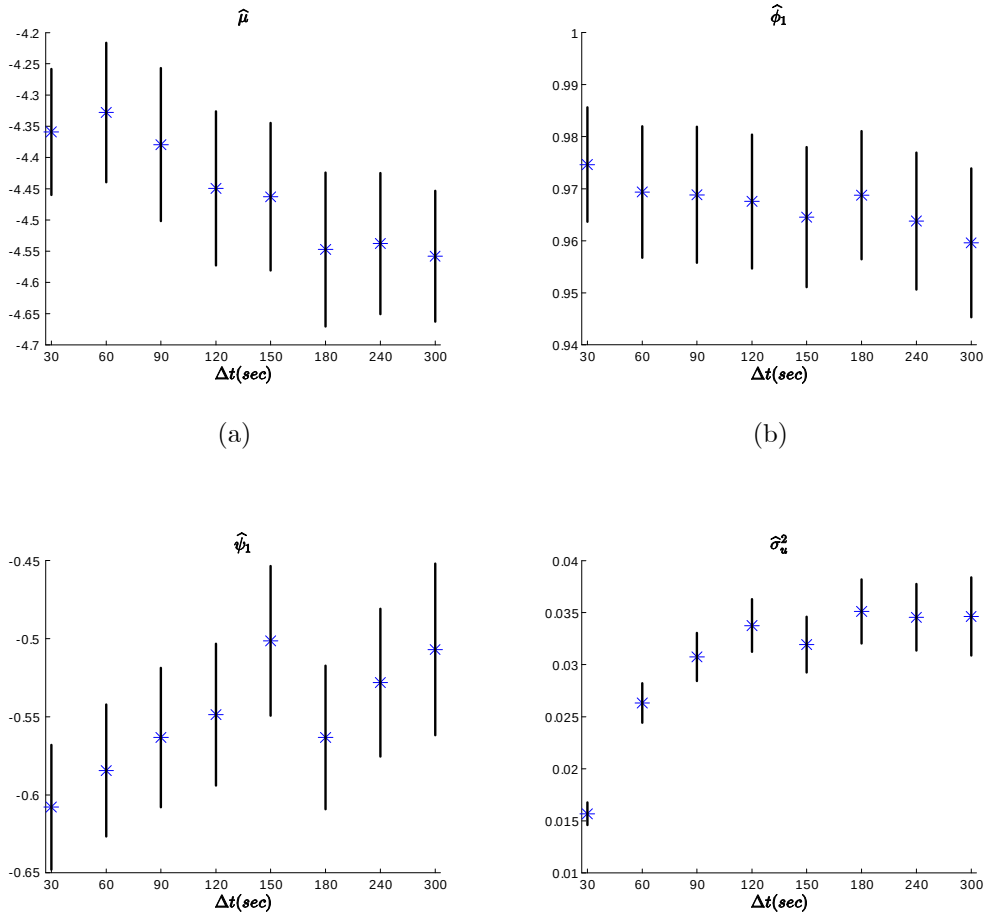
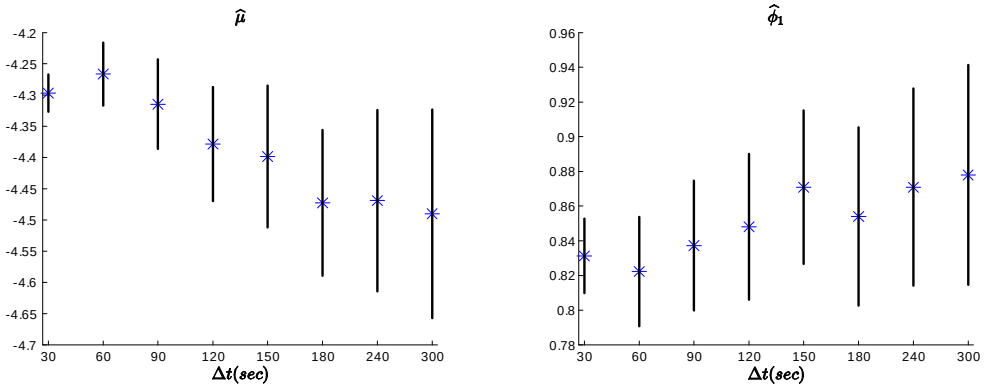
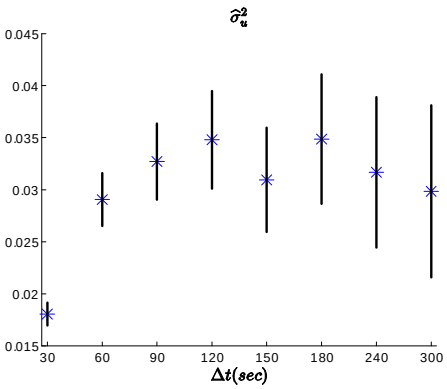


Figure 6.3.3: Estimates of ARMA$(1,1)$ SSV model parameters and their standard errors: (a) $\widehat{\mu}$, (b) $\widehat{\phi}_1$, (c) $\widehat{\psi}_1$ and (d) $\widehat{\sigma}_u^2$,



(a) (b)



(c)

Figure 6.3.4: Estimates of AR(1) SSV model parameters and their standard errors: (a) $\widehat{\mu}$, (b) $\widehat{\phi}_1$ and (c) $\widehat{\sigma}_u^2$,

The standard quantile levels p_{M_C} were used for stable distribution parameter estimation. In addition, the moving average order of the intraday log return processes was not assumed to be equal to one, but was instead determined using the Adler method, Adler et al. (1998), (see Section 3.8). No allowance for measurement error was made in Barker (2014) and similar values for $\widehat{\phi}_1$ were obtained.

6.4 Diagnostics

In this section we test whether the ASX200 index data supports some of the SSV model assumptions. The null hypotheses to be tested are

H_1 That the α parameter of the distribution of the intraday log returns is constant for all d,

H_2 That the θ_1 parameter of the SMA(1) intraday log return process is constant for all d,

H_3 That the moving average order of each of the intraday log return processes is less than or equal to one,

H_4 That the distribution of the intraday log returns are stationary for all t.

First, we test the null hypothesis H_1. Let $S_{\widehat{\alpha}}$ denote the sample variance of the estimators $\widehat{\alpha}_d$, so that

$$S_{\widehat{\alpha}} = (D-1)^{-1} \sum_{d=1}^{D} \left(\widehat{\alpha}_d - \overline{\alpha}\right)^2, \tag{6.4.1}$$

and let $V_{\widehat{\alpha}}$ denote the asymptotic variance of $\widehat{\alpha}_d$. If the null hypothesis H_1 is true, then the distribution of $S_{\widehat{\alpha}}$ can be approximated by

$$\frac{(D-1) * T_r * S_{\widehat{\alpha}}}{V_{\widehat{\alpha}}} \sim \chi^2_{D-1} \tag{6.4.2}$$

where χ^2_{D-1} is the chi-squared distribution with $D-1$ degrees of freedom. The test statistic used for testing the null hypothesis H_1 is

$$Z_1 = T_r * S_{\widehat{\alpha}} \tag{6.4.3}$$

and the rejection region at significance level a is

$$Z_1 > R_1(a) \tag{6.4.4}$$

	p_{M_A}			p_{M_B}		
T_r	Z_1	$R_1(a)$	Reject H_1	Z_1	$R_1(a)$	Reject H_1
720	11.276	4.211	Yes	7.562	3.339	Yes
360	5.215	4.117	Yes	3.345	3.339	Yes
240	4.742	4.309	Yes	3.214	3.332	No
180	4.496	4.604	No	3.592	3.329	Yes
144	4.543	4.776	No	3.452	3.344	Yes
120	3.935	5.212	No	3.056	3.371	No

Table 6.4.1: Test results for the null hypothesis, H_1 at significance level $a = 0.95$ based on the asymptotic variance using quantile levels p_{M_A} and p_{M_B}. The null hypothesis is rejected if. $Z_1 > R_1(a)$.

where

$$R_1(a) = \frac{V_{\widehat{\alpha}} * \chi^2_{D-1}(a)}{D-1} \tag{6.4.5}$$

and $\chi^2_{D-1}(a)$ is the a^{th} quantile of the χ^2_{D-1} distribution.

Initial results for the H_1 null hypothesis tests are listed in Table 6.4.1. Tests were conducted from $\widehat{\alpha}_d$ estimators calculated using the quantile levels p_{M_A} and p_{M_B}. Test results derived from the more accurate estimators of α_d using quantile levels p_{M_B} are preferred. Whilst H_1 is clearly rejected at $T_r = 720$ and clearly not rejected at $T_r = 120$, the results are mixed for the intermediate values of T_r.

The rejection region defined in (6.4.4) is based on the assumption that the small sample distribution of $S_{\widehat{\alpha}}$ can be approximated by its asymptotic distribution in (6.4.2) . In practice, this approximation is not perfect. Let $V_{\widehat{\alpha};T_r}$ denote the standardised small sample variance of $\widehat{\alpha}_d$ from samples of size T_r, the following opposing characteristics help to explain the differences between $V_{\widehat{\alpha}}$ and $V_{\widehat{\alpha};T_r}$:

1. The small sample standardised variance of quantile estimators from an SMA process tend to be higher than the asymptotic variance. This has the effect of making $V_{\widehat{\alpha};T_r}$ approximately 5% larger than $V_{\widehat{\alpha}}$ for $T_r = 720$ and approximately 25% larger than $V_{\widehat{\alpha}}$ for $T_r = 120$.

2. The maximum value of $\widehat{\alpha}_d$ is 2. Where the sample size is small and α is close to 2, this

	p_{M_A}			p_{M_B}		
T_r	Z_1	$R_{1;T}(a)$	Reject H_1	Z_1	$R_{1;T}(a)$	Reject H_1
720	11.276	4.548	Yes	7.562	3.539	Yes
360	5.215	4.734	Yes	3.345	3.640	No
240	4.742	4.912	No	3.214	3.898	No
180	4.496	4.973	No	3.592	3.928	No
144	4.543	5.015	No	3.452	3.946	No
120	3.935	4.483	No	3.056	4.012	No

Table 6.4.2: Test results for the null hypothesis, H_1 at significance level $a = 0.95$ based on the standardised small sample variance using quantile levels p_{M_A} and p_{M_B}. The null hypothesis is rejected if. $Z_1 > R_{1;T}(a)$.

has the effect of truncating the small sample distribution of $\widehat{\alpha}_d$ and reducing $V_{\widehat{\alpha};T_r}$ relative to $V_{\widehat{\alpha}}$. For p_{M_A} and $T_r = 120$, approximately 10% of the $\widehat{\alpha}_d$ estimates are equal to 2.

Simulations were run to estimate $V_{\widehat{\alpha};T}$ and those estimates were used to define a new rejection region $R_{1;T_r}(a)$ based on the small sample distribution of $S_{\widehat{\alpha}}$, (Table 6.4.2). These test results indicate that H_1 should be rejected at $T_r = 720$, but does not support the rejection of H_1 at other values of T_r.

Remark 6.4.1 *Tests conducted using the* ASX200 *index data do not reject the* SSV *model assumption that the α parameter of the distribution of the intraday log returns is constant for all d, except for intraday log returns at 30 second intervals. The reason that intraday log returns at 30 second intervals has a larger than expected variance of $\widehat{\alpha}_d$ estimates may be related to the discontinuous nature of the data at 30 second intervals, which is less significant at longer intervals.*

Next, we test the null hypothesis H_2. Let $S_{\widehat{\theta}_1}$ denote the sample variance of the estimators $\widehat{\theta}_{1;d}$, so that

$$S_{\widehat{\theta}} = (D-1)^{-1} \sum_{d=1}^{D} \left(\widehat{\theta}_{1;d} - \overline{\theta}_1 \right)^2, \tag{6.4.6}$$

T_r	Z_1	$R_2(a)$	Reject H_2
720	4.736	0.805	Yes
360	2.373	0.785	Yes
240	2.230	0.848	Yes
180	1.975	0.930	Yes
144	1.880	0.969	Yes
120	1.696	1.061	Yes

Table 6.4.3: Test results for the null hypothesis, H_2 at significance level $a = 0.95$. The null hypothesis is rejected if. $Z_2 > R_2(a)$.

and let $V_{\widehat{\theta}}$ denote the asymptotic variance of $\widehat{\theta}_{1;d}$. If the null hypothesis H_2 is true, then the distribution of $S_{\widehat{\theta}}$ can be approximated by

$$\frac{(D-1) * T_r * S_{\widehat{\theta}}}{V_{\widehat{\theta}}} \sim \chi^2_{D-1} \tag{6.4.7}$$

where χ^2_{D-1} is the chi-squared distribution with $D-1$ degrees of freedom. The test statistic used for testing the null hypothesis H_2 is

$$Z_2 = T_r * S_{\widehat{\theta}_1} \tag{6.4.8}$$

and the rejection region at significance level a is

$$Z_2 > R_2(a) \tag{6.4.9}$$

where

$$R_2(a) = \frac{V_{\widehat{\theta}} * \chi^2_{D-1}(a)}{(D-1)} \tag{6.4.10}$$

and $\chi^2_{D-1}(a)$ is the a^{th} quantile of the χ^2_{D-1} distribution. We do not have a formula for the standardised small sample variance $V_{\widehat{\theta}_1}$, nor the asymptotic variance of $\widehat{\theta}_1$, therefore we use simulations to estimate its value. Results for the H_2 null hypothesis tests are listed in Table 6.4.3.

Remark 6.4.2 *Tests conducted using the* ASX200 *index data reject the* SSV *model assumption that the θ_1 parameter of the distribution of the intraday log returns is constant for all d.*

T_r	(i)	(ii)	(iii)
720	59.6%	62.9%	55.4%
360	49.3%	44.2%	41.2%
240	39.3%	32.5%	28.2%
180	29.6%	22.3%	19.1%
144	25.3%	17.8%	14.2%
120	10.3%	10.3%	7.5%

Table 6.4.4: Q statistic test results at significance level $a = 0.95$. Listed in this table is the percentage of days on which (i) the null hypothesis of independence of the intraday log return process was rejected, (ii) the null hypothesis of independence of the residuals of SMA(1) model using $\overline{\theta}_1$ was rejected and (iii) the null hypothesis of independence of the residuals of SMA(1) model using $\widehat{\theta}_1$ was rejected.

To test the null hypothesis H_3, we use two different tests, both based on the sample autocorrelation function: the Q-statistic, (see Section 3.7) and the extended Adler test, (see Section 3.8).

The Q-statistic is defined by the following equation

$$Q_s = T_r \sum_{h=1}^{s} \widehat{\rho}\,(h)\,, \quad s < T_r \tag{6.4.11}$$

where $\widehat{\rho}\,(h)$ is the mean corrected sample autocorrelation function (Definition 3.6.1) at lag h of the residuals of the SMA(1) model fitted to the intraday log return data. Under the null hypothesis H_3, the residuals of the SMA(1) model fitted to the intraday log return data have the asymptotic distribution given by Theorem 3.7.1. Let $Q_{s;T_r}\,(a)$ denote the a^{th} quantile of the small sample distribution of Q_s from a sample size of T_r. The rejection region of this test is

$$Q_s > Q_{s;T_r}\,(a) \tag{6.4.12}$$

where we choose $s = 20$ and $a = 0.95$. As discussed in Section 3.7, convergence to the asymptotic distribution of Q_s is very slow and simulations were used to derive numerical approximations for $Q_{s;T_r}\,(a)$. See Table 3.7.1 for values of $Q_{s;T_r}\,(a)$ where $T_r = 720$. Results of the Q statistic tests for null hypothesis H_3 are presented in Table 6.4.4.

The autocorrelation function of an SMA(q) process is zero at all lags greater than q. If the absolute value of the sample autocorrelation function $\widehat{\rho}(h)$ is greater than some significance level, then that suggests the moving average order of the process is greater than or equal to h. In Adler et al. (1998), the asymptotic distribution of the sample autocorrelation function from an IID symmetric Cauchy process was used to determine a significance level for the test. As described in Section 3.8, it was found that better results could be achieved using significance levels determined by simulations from an IID process of the same size and with the same α and β stable distribution parameters.

Let $B_{\widehat{\rho}(1)}(n, a; \alpha, \beta)$ denote the a^{th} percentile of $\widehat{\rho}(1)$ from an IID sample of length n and distribution $S_\alpha(\beta, 1, 0)$. Let I_{Finite} denote the interval

$$I_{Finite} = \left[B_{\widehat{\rho}(1)}(T_r, a_1; \overline{\alpha}, 0), B_{\widehat{\rho}(1)}(T_r, a_2; \overline{\alpha}, 0)\right] \tag{6.4.13}$$

where $(a_1, a_2) = (1.0\%, 99.0\%)$. For this test we determine the order of each intraday log return process as follows:

1. If the sample autocorrelation function at any of the lags $h = 2, 3, \ldots, 8$ lies outside the interval I_{Finite}, then we say that the moving average order of the process is greater than one,

2. If the sample autocorrelation function at each of the lags $h = 1, 2, \ldots, 8$ lies inside the interval I_{Finite}, then we say that the moving average order of the process is equal to zero,

3. Otherwise we say that the moving average order of the process is equal to one.

The results of this test are presented in Table 6.4.5.

Remark 6.4.3 *Tests conducted using the* ASX200 *index data suggest that the moving average order of the intraday log return processes may be greater than one in many cases, though this appears to be less likely as the return interval increases, i.e. T decreases. The poor results at low return interval values may be due to the discreteness of the log returns at these intervals.*

Finally we consider null hypothesis H_4. We do not have a formal test for H_4, but instead display some graphical results. Let

$$Z_t^{(r)}(d) = \left\{S_{d;t}^{(r)}\right\},$$

T_r	$q = 0$	$q = 1$	$q > 1$
720	23.5%	9.8%	66.7%
360	33.9%	9.3%	56.8%
240	37.9%	21.5%	40.6%
180	46.8%	20.1%	33.1%
144	51.7%	21.5%	26.8%
120	64.1%	11.8%	24.1%

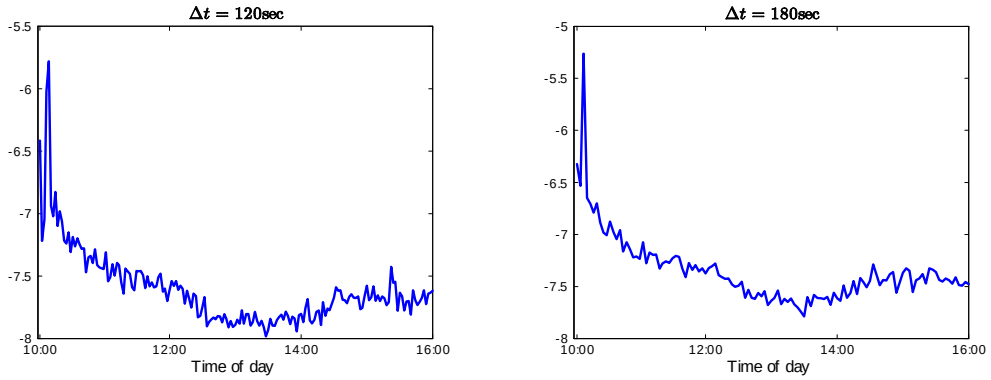
Table 6.4.5: Test results for the moving average order, q, of the intraday log return processes based on the extended Adler test.

where t is fixed, denote the intraday log returns at some particular time of day. In the SSV model,

$$Z_t^{(r)}(d) \sim S_\alpha^0\left(0, \gamma_d^{(1)}, 0\right)$$

that is $Z_t^{(r)}(d)$ is an independent mixture of stable distributions which does not depend on t. In Figure 6.4.1, plots are provided of the log quantile differences of $\left\{Z_t^{(r)}(d)\right\}$ at quantile level $p = (0.25, 0.75)$ and return intervals of 120 seconds and 180 seconds.

Remark 6.4.4 *A study of mixtures of stable distributions is beyond the scope of this thesis, so we do not attempt to conduct a formal statistical test on the null hypothesis H_5. However, the plots in Figure 6.4.1, do suggest that trading on the* ASX200 *index is more volatile at the beginning of the day than during the remainder of the day and therefore that the intraday log return processes are not stationary. Intuitively, this can be explained by the excess of information from overseas events which needs to be absorbed by the market at the beginning of the day.*



(a) (b)

Figure 6.4.1: Plots of the empirical LQD estimator of the distribution of the intraday log returns as a function of the time of day. The LQD estimators are calculated at quantile level $p = (0.25, 0.75)$ and return interval lengths (a) $\Delta t = 120(sec)$ and (b) $\Delta t = 180(sec)$.

Chapter 7

Conclusion

The main topic of this thesis is the estimation and behaviour of log quantile differences (LQD) of alpha-stable moving average (SMA) processes under temporal aggregation. Formulae for the LQDs are derived in Chapter 2 and these formulae are used to categorise the properties of the LQDs of the temporal aggregation of some SMA(2) processes by the value of their moving average parameters.

In Chapter 3, an extension is made to the quantile-based method from McCulloch (1986) for estimating alpha-stable distribution parameters to allow estimation from an SMA process. Optimisation of this method is done through the use of quantile levels other than the standard quantile levels originally proposed in McCulloch (1986). More accurate evaluations are made of the asymptotics of maximum likelihood stable distribution parameter estimators which allow the calculation of the relative asymptotic efficiencies of the optimal quantile-based estimators. Improved performance is demonstrated in statistical tests for the order identification and residual independence of SMA processes through the introduction of significance levels based on small sample rather than asymptotic distributions.

In Chapter 4, the formulae from Chapter 2 and the estimators of Chapter 3 are combined to define asymptotically normal estimators for the LQD of the temporal aggregation of an SMA process. Whilst the focus of Chapters 2, 3 and 4 is on SMA processes, an extension of many of these results may be possible to cover alpha-stable autoregressive and alpha-stable autoregressive moving average processes.

In Chapter 5, an application is developed for the LQD estimators proposed in Chapter 4 as a means for measuring the realised volatility of financial market asset returns. The

realised volatility in terms of LQDs is calculated under the assumption that the intraday log returns were SMA processes. That intraday log returns follow an SMA process is inconsistent with the stochastic volatility (SV) model which assumes that the daily log returns have a finite conditional variance. Consequently, an extension to the SV model is proposed, called the stable stochastic volatility (SSV) model, where the conditional distribution of the daily log returns is stable. Estimation of the SSV model parameters is done utilising the realised volaility in terms of LQDs together with an allowance for measurement error.

In Chapter 6, an empirical study on these methods is carried out on ASX200 index data from 2009 and 2010. For each day an SMA(1) process was fitted to the intraday log returns at selected return interval lengths. Estimates for the SSV model parameters were calculated from the parameters of the fitted SMA(1) processes and some diagnostic tests were conducted, including the residual independence and order identification tests proposed in Chapter 3. Estimation of the SSV model may be improved by allowing intraday log returns to be modelled by non-stationary SMA processes which account for higher volatilities at the start of the trading day.

Appendix A

Proofs

A.1 Proof of Theorem 2.3.1

The proof of Theorem 2.3.1 involves the collation of a number of lemmas, each of which proves the properties of $g_\alpha(\theta_1, \theta_2)$ for a particular part of the invertibility region. We commence with a version of Jensen's inequality, which we apply throughout this proof. (See for example Rudin (1987).)

Lemma A.1.1 ***(Jensen's Inequality).*** *If the function f is strictly convex on (a, b), then for $a < x, y < b$*

$$f(x) + f(y) < f(x + y) \tag{A.1.1}$$

If $1 < \alpha \leq 2$, then the function $f(x) = x^\alpha$ is strictly convex on the interval $(0, M)$ for $M > 0$. If $0 < \alpha < 1$, then the function $f(x) = -(x^\alpha)$ is strictly convex on the interval $(0, M)$ for $M > 0$. Applying Jensen's inequality to these functions gives

$$x^\alpha + y^\alpha < (x + y)^\alpha \quad \text{for } x, y > 0 \text{ and } 1 < \alpha \leq 2 \tag{A.1.2}$$

$$x^\alpha + y^\alpha > (x + y)^\alpha \quad \text{for } x, y > 0 \text{ and } 0 < \alpha < 1. \tag{A.1.3}$$

Since all elements of the invertibility region satisfy

$$\theta_1 + \theta_2 > -1, \tag{A.1.4}$$

we can rewrite the expression for $g_\alpha(\theta_1, \theta_2)$ in the invertibility region as

$$g_\alpha(\theta_1, \theta_2) = 1 + |1 + \theta_1|^\alpha - 2(1 + \theta_1 + \theta_2)^\alpha + |\theta_1 + \theta_2|^\alpha + |\theta_2|^\alpha. \tag{A.1.5}$$

For convenience we define two new functions $g_\alpha^{(1)}(\theta_1,\theta_2)$ and $g_\alpha^{(2)}(\theta_1,\theta_2)$ as follows

$$g_\alpha^{(a)}(\theta_1,\theta_2) = 1+|\theta_1+\theta_2|^\alpha-(1+\theta_1+\theta_2)^\alpha \tag{A.1.6}$$

$$g_\alpha^{(b)}(\theta_1,\theta_2) = |1+\theta_1|^\alpha+|\theta_2|^\alpha-(1+\theta_1+\theta_2)^\alpha \tag{A.1.7}$$

so that

$$g_\alpha(\theta_1,\theta_2) = g_\alpha^{(a)}(\theta_1,\theta_2)+g_\alpha^{(b)}(\theta_1,\theta_2). \tag{A.1.8}$$

The following lemmas show partial results for $g_\alpha^{(a)}(\theta_1,\theta_2)$ and $g_\alpha^{(b)}(\theta_1,\theta_2)$ in the various sub-regions.

Lemma A.1.2 *In sub-regions 1,2 and 4,* $g_\alpha^{(a)}(\theta_1,\theta_2)>0$.

Proof. In sub-regions 1,2 and 4

$$-1<\theta_1+\theta_2<0, \tag{A.1.9}$$

therefore

$$(1+\theta_1+\theta_2)^\alpha<1 \tag{A.1.10}$$

and

$$\begin{aligned} g_\alpha^{(a)}(\theta_1,\theta_2) &> |\theta_1+\theta_2|^\alpha \\ &> 0 \end{aligned} \tag{A.1.11}$$

■

Lemma A.1.3 *In sub-regions 4 and 5,* $g_\alpha^{(b)}(\theta_1,\theta_2)>0$.

Proof. In sub-regions 4 and 5

$$-1<\theta_2<0, \tag{A.1.12}$$

therefore

$$(1+\theta_1+\theta_2)^\alpha<|1+\theta_1|^\alpha \tag{A.1.13}$$

and

$$\begin{aligned} g_\alpha^{(b)}(\theta_1,\theta_2) &> |\theta_2|^\alpha \\ &> 0 \end{aligned} \tag{A.1.14}$$

■

Lemma A.1.4 *In sub-region 1,* $g_\alpha^{(b)}(\theta_1, \theta_2) > 0$.

Proof. In sub-region 1

$$1 + \theta_1 < 0, \tag{A.1.15}$$

$$1 + \theta_1 + \theta_2 > 0 \tag{A.1.16}$$

therefore

$$(1 + \theta_1 + \theta_2)^\alpha < |\theta_2|^\alpha \tag{A.1.17}$$

and

$$\begin{aligned} g_\alpha^{(b)}(\theta_1, \theta_2) &> |1 + \theta_1|^\alpha \\ &> 0 \end{aligned} \tag{A.1.18}$$

■

Lemma A.1.5 *In sub-regions 3 and 5,*

$$g_\alpha^{(a)}(\theta_1, \theta_2) \text{ is } \begin{cases} > 0 & \text{for } 0 < \alpha < 1 \\ = 0 & \text{for } \alpha = 1 \\ < 0 & \text{for } 1 < \alpha \leq 2 \end{cases}. \tag{A.1.19}$$

Proof. In sub-regions 3 and 5

$$\theta_1 + \theta_2 > 0. \tag{A.1.20}$$

If $0 < \alpha < 1$, then by Jensen's inequality

$$(1 + \theta_1 + \theta_2)^\alpha < 1 + |\theta_1 + \theta_2|^\alpha \tag{A.1.21}$$

and therefore

$$g_\alpha^{(a)}(\theta_1, \theta_2) > 0. \tag{A.1.22}$$

If $\alpha = 1$, then

$$\begin{aligned} g_\alpha^{(a)}(\theta_1, \theta_2) &= 1 + \theta_1 + \theta_2 - (1 + \theta_1 + \theta_2) \\ &= 0 \end{aligned} \tag{A.1.23}$$

If $1 < \alpha \leq 2$, then by Jensen's inequality

$$(1 + \theta_1 + \theta_2)^\alpha > 1 + |\theta_1 + \theta_2|^\alpha \tag{A.1.24}$$

and so

$$g_{\alpha}^{(a)}(\theta_1, \theta_2) < 0. \tag{A.1.25}$$

■

Lemma A.1.6 *In sub-regions 2 and 3,*

$$g_{\alpha}^{(b)}(\theta_1, \theta_2) \text{ is } \begin{cases} > 0 & \text{for } 0 < \alpha < 1 \\ = 0 & \text{for } \alpha = 1 \\ < 0 & \text{for } 1 < \alpha \leq 2 \end{cases}. \tag{A.1.26}$$

Proof. In sub-regions 2 and 3

$$\theta_2 > 0, \tag{A.1.27}$$

$$\theta_1 > -1. \tag{A.1.28}$$

If $0 < \alpha < 1$, then by Jensen's inequality

$$(1 + \theta_1 + \theta_2)^{\alpha} < |1 + \theta_1|^{\alpha} + |\theta_2|^{\alpha} \tag{A.1.29}$$

and therefore

$$g_{\alpha}^{(b)}(\theta_1, \theta_2) > 0. \tag{A.1.30}$$

If $\alpha = 1$, then

$$\begin{aligned} g_{\alpha}^{(b)}(\theta_1, \theta_2) &= 1 + \theta_1 + \theta_2 - (1 + \theta_1 + \theta_2) \\ &= 0 \end{aligned} \tag{A.1.31}$$

If $1 < \alpha \leq 2$, then by Jensen's inequality

$$(1 + \theta_1 + \theta_2)^{\alpha} > |1 + \theta_1|^{\alpha} + |\theta_2|^{\alpha} \tag{A.1.32}$$

and therefore

$$g_{\alpha}^{(b)}(\theta_1, \theta_2) < 0. \tag{A.1.33}$$

■

The following lemmas show results for $g_{\alpha}^{(a)}(\theta_1, \theta_2)$ and $g_{\alpha}^{(b)}(\theta_1, \theta_2)$ in each of the borders between the sub-regions.

Lemma A.1.7 *On the border of sub-regions 1 and 2,* $g_{\alpha}(\theta_1, \theta_2) > 0$.

Proof. On the border of sub-regions 1 and 2

$$\theta_1 = -1, \tag{A.1.34}$$

$$\theta_2 > 0. \tag{A.1.35}$$

Therefore

$$\begin{aligned} g_\alpha^{(a)}(\theta_1, \theta_2) &= 1 + |-1 + \theta_2|^\alpha - \theta_2^\alpha \\ &> |-1 + \theta_2|^\alpha \\ &> 0 \end{aligned} \tag{A.1.36}$$

and

$$\begin{aligned} g_\alpha^{(b)}(\theta_1, \theta_2) &= |\theta_2|^\alpha - (\theta_2)^\alpha \\ &= 0. \end{aligned} \tag{A.1.37}$$

Combining (*A*.1.36) and (*A*.1.37) proves the lemma. ■

Lemma A.1.8 *On the border of sub-regions 2 and 4,* $g_\alpha(\theta_1, \theta_2) > 0$.

Proof. On the border of sub-regions 2 and 4

$$-1 < \theta_1 < 0, \tag{A.1.38}$$

$$\theta_2 = 0. \tag{A.1.39}$$

Therefore

$$\begin{aligned} g_\alpha^{(a)}(\theta_1, \theta_2) &= 1 + |\theta_1|^\alpha - (1 + \theta_1)^\alpha \\ &> |\theta_1|^\alpha \\ &> 0 \end{aligned} \tag{A.1.40}$$

and

$$\begin{aligned} g_\alpha^{(b)}(\theta_1, \theta_2) &= |1 + \theta_1|^\alpha - (1 + \theta_1)^\alpha \\ &= 0. \end{aligned} \tag{A.1.41}$$

Combining (*A*.1.40) and (*A*.1.41) proves the lemma. ■

Lemma A.1.9 *On the border of sub-regions 4 and 5,* $g_\alpha(\theta_1, \theta_2) > 0$.

Proof. On the border of sub-regions 4 and 5

$$\theta_1 > 0, \tag{A.1.42}$$

$$\theta_1 + \theta_2 = 0. \tag{A.1.43}$$

Therefore

$$\begin{aligned} g_\alpha^{(a)}(\theta_1, \theta_2) &= 1 - (1)^\alpha \\ &= 0 \end{aligned} \tag{A.1.44}$$

and

$$\begin{aligned} g_\alpha^{(b)}(\theta_1, \theta_2) &= |1 + \theta_1|^\alpha + |\theta_2|^\alpha - (1)^\alpha \\ &> |\theta_2|^\alpha \\ &> 0 \end{aligned} \tag{A.1.45}$$

Combining $(A.1.44)$ and $(A.1.45)$ proves the lemma. ■

Lemma A.1.10 *On the border of sub-regions 2 and 3,* $g_\alpha^{(a)}(\theta_1, \theta_2) = 0$ *and*

$$g_\alpha^{(b)}(\theta_1, \theta_2) \text{ is } \begin{cases} > 0 & \text{for } 0 < \alpha < 1 \\ = 0 & \text{for } \alpha = 1 \\ < 0 & \text{for } 1 < \alpha \leq 2 \end{cases} \tag{A.1.46}$$

Proof. On the border of sub-regions 2 and 3

$$-1 < \theta_1 < 0, \tag{A.1.47}$$

$$\theta_1 + \theta_2 = 0. \tag{A.1.48}$$

Therefore

$$\begin{aligned} g_\alpha^{(a)}(\theta_1, \theta_2) &= 1 - (1)^\alpha \\ &= 0 \end{aligned} \tag{A.1.49}$$

and

$$\begin{aligned} g_\alpha^{(b)}(\theta_1, \theta_2) &= |1 + \theta_1|^\alpha + |\theta_2|^\alpha - (1)^\alpha \\ &= |1 - |\theta_1||^\alpha + |\theta_1|^\alpha - 1 \end{aligned} \tag{A.1.50}$$

If $0 < \alpha < 1$, then by Jensen's inequality

$$g_{\alpha}^{(b)}(\theta_1, \theta_2) < 0. \tag{A.1.51}$$

If $\alpha = 1$,then

$$g_{\alpha}^{(b)}(\theta_1, \theta_2) = 0 \tag{A.1.52}$$

If $1 < \alpha \leq 2$, then by Jensen's inequality

$$g_{\alpha}^{(b)}(\theta_1, \theta_2) > 0. \tag{A.1.53}$$

Combining $(A.1.49), (A.1.51), (A.1.52)$ and $(A.1.53)$ proves the lemma. ■

Lemma A.1.11 *On the border of sub-regions 3 and 5,*

$$g_{\alpha}^{(a)}(\theta_1, \theta_2) \ \textit{is} \begin{cases} > 0 & \textit{for } 0 < \alpha < 1 \\ = 0 & \textit{for } \alpha = 1 \\ < 0 & \textit{for } 1 < \alpha \leq 2 \end{cases} \tag{A.1.54}$$

and $g_{\alpha}^{(b)}(\theta_1, \theta_2) = 0$.

Proof. On the border of sub-regions 3 and 5

$$\theta_1 > 0, \tag{A.1.55}$$

$$\theta_2 = 0. \tag{A.1.56}$$

Therefore

$$g_{\alpha}^{(a)}(\theta_1, \theta_2) = 1 + \theta_1^{\alpha} - (1 + \theta_1)^{\alpha} \tag{A.1.57}$$

and

$$\begin{aligned} g_{\alpha}^{(b)}(\theta_1, \theta_2) &= |1 + \theta_1|^{\alpha} - (1 + \theta_1)^{\alpha} \\ &= 0 \end{aligned} \tag{A.1.58}$$

If $0 < \alpha < 1$, then by Jensen's inequality

$$g_{\alpha}^{(a)}(\theta_1, \theta_2) < 0. \tag{A.1.59}$$

If $\alpha = 1$,then

$$g_{\alpha}^{(a)}(\theta_1, \theta_2) = 0 \tag{A.1.60}$$

If $1 < \alpha \leq 2$, then by Jensen's inequality

$$g_{\alpha}^{(a)}(\theta_1, \theta_2) > 0. \tag{A.1.61}$$

Combining $(A.1.58), (A.1.59), (A.1.60)$ and $(A.1.61)$ proves the lemma. ■

The following lemma shows results for $g_{\alpha}^{(a)}(\theta_1, \theta_2)$ and $g_{\alpha}^{(b)}(\theta_1, \theta_2)$ at the origin.

Lemma A.1.12 *At the origin,* $g_{\alpha}(\theta_1, \theta_2) = 0$.

Proof. At the origin

$$\begin{aligned} g_{\alpha}(\theta_1, \theta_2) &= 1 + |1 + \theta_1|^{\alpha} - 2(1 + \theta_1 + \theta_2)^{\alpha} + |\theta_1 + \theta_2|^{\alpha} + |\theta_2|^{\alpha} \\ &= 1 + |1|^{\alpha} - 2(1)^{\alpha} + 0 + 0 \\ &= 0 \end{aligned} \tag{A.1.62}$$

■

The Lemmas A.1.2 though A.1.12 are combined to prove Theorem 2.3.1.

A.2 Proof of Theorem 4.2.2

The following lemma provides formulae for each of the partial derivatives $\frac{\partial \widehat{\omega}_i^{(r)}}{\partial \widehat{\varkappa}_j}$. Derivation of these formulae is achieved through the application of the standard chain and product rules of differentiation and is presented without proof.

Lemma A.2.1 *If $C_4 \neq 0$ and $\widehat{\alpha} \geq 1$ then*

$$\frac{\partial \widehat{\alpha}}{\partial \widehat{\alpha}} = 1, \tag{A.2.1}$$

$$\frac{\partial \widehat{\beta}^{(r)}}{\partial \widehat{\alpha}} = \widehat{\beta}^{(1)} \frac{\partial}{\partial \widehat{\alpha}} \left(\frac{C_2 C_3^{(r)}}{C_1^{(r)} C_4} \right), \tag{A.2.2}$$

$$\frac{\partial \widehat{\gamma}^{(r)}}{\partial \widehat{\alpha}} = \widehat{\gamma}^{(1)} \frac{\partial}{\partial \widehat{\alpha}} \left(\frac{C_1^{(r)}}{C_2} \right)^{1/\widehat{\alpha}}, \tag{A.2.3}$$

$$\frac{\partial \widehat{\alpha}}{\partial \widehat{\beta}^{(1)}} = 0, \tag{A.2.4}$$

$$\frac{\partial \widehat{\beta}^{(r)}}{\partial \widehat{\beta}^{(1)}} = \frac{C_2 C_3^{(r)}}{C_1^{(r)} C_4}, \tag{A.2.5}$$

$$\frac{\partial \widehat{\gamma}^{(r)}}{\partial \widehat{\beta}^{(1)}} = 0, \tag{A.2.6}$$

$$\frac{\partial \widehat{\alpha}}{\partial \widehat{\gamma}^{(1)}} = 0, \tag{A.2.7}$$

$$\frac{\partial \widehat{\beta}^{(r)}}{\partial \widehat{\gamma}^{(1)}} = 0, \tag{A.2.8}$$

$$\frac{\partial \widehat{\gamma}^{(r)}}{\partial \widehat{\gamma}^{(1)}} = \left(\frac{C_1^{(r)}}{C_2} \right)^{1/\widehat{\alpha}}, \tag{A.2.9}$$

$$\frac{\partial \widehat{\alpha}}{\partial \widehat{\theta}_k} = 0, \quad \textit{for } k = 1, \ldots, q \tag{A.2.10}$$

$$\frac{\partial \widehat{\beta}^{(r)}}{\partial \widehat{\theta}_k} = \widehat{\beta}^{(1)} \frac{\partial}{\partial \widehat{\theta}_k} \left(\frac{C_2 C_3^{(r)}}{C_1^{(r)} C_4} \right), \quad \textit{for } k = 1, \ldots, q \tag{A.2.11}$$

$$\frac{\partial \widehat{\gamma}^{(r)}}{\partial \widehat{\theta}_k} = \widehat{\gamma}^{(1)} \frac{\partial}{\partial \widehat{\theta}_k} \left(\frac{C_1^{(r)}}{C_2} \right)^{1/\widehat{\alpha}}, \quad \textit{for } k = 1, \ldots, q \tag{A.2.12}$$

where

$$\frac{\partial}{\partial\widehat{\alpha}}\left(\frac{C_2C_3^{(r)}}{C_1^{(r)}C_4}\right) = \frac{C_2}{C_1^{(r)}}\left[\frac{1}{C_4}\frac{\partial C_3^{(r)}}{\partial\widehat{\alpha}} - \frac{C_3^{(r)}}{C_4^2}\frac{\partial C_4}{\partial\widehat{\alpha}}\right] + \frac{C_3^{(r)}}{C_4}\left[\frac{1}{C_1^{(r)}}\frac{\partial C_2}{\partial\alpha} - \frac{C_2}{\left(C_1^{(r)}\right)^2}\frac{\partial C_1^{(r)}}{\partial\alpha}\right] \tag{A.2.13}$$

$$\frac{\partial}{\partial\widehat{\theta}_k}\left(\frac{C_2C_3^{(r)}}{C_1^{(r)}C_4}\right) = \frac{C_2}{C_1^{(r)}}\left[\frac{1}{C_4}\frac{\partial C_3^{(r)}}{\partial\widehat{\theta}_k} - \frac{C_3^{(r)}}{C_4^2}\frac{\partial C_4}{\partial\widehat{\theta}_k}\right] + \frac{C_3^{(r)}}{C_4}\left[\frac{1}{C_1^{(r)}}\frac{\partial C_2}{\partial\widehat{\theta}_k} - \frac{C_2}{\left(C_1^{(r)}\right)^2}\frac{\partial C_1^{(r)}}{\partial\widehat{\theta}_k}\right] \tag{A.2.14}$$

$$\frac{\partial}{\partial\widehat{\alpha}}\left(\frac{C_1^{(r)}}{C_2}\right)^{1/\widehat{\alpha}} = -\frac{1}{\widehat{\alpha}^2}\left(\frac{C_1^{(r)}}{C_2}\right)^{1/\widehat{\alpha}}\left[\begin{array}{c}\ln\left(\frac{C_1^{(r)}}{C_2}\right) - \\ \widehat{\alpha}\left(\frac{1}{C_1^{(r)}}\frac{\partial C_1^{(r)}}{\partial\widehat{\alpha}} - \frac{1}{C_2}\frac{\partial C_2}{\partial\widehat{\alpha}}\right)\end{array}\right] \tag{A.2.15}$$

$$\frac{\partial}{\partial\widehat{\theta}_k}\left(\frac{C_1^{(r)}}{C_2}\right)^{1/\widehat{\alpha}} = -\frac{1}{\widehat{\alpha}}\left(\frac{C_1^{(r)}}{C_2}\right)^{1/\widehat{\alpha}-1}\left[-\frac{C_1^{(r)}}{C_2^2}\frac{\partial C_2}{\partial\widehat{\theta}_k} + \frac{1}{C_2}\frac{\partial C_1^{(r)}}{\partial\widehat{\theta}_k}\right] \tag{A.2.16}$$

and

$$\frac{\partial C_1^{(r)}}{\partial\widehat{\alpha}} = \sum_{j=0}^{r+q-1}\left|c_j^{(r)}\right|^{\widehat{\alpha}}\ln\left|c_j^{(r)}\right| \tag{A.2.17}$$

$$\frac{\partial C_2}{\partial\widehat{\alpha}} = \sum_{j=0}^{q}\left|\widehat{\theta}_j\right|^{\alpha}\ln\left|\widehat{\theta}_j\right| \tag{A.2.18}$$

$$\frac{\partial C_3^{(r)}}{\partial\widehat{\alpha}} = \sum_{j=0}^{r+q-1} sign\left(c_j^{(r)}\right)\left|c_j^{(r)}\right|^{\widehat{\alpha}}\ln\left|c_j^{(r)}\right| \tag{A.2.19}$$

$$\frac{\partial C_4}{\partial\widehat{\alpha}} = \sum_{j=0}^{q} sign\left(\widehat{\theta}_j\right)\left|\widehat{\theta}_j\right|^{\widehat{\alpha}}\ln\left|\widehat{\theta}_j\right| \tag{A.2.20}$$

$$\frac{\partial C_1^{(r)}}{\partial\widehat{\theta}_k} = \widehat{\alpha}\sum_{j=0}^{r+q-1} sign\left(c_j^{(r)}\right)\left|c_j^{(r)}\right|^{\widehat{\alpha}-1}\frac{\partial c_j^{(r)}}{\partial\widehat{\theta}_k} \tag{A.2.21}$$

$$\frac{\partial C_2}{\partial\widehat{\theta}_k} = \widehat{\alpha} sign\left(\widehat{\theta}_k\right)\left|\widehat{\theta}_k\right|^{\widehat{\alpha}-1} \tag{A.2.22}$$

$$\frac{\partial C_3^{(r)}}{\partial\widehat{\theta}_k} = \widehat{\alpha}\sum_{j=0}^{r+q-1} sign^2\left(c_j^{(r)}\right)\left|c_j^{(r)}\right|^{\widehat{\alpha}-1}\frac{\partial c_j^{(r)}}{\partial\widehat{\theta}_k} \tag{A.2.23}$$

$$\frac{\partial C_4}{\partial\widehat{\theta}_k} = \widehat{\alpha}\left|\theta_k\right|^{\widehat{\alpha}-1} sign^2\left(\widehat{\theta}_k\right) \tag{A.2.24}$$

and

$$\frac{\partial c_j^{(r)}}{\partial \widehat{\theta}_k} = \begin{cases} 1 & \textit{for } j = k, \ldots, k+r-1 \textit{ and } 0 \leq k \leq q \\ 0 & \textit{otherwise} \end{cases} \tag{A.2.25}$$

Before proceeding with the proof of Theorem 4.2.2, we make the following observations

1. that l'Hopital's rule can be used to show that $\lim_{x \to 0} x^{\alpha} \ln x^{\alpha} = 0$ for all $\alpha \in (0, 2]$, and therefore that $|x|^{\alpha} \ln |x|^{\alpha}$ is a continuous function
2. that $sign(x)|x|$ is a continuous function
3. that $C_1^{(r)}$ and C_2 are strictly positive functions.

Proof. Define the function $g^{(r)}(\varkappa) : \mathbb{R}^{3+q} \to \mathbb{R}^3$ by

$$g^{(r)}(\varkappa) = \left(\alpha, \beta^{(r)}, \gamma^{(r)}\right)' \tag{A.2.26}$$

where

$$\varkappa = \left(\alpha, \beta^{(1)}, \gamma^{(1)}, \theta_1, \ldots, \theta_q\right)' \tag{A.2.27}$$

and

$$\beta^{(r)} = \frac{C_2 C_3^{(r)}}{C_1^{(r)} C_4} \beta^{(1)}, \tag{A.2.28}$$

$$\gamma^{(r)} = \frac{C_2 C_3^{(r)}}{C_1^{(r)} C_4} \gamma^{(1)} \tag{A.2.29}$$

and $C_1^{(r)}, C_2, C_3^{(r)}, C_4$ are defined in (4.2.11) – (4.2.14). Let $g_i^{(r)}(\varkappa)$, $i = 1, \ldots, 3$ denote the component functions of $g^{(r)}(\varkappa)$.

For sample size T we have

$$\sqrt{T}\left(\widehat{\varkappa} - \varkappa_0\right) \xrightarrow{d} N\left(0, V_{\widehat{\varkappa}}\right), \tag{A.2.30}$$

where $V_{\widehat{\varkappa}}$ is given in (4.2.22).

In (4.2.16) and (4.2.17) it was established that

$$g^{(r)}(\varkappa_0) = \omega_0^{(r)}. \tag{A.2.31}$$

To prove the theorem it remains to be shown that each component function $g_i^{(r)}$ has a non-zero differential at $\widehat{\varkappa} = \varkappa_0$, (Theorem B.2.3). To show that $g_i^{(r)}$ has a non-zero differential

it is sufficient to show that the partial derivatives exist at $\widehat{\varkappa} = \varkappa_0$, are continuous at $\widehat{\varkappa} = \varkappa_0$ and that at least one of the partial derivatives is non-zero at $\widehat{\varkappa} = \varkappa_0$.

If $C_4 \neq 0$, and either $\widehat{\alpha} \geq 1$ or $c_j^{(r)} \neq 0$ for $j = 0, \ldots, r+q+1$ and $\theta_j \neq 0$ for $j = 1, \ldots, q$, then it is clear from the results of Lemma A.2.1 that all partial derivatives $\dfrac{\partial \widehat{\omega}_i^{(r)}}{\partial \widehat{\varkappa}_j}$ exist. Moreover, the partial derivatives being the sum and/or product of continuous functions are themselves continuous.

It remains to find at least one partial derivative for each of the component functions $g_i^{(r)}$ which is non-zero at $\widehat{\varkappa} = \varkappa_0$.

For $\widehat{\alpha} = g_1^{(r)}$, $\dfrac{\partial \widehat{\alpha}}{\partial \widehat{\alpha}} = 1$ is non-zero.

For $\widehat{\beta}^{(r)} = g_2^{(r)}$, if $C_3^{(r)}, C_4^{(r)} \neq 0$, then $\dfrac{\partial \widehat{\beta}^{(r)}}{\partial \widehat{\beta}^{(1)}} = \dfrac{C_2 C_3^{(r)}}{C_1^{(r)} C_4}$ is non-zero

For $\gamma^{(r)} = g_3^{(r)}$, $\dfrac{\partial \widehat{\gamma}^{(r)}}{\partial \widehat{\gamma}^{(1)}} = \left(\dfrac{C_1^{(r)}}{C_2} \right)^{1/\widehat{\alpha}}$ is non-zero.

This completes the proof. ■

A.3 Proof of Theorem 4.3.1

In order to prove Theorem 4.3.1 it is necessary to prove first that the partial derivatives of the quantile function of a stable distribution are continuous with respect to the parameter of the distribution.

Let $\xi_p(\omega)$ denote the p^{th} quantile of a stable distribution with parameters

$$\omega = (\alpha, \beta, \gamma, \delta) \tag{A.3.1}$$

using the $S_\alpha^0(\beta, \gamma, \delta)$ parameterisation of Nolan (1998). To show that the partial derivatives of $\xi_p(\omega)$ with respect to the parameters γ and δ are continuous follows immediately from the location and scale properties of those parameters.

Lemma A.3.1 *Let $X \sim S_\alpha(\beta, \gamma, \delta)$ and $X^* \sim S_\alpha(\beta, 1, 0)$ be stable random variables. Let ξ_p and ξ_p^* denote the p^{th} quantile of X and X^* respectively. Then*

$$\xi_p = \gamma \xi_p^* + \delta \tag{A.3.2}$$

(from Lemma E.1.2). Since ξ_p^* does not depend on either γ or δ, it follows that

$$\frac{\partial \xi_p}{\partial \gamma} = \xi_p^* \tag{A.3.3}$$

and

$$\frac{\partial \xi_p}{\partial \delta} = 1. \tag{A.3.4}$$

Remark A.3.1 *It was not possible in this thesis to prove that the partial derivatives* $\frac{\partial \xi_p}{\partial \alpha}$ *and* $\frac{\partial \xi_p}{\partial \beta}$ *are continuous functions however, the numerical evidence does support this assertion. Examples of plots of* ξ_p *as a function of* α *and* β *are shown in Figure A.3.1. In each case the plot appears to be of a continuously differentiable function. In each case,* ξ_p *appears to be a strictly montonic function of* β*, but* ξ_p *is not a strictly monotonic function of* α*. Whilst the absence of strict monotonicity in* ξ_p *as a function of* α *is not surprising at* $p = 0.5$ *and* $\beta = 0$ *where* ξ_p *is zero, the case where* $p = 0.7$ *and* $\beta = 0$ *in Figure A.3.1(b) is a non-constant function with distinct maxima and minima.*

We are now able to complete the proof of Theorem 4.3.1.

Proof. Define the function $g_p^{(r)}\left(\omega^{(r)}\right) : \mathbb{R}^4 \to \mathbb{R}$ by

$$g_p^{(r)}\left(\omega^{(r)}\right) = \widetilde{\zeta}_p^{(r)} \tag{A.3.5}$$

where

$$\omega^{(r)} = \left(\alpha, \beta^{(r)}, \gamma^{(r)}\right)$$

and $\widetilde{\zeta}_p^{(r)}$ is the log quantile difference of the distribution $S_\alpha\left(\beta^{(r)}, \gamma^{(r)}, 0\right)$ at quantile levels $p = (p_1, p_2)$.

For sample size T we have

$$\sqrt{T}\left(\widehat{\omega}^{(r)} - \omega_0^{(r)}\right) \xrightarrow{d} N\left(0, V_{\widehat{\omega}^{(r)}}\right), \tag{A.3.6}$$

where $V_{\widehat{\omega}^{(r)}}$ is given in (4.2.30) and $\omega_0^{(r)}$ is the true value of the parameters $\omega^{(r)}$.

By definition

$$g_{1;p}\left(\omega_0^{(r)}\right) = \zeta_{p;0}^{(r)}. \tag{A.3.7}$$

To prove the theorem it remains to be shown that the function $g_p^{(r)}$ has a non-zero differential at $\widehat{\omega}^{(r)} = \omega_0^{(r)}$, (Theorem B.2.3). To show that $g_p^{(r)}$ has a non-zero differential it is sufficient to show that the partial derivatives exist at $\widehat{\omega}^{(r)} = \omega_0^{(r)}$, are continuous at $\widehat{\omega}^{(r)} = \omega_0^{(r)}$ and that at least one of the partial derivatives is non-zero at $\widehat{\omega}^{(r)} = \omega_0^{(r)}$.

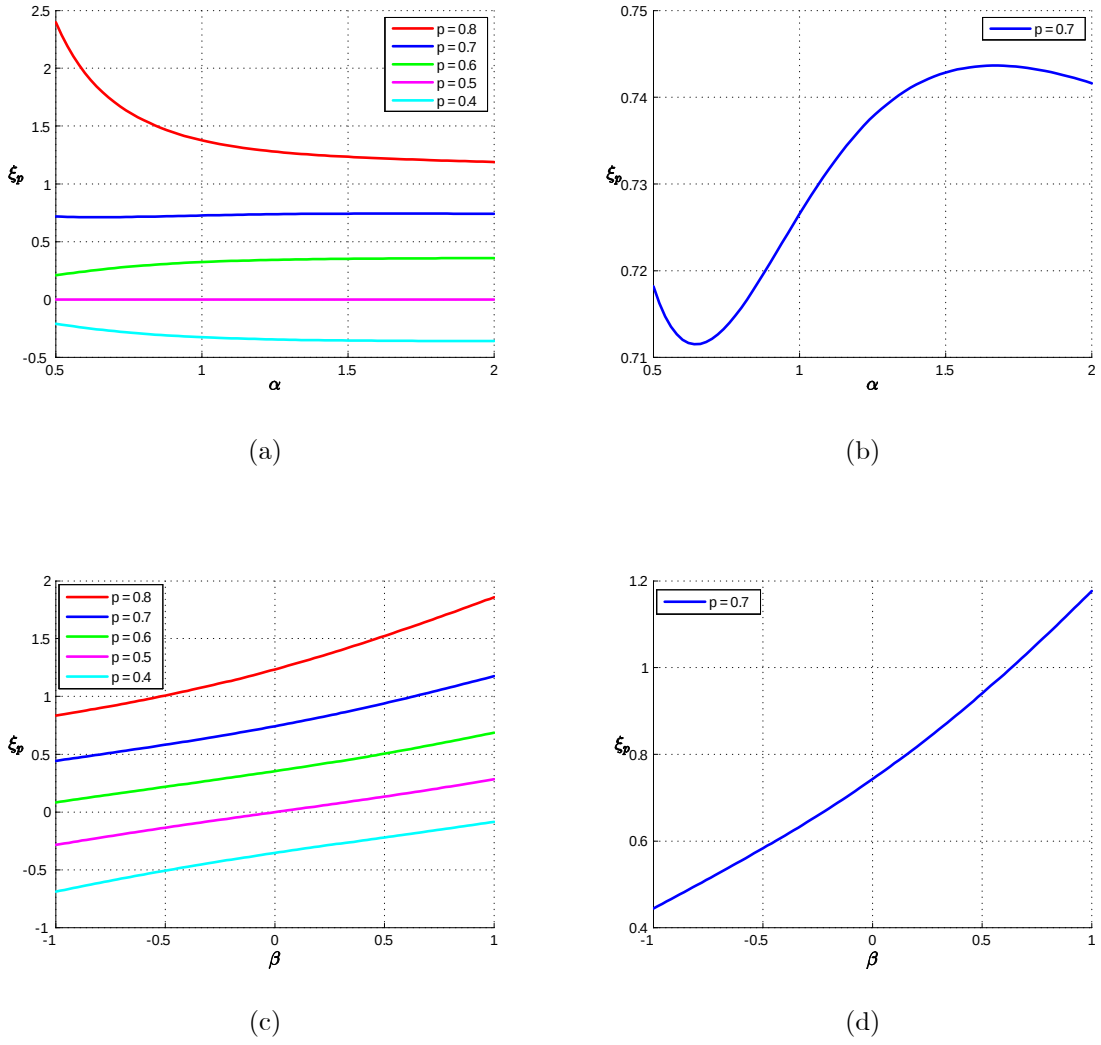


Figure A.3.1: A plot of the quantiles of the stable distribution (a), (b) as a function of α where $\beta = 0$ and (c), (d) as a function of β where $\alpha = 1.5$.

We have shown in Lemma A.3.1 that the partial derivative $\dfrac{\partial\widetilde{\xi}_p^{*(r)}}{\partial\widehat{\gamma}^{(r)}}$ is continuous at $\widehat{\omega}^{(r)} = \omega_0^{(r)}$ and assumed in Remark A.3.1 that the partial derivatives $\dfrac{\partial\widetilde{\xi}_p^{*(r)}}{\partial\widehat{\alpha}}$ and $\dfrac{\partial\widetilde{\xi}_p^{*(r)}}{\partial\widehat{\beta}^{(r)}}$ are continuous at $\widehat{\omega}^{(r)} = \omega_0^{(r)}$. The partial derivative $\dfrac{\partial\widetilde{\xi}_p^{*(r)}}{\partial\widehat{\gamma}^{(r)}}$ is non-zero except where $\widetilde{\xi}_p^{*(r)}$ is zero. The partial derivative $\dfrac{\partial\widetilde{\xi}_p^{*(r)}}{\partial\widehat{\beta}^{(r)}}$ is assumed to be non-zero.

The partial derivatives $\dfrac{\partial\widetilde{\zeta}_p^{(r)}}{\partial\widetilde{\xi}_{p_j}^{*(r)}}$ are given by

$$\frac{\partial\widetilde{\zeta}_p^{(r)}}{\partial\widetilde{\xi}_{p_j}^{*(r)}} = \begin{cases} \dfrac{-1}{\widetilde{\xi}_{p_2}^{*(r)} - \widetilde{\xi}_{p_1}^{*(r)}} & j = 1 \\ \dfrac{1}{\widetilde{\xi}_{p_2}^{*(r)} - \widetilde{\xi}_{p_1}^{*(r)}} & j = 2 \end{cases}$$

which are clearly continuous and non-zero wherever $p_1 \neq p_2$.

This completes the proof. ■

A.4 Proof of Theorem 4.7.2

The following lemma provides formulae for each of the partial derivatives $\dfrac{\partial\check{\zeta}_p^{(r)}}{\partial\widehat{\chi}_j}$. Derivation of these formulae is achieved through the application of the standard chain and product rules of differentiation and is presented without proof. Note that some of these formulae also appeared in Lemma A.2.1.

Lemma A.4.1 *The partial derivatives* $\dfrac{\partial\check{\zeta}_p^{(r)}}{\partial\widehat{\chi}_j}$ *are given by*

$$\frac{\partial\check{\zeta}_p^{(r)}}{\partial\widehat{\alpha}} = -\widehat{\alpha}^{-2}\ln\left(\frac{C_1^{(r)}}{C_2}\right) + \widehat{\alpha}^{-1}\left(\frac{1}{C_1^{(r)}}\frac{\partial C_1^{(r)}}{\partial\widehat{\alpha}} - \frac{1}{C_2}\frac{\partial C_2}{\partial\widehat{\alpha}}\right), \tag{A.4.1}$$

$$\frac{\partial\check{\zeta}_p^{(r)}}{\partial\widehat{\zeta}_p^{(1)}} = 1 \tag{A.4.2}$$

$$\frac{\partial\check{\zeta}_p^{(r)}}{\partial\widehat{\theta}_k} = \widehat{\alpha}^{-1}\left(\frac{1}{C_1^{(r)}}\frac{\partial C_1^{(r)}}{\partial\widehat{\theta}_k} - \frac{1}{C_2}\frac{\partial C_2}{\partial\widehat{\theta}_k}\right), \quad \textit{for } k = 1, \ldots, q \tag{A.4.3}$$

where

$$\frac{\partial C_1^{(r)}}{\partial \widehat{\alpha}} = \sum_{j=0}^{r+q-1} \left|c_j^{(r)}\right|^{\widehat{\alpha}} \ln \left|c_j^{(r)}\right| \tag{A.4.4}$$

$$\frac{\partial C_2}{\partial \widehat{\alpha}} = \sum_{j=0}^{q} \left|\widehat{\theta}_j\right|^{\alpha} \ln \left|\widehat{\theta}_j\right| \tag{A.4.5}$$

$$\frac{\partial C_1^{(r)}}{\partial \widehat{\theta}_k} = \widehat{\alpha} \sum_{j=0}^{r+q-1} sign\left(c_j^{(r)}\right) \left|c_j^{(r)}\right|^{\widehat{\alpha}-1} \frac{\partial c_j^{(r)}}{\partial \widehat{\theta}_k} \tag{A.4.6}$$

$$\frac{\partial C_2}{\partial \widehat{\theta}_k} = \widehat{\alpha} sign\left(\widehat{\theta}_k\right) \left|\widehat{\theta}_k\right|^{\widehat{\alpha}-1} \tag{A.4.7}$$

and

$$\frac{\partial c_j^{(r)}}{\partial \widehat{\theta}_k} = \begin{cases} 1 & for\ j = k, \ldots, k+r-1\ and\ 0 \leq k \leq q \\ 0 & otherwise \end{cases} \tag{A.4.8}$$

Before proceeding with the proof of Theorem 4.7.2, we make the following observations

1. that l'Hopital's rule can be used to show that $\lim_{x\to 0} x^{\alpha} \ln x^{\alpha} = 0$ for all $\alpha \in (0, 2]$, and therefore that $|x|^{\alpha} \ln |x|^{\alpha}$ is a continuous function
2. that $sign(x)|x|$ is a continuous function
3. that $C_1^{(r)}$ and C_2 are strictly positive functions.

Proof. Define the function $g^{(r)}(\chi) : \mathbb{R}^{2+q} \to \mathbb{R}$ by

$$g^{(r)}(\chi) = \check{\zeta}_p^{(r)} \tag{A.4.9}$$

where

$$\chi = \left(\alpha, \zeta_p^{(1)}, \theta_1, \ldots, \theta_q\right)' \tag{A.4.10}$$

and

$$\check{\zeta}_p^{(r)} = \widehat{\alpha}^{-1} \ln \left(\frac{C_1^{(r)}(\chi)}{C_2(\chi)}\right) + \widehat{\zeta}_p^{(1)} \tag{A.4.11}$$

and $C_1^{(r)}, C_2$ are defined in (4.7.2) – (4.7.3).

For sample size T we have

$$\sqrt{T}\left(\widehat{\chi} - \chi_0\right) \xrightarrow{d} N\left(0, V_{\widehat{\chi}}\right), \tag{A.4.12}$$

where $V_{\widehat{\chi}}$ is given in (4.2.22).

By assumption, either $\beta = 0$ or $\theta_j \geq 0$ for $j = 1, \ldots, q$, therefore we have from Corollary 2.2.2 that

$$g^{(r)}(\chi_0) = \zeta_{p;0}^{(r)}. \tag{A.4.13}$$

To prove the theorem it remains to be shown that $g^{(r)}$ has a non-zero differential at $\widehat{\chi} = \chi_0$, (Theorem B.2.2). To show that $g^{(r)}$ has a non-zero differential it is sufficient to show that the partial derivatives exist at $\widehat{\chi} = \chi_0$, are continuous at $\widehat{\chi} = \chi_0$ and that at least one of the partial derivatives is non-zero at $\widehat{\chi} = \chi_0$.

By assumption, either $\widehat{\alpha} \geq 1$ or $c_j^{(r)} \neq 0$ for $j = 0, \ldots, r+q+1$ and $\theta_j \neq 0$ for $j = 1, \ldots, q$,and thus it is clear from the results of Lemma A.4.1 that all partial derivatives $\dfrac{\partial \check{\zeta}_p^{(r)}}{\partial \widehat{\chi}_j}$ exist. Moreover, the partial derivatives being the sum and/or product of continuous functions are themselves continuous.

Since the partial derivative $\dfrac{\partial \check{\zeta}_p^{(r)}}{\partial \widehat{\zeta}_p^{(1)}}$ is always non-zero, the proof is complete. ■

A.5 Proof of Theorem 5.1.1

Define the function $g^{(r)}\left(\widehat{\lambda}_0, \widehat{\lambda}_1\right) : \mathbb{R}^2 \to \mathbb{R}$ by

$$g^{(r)}\left(\widehat{\lambda}_0, \widehat{\lambda}_1\right) = \frac{1}{2} \ln\left(r\widehat{\lambda}_0 + 2(r-1)\widehat{\lambda}_1\right) \tag{A.5.1}$$

Lemma A.5.1 *Under the assumption that the mean of the invertible* MA(1) *process* $\{X_t\}$ *is known to be zero*

$$\begin{aligned} \left(\sigma^{(r)}\right)^2 &= V\left[S_t^{(r)}\right] \\ &= rE\left[X_j^2\right] + 2(r-1)E\left[X_j X_{j-1}\right] \\ &= r\lambda_0 + 2(r-1)\lambda_1 \end{aligned} \tag{A.5.2}$$

and therefore

$$\begin{aligned} \ln \sigma^{(r)} &= \frac{1}{2} \ln\left(r\lambda_0 + 2(r-1)\lambda_1\right) \\ &= g^{(r)}(\lambda_0, \lambda_1) \end{aligned} \tag{A.5.3}$$

We now commence the proof of Theorem 5.1.1.

Proof. Let $\widehat{\lambda}_0, \widehat{\lambda}_1$ denote the sample autocovariance function of $\{X_{d;t}\}$ and lags $0, 1$ respectively. From Proposition 7.3.4 in Brockwell and Davis (1991), we have as the sample

size $n \to \infty$ that

$$\sqrt{n}\left(\begin{pmatrix} \widehat{\lambda}_0 \\ \widehat{\lambda}_1 \end{pmatrix} - \begin{pmatrix} \lambda_0 \\ \lambda_1 \end{pmatrix}\right) \xrightarrow{d} N\left(0, V_\lambda\right) \tag{A.5.4}$$

where

$$V_\lambda = \begin{pmatrix} 2\lambda_0^2 + 4\lambda_1^2 & 4\lambda_0\lambda_1 \\ 4\lambda_0\lambda_1 & \lambda_0^2 + 3\lambda_1^2 \end{pmatrix}. \tag{A.5.5}$$

From Lemma A.5.1, we have

$$g^{(r)}\left(\lambda_0, \lambda_1\right) = \ln \sigma^{(r)} \tag{A.5.6}$$

Let $\widehat{\lambda} = \left(\widehat{\lambda}_0, \widehat{\lambda}_1\right)'$ and $\lambda = (\lambda_0, \lambda_1)'$. To prove the theorem it remains to be shown that $g^{(r)}$ has a non-zero differential at $\widehat{\lambda} = \lambda$, (Theorem B.2.2). To show that $g^{(r)}$ has a non-zero differential it is sufficient to show that the partial derivatives exist at $\widehat{\lambda} = \lambda$, are continuous at $\widehat{\lambda} = \lambda$ and that at least one of the partial derivatives is non-zero at $\widehat{\lambda} = \lambda$.

The partial derivatives of $g^{(r)}$ are given by

$$\frac{\partial g^{(r)}}{\partial \widehat{\lambda}_0} = \left(\frac{r}{2}\right) \frac{1}{r\widehat{\lambda}_0 + 2\left(r-1\right)\widehat{\lambda}_1}, \tag{A.5.7}$$

$$\frac{\partial g^{(r)}}{\partial \widehat{\lambda}_1} = \left(r-1\right) \frac{1}{r\widehat{\lambda}_0 + 2\left(r-1\right)\widehat{\lambda}_1}. \tag{A.5.8}$$

The partial derivatives $\dfrac{\partial g^{(r)}}{\partial \widehat{\lambda}_0}$ and $\dfrac{\partial g^{(r)}}{\partial \widehat{\lambda}_1}$ exist for all $\widehat{\lambda}$ except where

$$\frac{\widehat{\lambda}_1}{\widehat{\lambda}_0} = \frac{r}{2\left(r-1\right)} > \frac{1}{2}. \tag{A.5.9}$$

For an invertible MA(1) process,

$$\frac{\lambda_1}{\lambda_0} = \frac{\theta_1}{1+\theta_1^2} < \frac{1}{2}, \tag{A.5.10}$$

and therefore the partial derivatives exist at $\widehat{\lambda} = \lambda$.

Clearly the partial derivatives are continuous at $\widehat{\lambda} = \lambda$, and at least one of the partial derivatives is non-zero at $\widehat{\lambda} = \lambda$, which completes the proof. ■

A.6 Proof of Theorem 5.2.1

Applying Theorems 4.3.1 and 4.2.2, to the Mean Symmetric Log Quantile Difference Estimator $\overline{\zeta}_{p;d}^{(r)}$, we have as the sample size $T \to \infty$

$$\sqrt{T}\left(\overline{\zeta}_{p;d}^{(r)} - \zeta_{p;d;0}^{(r)}\right) \xrightarrow{} N\left(0, V_{\overline{\zeta}_{p;d}^{(r)}}\right) \tag{A.6.1}$$

where

$$V_{\overline{\zeta}_{p;d}^{(r)}} = D_{\overline{\zeta}_{p;d}^{(r)}} D_{\overline{\xi}_{p;d}^{*(r)}} D_{\overline{\omega}_d^{(r)}} V_{\overline{\varkappa}_d^{(1)}} D'_{\overline{\omega}_d^{(r)}} D'_{\overline{\xi}_{p;d}^{*(r)}} D'_{\overline{\zeta}_{p;d}^{(r)}}, \tag{A.6.2}$$

$$V_{\overline{\varkappa}_d^{(1)}} = \begin{pmatrix} D_{\overline{\omega}_d^{(1)}} V_{\widehat{\xi}_{p_M;d}} D'_{\overline{\omega}_d^{(1)}} & 0 \\ 0 & V_{\overline{\theta}} \end{pmatrix} \tag{A.6.3}$$

and

$$D_{\overline{\zeta}_{p;d}^{(r)}} = \left[\frac{\partial \overline{\zeta}_{p;d}^{(r)}}{\partial \overline{\xi}_{p;d}^{*(r)}} \Big|_{\overline{\xi}_{p;d}^{*(r)} = \overline{\xi}_{p;d;0}^{*(r)}} \right]_{j=1,\ldots,2}, \tag{A.6.4}$$

$$D_{\overline{\xi}_{p;d}^{*(r)}} = \left[\frac{\partial \overline{\xi}_{p;d}^{*(r)}}{\partial \overline{\omega}_j^{(r)}} \Big|_{\overline{\omega}_d^{(r)} = \omega_{d;0}^{(r)}} \right]_{i=1,\ldots,2, j=1,\ldots,3}, \tag{A.6.5}$$

$$D_{\overline{\omega}_d^{(r)}} = \left[\frac{\partial \overline{\omega}_{d;i}^{(r)}}{\partial \overline{\varkappa}_{d;j}^{(1)}} \Big|_{\overline{\varkappa}_d^{(1)} = \varkappa_{d;0}^{(1)}} \right]_{i=1,\ldots,3, j=1,\ldots,4} \tag{A.6.6}$$

$$D_{\overline{\omega}_d^{(1)}} = \left[\frac{\partial \overline{\omega}_{d;i}^{(1)}}{\partial \widehat{\xi}_{p_{M_j};d}} \Big|_{\widehat{\xi}_{p_M;d} = \xi_{p_M;d;0}} \right]_{i=1,\ldots,3, j=1,\ldots,5} \tag{A.6.7}$$

and $\zeta_{p;d;0}^{(r)}, \overline{\xi}_{p;d;0}^{*(r)}, \omega_{d;0}^{(r)}, \varkappa_{d;0}^{(1)}, \xi_{p_M;d;0}$ are the true values of $\zeta_{p;d}^{(r)}, \xi_{p;d}^{*(r)} \omega_d^{(r)}, \varkappa_d^{(1)}, \xi_{p_M;d}$ respectively.

Throughout this proof we write $z \propto \gamma_d^a$ to mean that $z = c \left(\gamma_d^{(1)}\right)^a$ for some constant $c \neq 0$ which has the same value for all $\gamma_d^{(1)} > 0$.

Proof. Some consideration of each of the terms in the partial derivative matrices gives

$$D_{\overline{\zeta}_{p;d}^{(r)}} \propto \begin{pmatrix} \gamma_d^{-1} & \gamma_d^{-1} \end{pmatrix}, \tag{A.6.8}$$

$$D_{\overline{\xi}_{p;d}^{*(r)}} \propto \begin{pmatrix} \gamma_d^1 & \gamma_d^1 & \gamma_d^0 \\ \gamma_d^1 & \gamma_d^1 & \gamma_d^0 \end{pmatrix} \tag{A.6.9}$$

$$D_{\overline{\omega}_d^{(r)}} \propto \begin{pmatrix} \gamma_d^0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \gamma_d^1 & 0 & \gamma_d^0 & \gamma_d^1 \end{pmatrix} \tag{A.6.10}$$

$$D_{\overline{\omega}_d^{(1)}} \propto \begin{pmatrix} \gamma_d^{-1} & \gamma_d^{-1} & 0 & \gamma_d^{-1} & \gamma_d^{-1} \\ 0 & 0 & 0 & 0 & 0 \\ \gamma_d^0 & \gamma_d^0 & 0 & \gamma_d^0 & \gamma_d^0 \end{pmatrix} \tag{A.6.11}$$

Each element in the matrix $V_{\widehat{\xi}_{p_M;d}}$ is inversely proportional to the square of the density

function and therefore

$$V_{\widehat{\xi}_{p_M;d}} \propto \gamma_d^2. \tag{A.6.12}$$

Finally, from $(C.3.9)$,it can be seen that

$$V_{\overline{\theta}} \propto \gamma_d^0. \tag{A.6.13}$$

Use of the above formulae gives

$$V_{\overline{\varkappa}_d^{(1)}} \propto \begin{pmatrix} \gamma_d^0 & 0 & \gamma_d^1 & 0 \\ 0 & 0 & 0 & 0 \\ \gamma_d^1 & 0 & \gamma_d^2 & 0 \\ 0 & 0 & 0 & \gamma_d^0 \end{pmatrix} \tag{A.6.14}$$

and

$$V_{\overline{\zeta}_{p;d}^{(r)}} \propto \gamma_d^0 \tag{A.6.15}$$

which proves the theorem. ■

A.7 Proof of Theorem 5.3.1

Recall that the model equation of the unobserved process is

$$X_d - \mu = \phi_1 (X_{d-1} - \mu) + u_d + \psi_1 u_{d-1}, \tag{A.7.1}$$

that the model equation of the observed process is

$$Y_d - \mu = \phi_1 (Y_{d-1} - \mu) + v_d + \eta_1 v_{d-1}, \tag{A.7.2}$$

and that the unobserved and observed processes are related by

$$Y_d = X_d + w_d \tag{A.7.3}$$

where $\{u_d\}, \{v_d\}$ and $\{w_d\}$ are IID sequences of random variables with distributions $N\left(0, \sigma_u^2\right), N\left(0, \sigma_v^2\right)$ and $N\left(0, \sigma_w^2\right)$ respectively. We assume that u_d and w_d are uncorrelated and that σ_w^2 is known.

Let $\Omega = \left(\mu, \phi_1, \psi_1, \sigma_u^2\right)'$ denote the vector of parameters in the model for $\{X_d\}$ and $\Theta = \left(\mu, \phi_1, \eta_1, \sigma_v^2\right)'$ denote the vector of parameters in the model for $\{Y_d\}$. Define the function $g(\Theta) : \mathbb{R}^4 \to \mathbb{R}^4$ by

$$g(\Theta) = (g_1(\Theta), g_2(\Theta), g_3(\Theta), g_4(\Theta))' \tag{A.7.4}$$

where

$$g_1(\Theta) = \mu \tag{A.7.5}$$

$$g_2(\Theta) = \phi_1 \tag{A.7.6}$$

$$g_3(\Theta) = \psi_1 \tag{A.7.7}$$

$$g_4(\Theta) = \sigma_u^2 \tag{A.7.8}$$

and

$$\psi_1 = \begin{cases} \dfrac{\kappa_0/\kappa_1 + \sqrt{(\kappa_0/\kappa_1)^2 - 4}}{2} & \text{if } \kappa_1 < 0 \\ 0 & \text{if } \kappa_1 = 0 \\ \dfrac{\kappa_0/\kappa_1 - \sqrt{(\kappa_0/\kappa_1)^2 - 4}}{2} & \text{if } \kappa_1 > 0 \end{cases}, \tag{A.7.9}$$

$$\sigma_u^2 = \frac{\kappa_0}{1+\psi_1^2} \tag{A.7.10}$$

$$\kappa_0 = \left(1+\eta_1^2\right)\sigma_v^2 - \left(1+\phi_1^2\right)\sigma_w^2 \tag{A.7.11}$$

$$\kappa_1 = \eta_1\sigma_v^2 + \phi_1\sigma_w^2 \tag{A.7.12}$$

Lemma A.7.1 *The partial derivatives of the function $g(\Theta)$ are given by*

$$\frac{\partial g_1}{\partial \phi_1}, \frac{\partial g_1}{\partial \eta_1}, \frac{\partial g_1}{\partial \sigma_v^2}, \frac{\partial g_2}{\partial \mu}, \frac{\partial g_2}{\partial \eta_1}, \frac{\partial g_2}{\partial \sigma_v^2}, \frac{\partial g_3}{\partial \mu}, \frac{\partial g_4}{\partial \mu} = 0 \tag{A.7.13}$$

$$\frac{\partial g_1}{\partial \mu}, \frac{\partial g_2}{\partial \phi_1} = 1 \tag{A.7.14}$$

and

$$\frac{\partial g_3}{\partial \phi_1} = \frac{\partial g_3}{\partial \kappa_0}\frac{\partial \kappa_0}{\partial \phi_1} + \frac{\partial g_3}{\partial \kappa_1}\frac{\partial \kappa_1}{\partial \phi_1} \tag{A.7.15}$$

$$\frac{\partial g_3}{\partial \eta_1} = \frac{\partial g_3}{\partial \kappa_0}\frac{\partial \kappa_0}{\partial \eta_1} + \frac{\partial g_3}{\partial \kappa_1}\frac{\partial \kappa_1}{\partial \eta_1} \tag{A.7.16}$$

$$\frac{\partial g_3}{\partial \sigma_v^2} = \frac{\partial g_3}{\partial \kappa_0}\frac{\partial \kappa_0}{\partial \sigma_v^2} + \frac{\partial g_3}{\partial \kappa_1}\frac{\partial \kappa_1}{\partial \sigma_v^2} \tag{A.7.17}$$

$$\frac{\partial g_4}{\partial \phi_1} = \frac{\partial g_4}{\partial \kappa_0}\frac{\partial \kappa_0}{\partial \phi_1} + \frac{\partial g_4}{\partial \kappa_1}\frac{\partial \kappa_1}{\partial \phi_1} \tag{A.7.18}$$

$$\frac{\partial g_4}{\partial \eta_1} = \frac{\partial g_4}{\partial \kappa_0}\frac{\partial \kappa_0}{\partial \eta_1} + \frac{\partial g_4}{\partial \kappa_1}\frac{\partial \kappa_1}{\partial \eta_1} \tag{A.7.19}$$

$$\frac{\partial g_4}{\partial \sigma_v^2} = \frac{\partial g_4}{\partial \kappa_0}\frac{\partial \kappa_0}{\partial \sigma_v^2} + \frac{\partial g_4}{\partial \kappa_1}\frac{\partial \kappa_1}{\partial \sigma_v^2} \tag{A.7.20}$$

where

$$\frac{\partial g_3}{\partial \kappa_0} = \begin{cases} \dfrac{1-\kappa_0/\sqrt{\kappa_0^2-4\kappa_1^2}}{2\kappa_1} & \text{if } \kappa_1 \neq 0 \\ 0 & \text{if } \kappa_1 = 0 \end{cases} \tag{A.7.21}$$

$$\frac{\partial g_3}{\partial \kappa_1} = \begin{cases} \dfrac{\kappa_0}{2\kappa_1^2}\left(\kappa_0/\sqrt{\left(\kappa_0^2-4\kappa_1^2\right)}-1\right) & \text{if } \kappa_1 \neq 0 \\ 1 & \text{if } \kappa_1 = 0 \end{cases} \tag{A.7.22}$$

$$\frac{\partial g_4}{\partial \kappa_0} = \begin{cases} \dfrac{2\kappa_1^2}{4\kappa_1^2-\kappa_0^2+\kappa_0\sqrt{\left(\kappa_0^2-4\kappa_1^2\right)}} & \text{if } \kappa_1 \neq 0 \\ 1 & \text{if } \kappa_1 = 0 \end{cases} \tag{A.7.23}$$

$$\frac{\partial g_4}{\partial \kappa_1} = \begin{cases} \dfrac{-2\kappa_1}{\sqrt{\left(\kappa_0^2-4\kappa_1^2\right)}} & \text{if } \kappa_1 \neq 0 \\ 0 & \text{if } \kappa_1 = 0 \end{cases} \tag{A.7.24}$$

and

$$\frac{\partial \kappa_0}{\partial \phi_1} = -2\phi_1\sigma_w^2 \tag{A.7.25}$$

$$\frac{\partial \kappa_0}{\partial \eta_1} = 2\eta_1\sigma_v^2 \tag{A.7.26}$$

$$\frac{\partial \kappa_0}{\partial \sigma_v^2} = \left(1+\eta_1^2\right) \tag{A.7.27}$$

$$\frac{\partial \kappa_1}{\partial \phi_1} = \sigma_w^2 \tag{A.7.28}$$

$$\frac{\partial \kappa_1}{\partial \eta_1} = \sigma_v^2 \tag{A.7.29}$$

$$\frac{\partial \kappa_1}{\partial \sigma_v^2} = \eta_1 \tag{A.7.30}$$

Proof. Let $\widehat{\Theta} = \left(\widehat{\mu}, \widehat{\phi}_1, \widehat{\eta}_1, \widehat{\sigma}_v^2\right)'$ denote the maximum likelihood estimators of the parameters Θ and let Ω_0, Θ_0 denote the true values of Ω, Θ respectively. From Brockwell and Davis (1991), Sections 7.1, 8.7 and 8.8, the maximum likelihood estimators of the parameters of an ARMA$(1,1)$ process have the following asymptotic properties as the sample size $n \to \infty$

$$\sqrt{n}\left(\widehat{\Theta} - \Theta_0\right) \xrightarrow{d} N\left(0, V\right) \tag{A.7.31}$$

where

$$V = \begin{pmatrix} \sigma_v^2 \dfrac{(1+\eta_1)^2}{(1-\phi_1)^2} & 0 & 0 \\ 0 & \begin{pmatrix} \left(1-\phi_1^2\right)^{-1} & (1+\phi_1\eta_1)^{-1} \\ (1+\phi_1\eta_1)^{-1} & \left(1-\eta_1^2\right)^{-1} \end{pmatrix}^{-1} & 0 \\ 0 & 0 & 2\left(\sigma_v^2\right)^2 \end{pmatrix}. \tag{A.7.32}$$

In Section 5.3 it was established that

$$g\left(\Theta_0\right) = \Omega_0. \tag{A.7.33}$$

To prove the theorem it remains to be shown that each component function g_i has a non-zero differential at $\Theta = \Theta_0$, (Theorem B.2.3). To show that g_i has a non-zero differential it is sufficient to show that the partial derivatives exist at $\Theta = \Theta_0$, are continuous at $\Theta = \Theta_0$ and that at least one of the partial derivatives is non-zero at $\Theta = \Theta_0$.

From Lemma A.7.1, it is clear that the all the partial derivatives exist and are continuous at $\Theta = \Theta_0$.

It remains to find at least one partial derivative for each of the component functions g_i which is non-zero at $\Theta = \Theta_0$.

For g_1, $\dfrac{\partial g_1}{\partial \mu} = 1$ is non-zero.

For g_2, $\dfrac{\partial g_2}{\partial \phi_1} = 1$ is non-zero.

For g_3, if $\kappa_1 = 0$, then

$$\frac{\partial g_3}{\partial \phi_1} = \sigma_w^2 \text{ and } \frac{\partial g_3}{\partial \eta_1} = \sigma_v^2 \tag{A.7.34}$$

which by assumption are both non-zero. If $\kappa_1 \neq 0$, then

$$\frac{\partial g_3}{\partial \phi_1} = 0 \Longleftrightarrow \phi_1 = -\frac{\kappa_0}{2\kappa_1} \tag{A.7.35}$$

and

$$\frac{\partial g_3}{\partial \eta_1} = 0 \Longleftrightarrow \eta_1 = \frac{\kappa_0}{2\kappa_1}. \tag{A.7.36}$$

It follows that both $\dfrac{\partial g_3}{\partial \phi_1}$ and $\dfrac{\partial g_3}{\partial \eta_1}$ can both be zero if and only if $\phi_1 = -\eta_1$, which we do not allow.

For g_4, if $\kappa_1 = 0$, then

$$\frac{\partial g_4}{\partial \sigma_v^2} = 1 + \eta_1^2 \tag{A.7.37}$$

which is non-zero. If $\kappa_1 \neq 0$, then

$$\frac{\partial g_4}{\partial \phi_1} = 0 \Longleftrightarrow \phi_1 = \psi_1 \tag{A.7.38}$$

and

$$\frac{\partial g_4}{\partial \eta_1} = 0 \Longleftrightarrow \eta_1 = -\psi_1. \tag{A.7.39}$$

It follows that both $\frac{\partial g_4}{\partial \phi_1}$ and $\frac{\partial g_4}{\partial \eta_1}$ can both be zero if and only if $\phi_1 = -\eta_1$, which we do not allow.

This completes the proof. ■

Appendix B

Stochastic convergence

B.1 Modes of convergence

In this section is defined the various modes of convergence which are referred to throughout this thesis.

Definition B.1.1 ***(Convergence in Probability).*** *A sequence of random variables* $\{X_n\}$ *is said to converge in probability to* X *if*

$$\lim_{n\to\infty} P\left(|X_n - X| < \varepsilon\right) = 1$$

for all $\varepsilon > 0$*. If this is the case, we write* $X_n \xrightarrow{p} X$*.*

Definition B.1.2 ***(Convergence with Probability 1).*** *A sequence of random variables* $\{X_n\}$ *is said to converge with probability 1 to* X *if*

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1.$$

If this is the case, we write $X_n \xrightarrow{wp1} X$*.*

Definition B.1.3 ***(Convergence in Distribution).*** *A sequence of random variables* $\{X_n\}$ *is said to converge in distribution to* X *if*

$$\lim_{n\to\infty} P\left(X_n \leq x\right) = P\left(X \leq x\right)$$

for each point x *such that* $P\left(X = x\right) = 0$*. If this is the case, we write* $X_n \xrightarrow{d} X$*.*

As shown in Billingsley (1995) convergence with probability 1 implies convergence in probability which in turn implies convergence in distribution. The reverse of these implications does not necessarily hold. For more information on these and other modes of convergence see Billingsley (1995) and Serfling (1980).

Frequently a sequence of random variables which converges in distribution, converges to the normal distribution, in which case it is said to be asymptotically normal.

Definition B.1.4 ***(Asymptotically Normal).*** *A sequence of random variables* $\{X_n\}$ *is said to be asymptotically normal with mean* μ_n *and variance* $\sigma_n^2 > 0$ *if*

$$\frac{X_n - \mu_n}{\sigma_n} \xrightarrow{d} N(0,1)$$

where $N(0,1)$ *is the standard normal distribution and is denoted as such by writing* $X_n \sim AN\left(\mu_n, \sigma_n^2\right)$.

The various modes of covergence are used to categorise the manner in which a sequence of estimators converges to the object being estimated.

Definition B.1.5 ***(Weak Consistency).*** *A sequence of estimators* $\{T_n\}$ *is said to be weakly consistent with a parametric function* $g(\theta)$ *if* T_n *converges in probability to* $g(\theta)$, *i.e.* $T_n \xrightarrow{p} g(\theta)$.

Definition B.1.6 ***(Strong Consistency).*** *A sequence of estimators* $\{T_n\}$ *is said to be strongly consistent with a parametric function* $g(\theta)$ *if* T_n *converges with probability* 1 *to* $g(\theta)$, *i.e.* $T_n \overset{wp1}{\to} g(\theta)$.

Clearly any strongly consistent sequence of estimators is also weakly consistent.

B.2 Convergence theorems

The following theorem from Serfling (1980), using the Cramer-Wold device, seen also in Billingsley (1995) (Theorem 29.4) and elsewhere, allows one to derive the asymptotic distribution of a linear combination of random variables from their multivariate asymptotic distribution.

Theorem B.2.1 ***(Asymptotic Distribution of a Linear Combination of Random Variables - Serfling (1980), Theorem 1.5.2).*** *In $\mathbb{R}^k$, the random vectors X_n converge in distribution to the random vector X if and only if each linear combination of the components of X_n converges in distribution to the same linear combination of the components of X.*

The following theorem can be used to show that certain functions of an asymptotically normal random variable are also asymptotically normal.

Theorem B.2.2 ***(Univariate Delta Theorem - Serfling (1980), Theorem 3.1A).*** *Suppose that X_n is $AN\left(\mu, \sigma_n^2\right)$ with $\sigma_n \to 0$. Let g be a real-valued function differentiable at $x = \mu$, with $g'(\mu) \neq 0$. Then*

$$g(X_n) \text{ is } AN\left(g(\mu), \left[g'(\mu)\right]^2 \sigma_n^2\right). \tag{B.2.1}$$

Extensions to Theorem B.2.2 are available to deal with the situation where $g'(\mu) = 0$, see for example Serfling (1980) (Theorem 3.1B). However, we do not consider that situation here. The following theorem deals with vector-valued functions g.

Theorem B.2.3 ***(Multivariate Delta Theorem - Serfling (1980), Theorem 3.3A).*** *Suppose that $\mathbf{X}_n = (X_{n1}, \ldots, X_{nk})$ is $AN\left(\boldsymbol{\mu}, b_n^2 \boldsymbol{\Sigma}\right)$, with $\boldsymbol{\Sigma}$ a covariance matrix and $b_n \to 0$. Let $g(\mathbf{x}) = (g_1(\mathbf{x}), \ldots, g_m(\mathbf{x}))$, where $\mathbf{x} = (x_1, \ldots, x_k)$, be a vector-valued function for which each component function $g_i(\mathbf{x})$ is real-valued and has a non-zero differential $g_i(\mathbf{x}; \mathbf{t})$, where $\mathbf{t} = (t_1, \ldots, t_k)$, at $\mathbf{x} = \boldsymbol{\mu}$, with $g'(\mu) \neq 0$. Put*

$$\mathbf{D} = \left[\frac{\partial g_i}{\partial x_j} \mid_{\mathbf{x}=\boldsymbol{\mu}}\right]_{m \times k} \tag{B.2.2}$$

Then

$$g(\mathbf{X}_n) \text{ is } AN\left(g(\boldsymbol{\mu}), b_n^2 \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}'\right). \tag{B.2.3}$$

Remark B.2.1 ***(Serfling (1980), Remark 3.3B(i)).*** *A sufficient condition for the component function g_i to have a non-zero differential at $\mathbf{x} = \boldsymbol{\mu}$ is that the partial derivatives $\frac{\partial g_i}{\partial x_j}$, for $j = 1, \ldots, k$ are all continuous at $\mathbf{x} = \boldsymbol{\mu}$ and not all zero at $\mathbf{x} = \boldsymbol{\mu}$.*

Appendix C

Time series

C.1 Overview

A time series is a process indexed by some indicator of time, $\mathcal{T}$. The process is defined by a model equation of the form

$$X_t = f_t\left(\{X_t : t \in \mathcal{T}\}, \{u_t : t \in \mathcal{T}\}\right) \tag{C.1.1}$$

where each X_t is an element of some space X, f_t is the time series model function and u_t is an innovation process used to perturb the time series. The time series may also include some function of exogenous data, which we do not consider at this stage. A model such as that specified in equation (C.1.1) is too broad to be of much value, so to allow the development of a useful theory for the estimation and prediction, some restrictions on the nature of f_t, u_t and $\mathcal{T}$ are required.

In general, and throughout this thesis, $\mathcal{T}$ is restricted to the set of integers $\mathbb{Z}$. The innovation process may be restricted to being a white noise or strong white noise, which we define below, though conditional heteroskedastic models such as ARCH, GARCH etc have had a huge impact on the study of financial market data (Gourieroux (1997), Tsay (2010)).

Definition C.1.1 ***(White Noise).*** *The process $\{u_t\}$ is said to be white noise with mean 0 and variance σ^2 if and only if $E[u_t] = 0$ for all t*

$$E[u_t u_{t-h}] = \begin{cases} \sigma^2 < \infty & \text{if } h = 0 \\ 0 & \text{if } h \neq 0 \end{cases} \tag{C.1.2}$$

for all t. If in addition to the above conditions, u_t is an independent and identically distributed, then u_t is said to be strong white noise.

The more important and interesting restrictions are those placed on the function f_t. A time series process is said to be causal if the function f_t takes as parameters only those elements in the data set from the past. Time invariant functions $f := f_t$ are more commonly employed than time variant functions. However, there is a good deal of literature about processes with time varying coefficients, that is models where the structure of the function is time invariant but the process coefficients may change. See Lütkepohl (2005) and Kitigawa (2010) for more information on, in particular time-varying coefficient autoregressive time series processes.

Another important restriction on the function f_t is that of linearity. The class of linear time series processes known as autoregressive moving average (ARMA) processes and the subclasses of autoregressive (AR) and moving average (MA) processes provide the most widely studied and used processes in time series analysis. The ARMA equation,

$$X_t = \mu + \sum_{j=1}^{p} \phi_j \left(X_{t-j} - \mu\right) + \sum_{j=0}^{q} \theta_j u_{t-j} \tag{C.1.3}$$

includes the AR terms $\phi_j \left(X_{t-j} - \mu\right)$ and MA terms $\theta_j u_{t-j}$. An AR process is simply an ARMA process where $q = 0$.Simlarly an MA process is an ARMA process where $p = 0$. Most standard time series textbooks contains a substantial section on ARMA processes (Brockwell and Davis (1991), Hamilton (1994), and Fuller (1996)). To denote that $\{X_t\}$ is an ARMA(p, q) process, we may write

$$X_t \sim \text{ARMA}\left(p, q\right). \tag{C.1.4}$$

Similar notation may be used for MA(q) and AR(p) processes.

The bilinear process described by the following equation

$$X_t = \mu + \sum_{j=1}^{p} \phi_j \left(X_{t-j} - \mu\right) + \sum_{k=0}^{q} \theta_k u_{t-k} + \sum_{j=1}^{r} \sum_{k=0}^{s} \psi_{jk} \left(X_{t-j} - \mu\right) u_{t-k} \tag{C.1.5}$$

is another class of time series processesls which has been extensively studied. Bilinear models offer the ability to model data sets which include sudden bursts away from the mean, such as might be seen in seismic data. For more information on bilinear time series processes, see Subba Rao and Gabr (1984). Other classes of nonlinear time series processes are discussed in Fan and Yao (2005).

Although in this thesis we are primarily concerned with univariate ARMA processes, interesting extensions to much of this thesis may be possible onto both the finite dimensional vector space $\mathbb{R}^K$ and the infinite dimensional Hilbert space $L_2[0,1]$.

We complete this introduction by stating the following definitions which are frequently used in time series analysis.

Definition C.1.2 ***(Stationarity).*** *The time series $\{X_t, t \in \mathbb{Z}\}$ is said to be stationary, or second order stationary, if*

(a) $E\left[X_t^2\right] < \infty$ *for all* $t \in \mathbb{Z}$,

(b) $E[X_t] = m$ *for all* $t \in \mathbb{Z}$ *and*

(c) $Cov(X_r, X_s) = Cov(X_{r+t}, X_{s+t})$ *for all* $r, s, t \in \mathbb{Z}$.

Definition C.1.3 ***(Strict Stationarity).*** *The time series $\{X_t, t \in \mathbb{Z}\}$ is said to be strictly stationary if the joint distributions of $(X_{t_1}, \ldots, X_{t_k})$ and $(X_{t_1+h}, \ldots, X_{t_k+h})$ are the same for all positive integers k and for all $t_1, \ldots, t_k, h \in \mathbb{Z}$.*

The autocovariance and autocorrelation functions are used to describe the dependency of time series processes.

Definition C.1.4 ***(Autocovariance Function).*** *If $\{X_t, t \in \mathbb{Z}\}$ is a process such that $Var(X_t) < \infty$ for each $t \in \mathbb{Z}$,then the autocovariance function $\lambda(\cdot,\cdot)$ of $\{X_t\}$ is defined by*

$$\lambda(r,s) = Cov[X_r, X_s] \tag{C.1.6}$$

for $r, s \in \mathbb{Z}$. If $\{X_t\}$ is stationary, then for convenience we can redefine the autocovariance function to be

$$\lambda(h) = Cov[X_t, X_{t+h}] \tag{C.1.7}$$

Definition C.1.5 ***(Autocorrelation Function).*** *If $\{X_t, t \in \mathbb{Z}\}$ is a process such that $Var(X_t) < \infty$ for each $t \in \mathbb{Z}$,then the autocorrelation function $\rho(\cdot,\cdot)$ of $\{X_t\}$ is defined by*

$$\rho(r,s) = \frac{\lambda(r,s)}{\lambda(r,r)} \tag{C.1.8}$$

for $r, s \in \mathbb{Z}$. If $\{X_t\}$ is stationary, then for convenience we can redefine the autocorrelation function to be

$$\rho(h) = \frac{\lambda(h)}{\lambda(0)} \tag{C.1.9}$$

The lag operator is commonly used in time series literature for ease of notation.

Definition C.1.6 ***(Lag Operator).*** *The lag operator L on an element X_t of a time series process is defined by*

$$L\left(X_t\right) = X_{t-1}. \tag{C.1.10}$$

The autoregressive and moving average polynomials of a play an important role in the analysis of ARMA *processes.*

Definition C.1.7 ***(Autoregressive Polynomial).*** *The autoregressive polynomial of the* ARMA(p, q) *process in* $(C.1.3)$ *is defined to be*

$$\phi\left(z\right) = 1 - \phi_1 z - \cdots - \phi_p z^p. \tag{C.1.11}$$

Definition C.1.8 ***(Moving Average Polynomial).*** *The moving average polynomial of the* ARMA(p, q) *process in* $(C.1.3)$ *is defined to be*

$$\theta\left(z\right) = 1 + \theta_1 z + \cdots + \theta_q z^q. \tag{C.1.12}$$

The ARMA process equation $(C.1.3)$ can be succinctly written using the autoregressive and moving average polynomials and the lag operator as

$$\phi\left(L\right) X_t = \theta\left(L\right) u_t. \tag{C.1.13}$$

C.2 Summation of autoregressive moving average processes

In this section we describe some properties of the sums of ARMA processes. The following theorem shows that ARMA processes are closed under addition and provides upper bounds on the order of the summed process.

Theorem C.2.1 ***(Granger and Morris (1976) - Section 2).*** *Suppose $X_t \sim$ ARMA(k_1, m_1) and $Y_t \sim$ ARMA(k_2, m_2) are indepedendent, zero-mean, stationary processes and define the process $\{Z_t\}$ by*

$$Z_t = X_t + Y_t. \tag{C.2.1}$$

Then

$$Z_t \sim \text{ARMA}\left(k, m\right) \tag{C.2.2}$$

where

$$k \leq k_1 + k_2 \text{ and } m \leq \max\left(k_1 + m_2, k_2 + m_1\right). \tag{C.2.3}$$

The proof of Theorem C.2.1 involves the use of the autoregressive and moving average polynomials of $\{X_t\}$ and $\{Y_t\}$ to construct formulae for the autoregressive and moving average polynomials of $\{Z_t\}$. Suppose

$$\phi_1(L) X_t = \theta_1(L) e_t \tag{C.2.4}$$

and

$$\phi_2(L) Y_t = \theta_2(L) u_t \tag{C.2.5}$$

then

$$Z_t = \phi_1^{-1}(L) \theta_1(L) e_t + \phi_2^{-1}(L) \theta_2(L) u_t \tag{C.2.6}$$

or

$$\begin{aligned} \phi(L) Z_t &= \phi_1(L) \phi_2(L) Z_t \\ &= \phi_2(L) \theta_1(L) e_t + \phi_1(L) \theta_2(L) u_t \\ &= \theta(L) v_t \end{aligned} \tag{C.2.7}$$

The process $\{\theta(L) v_t\}$ is the sum of an $\text{MA}(k_1 + m_2)$ process and an $\text{MA}(k_2 + m_1)$ process which can be shown to be an $\text{MA}(m)$ process where m is given by $(C.2.3)$. The autoregressive polynomial, $\phi(L)$, of $\{Z_t\}$ is product $\phi_1(L) \phi_2(L)$ which has order $k \leq k_1 + k_2$. The inequalities in $(C.2.3)$ allow for the possibility that there may be common roots in the polynomials of $\{X_t\}$ and $\{Y_t\}$.

Whilst calculation of $\phi(L)$ is straightforward, an analytic formula for the coefficients of $\theta(L)$ may not be available. It is possible to calculate the autocovariance function of $\{\theta(L) v_t\}$, even where there is a correlation between $\{e_t\}$ and $\{u_t\}$. Given the autocovariance function of a moving average process, it is possible to calculate the unique moving average polynomial of the process using a Newton-Raphson method proposed in Wilson (1969), known as the Wilson Factorisation Algorithm.

C.3 Self-weighted least absolute deviation estimation

In this section we describe the self-weighted least absolute deviation (SLAD) method for the estimation of the parameters of an $\text{SMA}(q)$ process. Although the SLAD method is used in

this thesis only for the estimation of SMA(q) processes, the SLAD method is applicable to more general infinite variance ARMA processes. Let $\{X_t\}$ be the following SMA(q) process

$$X_t = \mu + \sum_{j=1}^{q} \theta_j e_{t-j} + e_t \tag{C.3.1}$$

where $\{e_t\}$ is an IID sequence of stable random variables such that

$$e_t \sim S_\alpha^0 \left(\beta_e, \gamma_e, \delta_e\right). \tag{C.3.2}$$

Denote the vector of process parameters by

$$\psi = \left(\mu, \theta_1, \ldots, \theta_q\right)'. \tag{C.3.3}$$

For a set of observations $\{x_1, \ldots, x_n\}$ and a set of initial values $\{x_0, x_{-1}, \ldots\}$, we define the objective function

$$H_n\left(\psi\right) = \frac{1}{n} \sum_{t=1}^{n} w_t \left|e_t\left(\psi\right)\right| \tag{C.3.4}$$

where

$$e_t\left(\psi\right) = x_t - \mu - \sum_{j=1}^{q} \theta_j e_{t-j}\left(\psi\right) \tag{C.3.5}$$

and

$$w_t \equiv w\left(x_{t-1}, x_{t-2}, \ldots\right) \tag{C.3.6}$$

is a weight function. If the initial values $\{x_0, x_{-1}, \ldots\}$ are not available we set them to zero. The SLAD estimator $\widehat{\psi}_n$ is defined as the minimiser of the objective function over the parameter space $\Psi \subset \mathbb{R}^k$, where $k = q + 1$, i.e.

$$\widehat{\psi}_n = \arg\min_{\psi \in \Psi} H_n\left(\psi\right). \tag{C.3.7}$$

Asymptotic properties of the SLAD estimators are given in the following theorem.

Theorem C.3.1 ***(Zhu and Ling (2012) - Theorem 2).*** *Let ψ_0 denote the true value of ψ for the* SMA(q) *process in* (C.3.1). *If the following assumptions are satisfied:*

AC.3.1 *That ψ_0 is an interior point in a compact parameter space Ψ;*

AC.3.2 *For each $\psi \in \Psi$, the moving average polynomials have no zero roots in the unit circle;*

AC.3.3 *That the weight function w_t is a positive, measurable and bounded function and that $E\left[\left(w_t + w_t^2\right)\pi_{\rho,t-1}^2\right] < \infty$ for any $\rho \in (0,1)$, where*

$$\pi_{\rho,t-1} = 1 + \sum_{i=0}^{\infty} \rho^i \left|X_{t-1-i}\right| \quad a.s.; \tag{C.3.8}$$

AC.3.4 *That ε_t has a zero median with a continuous density function $g(x)$ satisfying $g(0) > 0$ and $\sup_{x\in\mathbb{R}} g(x) < \infty$.*

then

1. $\sqrt{n}\left(\widehat{\psi}_n - \psi_0\right) = O_p(1)$
2. $\sqrt{n}\left(\widehat{\psi}_n - \psi_0\right) \xrightarrow{d} N(0, V_\psi) \quad as\ n \to \infty$

where

$$V_\psi = [2g(0)]^{-2}\,\Sigma_0^{-1}\Omega_0\Sigma_0^{-1} \tag{C.3.9}$$

and

$$\Sigma_0 = E\left[w_t\left(\partial e_t(\psi_0)/\partial\psi\right)\left(\partial e_t(\psi_0)/\partial\psi'\right)\right] \tag{C.3.10}$$

$$\Omega_0 = E\left[w_t^2\left(\partial e_t(\psi_0)/\partial\psi\right)\left(\partial e_t(\psi_0)/\partial\psi'\right)\right] \tag{C.3.11}$$

We denote the components of V_ψ by

$$V_\psi = \begin{pmatrix} V_\mu & V_{\mu,\theta} \\ V_{\mu,\theta}' & V_\theta \end{pmatrix} \tag{C.3.12}$$

where $V_\mu, V_{\mu,\theta}$ and V_θ are $1\times 1, 1\times q$ and $q\times q$ matrices respectively.

Assumptions AC.3.1 and AC.3.2 are satisfied by the collection of all invertible SMA(q) processes. If e_t has a non-zero median, then Assumption AC.3.4 can be satisfied by applying an appropriate location shift to the sequence $\{e_t\}$ and the parameter μ. All stable distributions have a non-zero density at the median.

There are many choices available for the weight function w_t which satisfy Assumption AC.3.3. The following weight function is suggested in Zhu and Ling (2012). If $E\left[e_t^2\right] < \infty$, then

$$w_t = 1. \tag{C.3.13}$$

If $E\left[e_t^2\right] = \infty$ and $E\left[|e_t|\right] < \infty$, then

$$w_t = \left[\max\left(1, C^{-1}\sum_{j=1}^{\infty}\frac{1}{j^9}\left|X_{t-j}\right| I\left\{\left|X_{t-j}\right| > C\right\}\right)\right]^{-4} \tag{C.3.14}$$

where $C > 0$ is a constant. If $E\left[|e_t|\right] = \infty$, then

$$w_t = \left[\max\left(1, C^{-1}\sum_{j=1}^{\infty}\frac{1}{j^{1+16/\alpha_1}}\left|X_{t-j}\right| I\left\{\left|X_{t-j}\right| > C\right\}\right)\right]^{-4} \tag{C.3.15}$$

where $0 < \alpha_1 < \alpha$.

It is claimed in Zhu and Ling (2012), that choosing C to be the $0.90th$ quantile of X_t works well in practice. However, in running the simulations for this thesis it was found that choosing C to be the $0.99th$ quantile of $|X_t|$ produced estimates with a lower variance and a distribution which more closely resembled the Gaussian distribution. Using a quantile of $|X_t|$ rather than X_t produced better estimates of an SMA(q) process with a skewed distribution. In the remainder of this thesis, unless stated otherwise, the weight functions used by the SLAD estimator are those defined in $(C.3.13), (C.3.14)$ and $(C.3.15)$ and are used with C equal to the $0.99th$ empirical quantile estimate of $\{|X_t|\}$ and $\alpha_1 = 2\alpha/3$.

In Zhu and Ling (2012), it is suggested that an estimate for the asymptotic variance V_ψ can be obtained by using

$$\widehat{V}_\psi = \left[2\widehat{g}\left(0\right)\right]^{-2}\widehat{\Sigma}_n^{-1}\widehat{\Omega}_n\widehat{\Sigma}_n^{-1} \tag{C.3.16}$$

where

$$\widehat{\Sigma}_n = \frac{1}{n}\sum_{t=1}^{n} w_t \frac{\partial e_t\left(\widehat{\psi}_n\right)}{\partial\psi}\frac{\partial e_t\left(\widehat{\psi}_n\right)}{\partial\psi'} \tag{C.3.17}$$

$$\widehat{\Omega}_n = \frac{1}{n}\sum_{t=1}^{n} w_t^2 \frac{\partial e_t\left(\widehat{\psi}_n\right)}{\partial\psi}\frac{\partial e_t\left(\widehat{\psi}_n\right)}{\partial\psi'} \tag{C.3.18}$$

and $\widehat{g}\left(0\right)$ is a density estimate of e_t at 0.

C.4 Generalised method of moments estimation

In this section we describe how the generalised method of moments (GMM) can be used to estimate the parameters of a stochastic volatility (SV) model. The GMM is an extension of the classical method of moments (CMM) as an alternative to maximum likelihood estimation.

The advantage of GMM estimation is that it does not require knowledge of the complete density function, only knowledge of selected moment functions. A moment function is just a function of the moments.

In CMM the number of moment functions used equals the number of parameters to be estimated and there is a unique solution. An example of CMM is the estimation of the parameters of an AR(1) process. Let $\{X_t\}$ be the AR(1) process given by

$$X_t = \phi_1 X_{t-1} + e_t \tag{C.4.1}$$

where $\{e_t\}$ is an IID sequence of $N\left(0, \sigma_e^2\right)$ random variables. The parameters to be estimated are ϕ_1 and σ_e^2. In CMM, we might use the autocovariance function, λ_h, at lags $h = 0, 1$ as our moment functions. Since

$$\lambda_0 = \frac{\sigma_e^2}{1-\phi_1^2} \text{ and } \lambda_1 = \phi_1 \frac{\sigma_e^2}{1-\phi_1^2}, \tag{C.4.2}$$

substitution of the sample autocovariance functions λ_0 and λ_1 into $(C.4.2)$ produces a unique set of estimates for the parameters ϕ_1 and σ_e^2. This example is also known as Yule-Walker estimation. Similar results could be achieved using the autocovariance function at any distinct pair of lags. There may be some benefit in including information from the sample autocovariance functions at many lags. However, in general there is not a single set of parameters which exactly matches the sample autocovariance function at more than two distinct lags.

In GMM the number of moment functions used may be greater than the number of parameters to be estimated. Instead of searching for a set of parameters which exactly matches each of the sample moment functions, the GMM returns the set of parameters which best matches all of the sample moment functions. Let $g(\theta) = (g_1(\theta), \ldots, g_k(\theta))'$ denote the vector of moment functions used in a GMM estimator of the vector of parameters $\theta = (\theta_1, \ldots, \theta_m)'$ where $\theta \in \Theta$ and Θ is some parameter space. The GMM estimator, $\widehat{\theta}$, from a sample of size n is then defined by

$$\widehat{\theta} = \arg\min_{\theta \in \Theta} \left(n^{-1} \sum_{t=1}^{n} \widehat{g}(t) - g(\theta) \right) W \left(n^{-1} \sum_{t=1}^{n} \widehat{g}(t) - g(\theta) \right)' \tag{C.4.3}$$

where $\widehat{g}(t) = (\widehat{g}_1(t), \ldots, \widehat{g}_k(t))'$ is the vector of sample moment functions evaluated at t and W is a symmetric weight matrix which assigns an importance to each moment function

of the GMM estimator. Hence, in the design of a GMM estimator there are two important decisions to be made: i) which moment functions to use and ii) which weight matrix to use.

The particular estimation problem of interest to this thesis to which we wish to apply the GMM is the estimation of the parameters of an SV model. In our SV model, the observable process $\{X_t\}$ is given by

$$X_t = \sigma_t \varepsilon_t, \tag{C.4.4}$$

$$\ln \sigma_t = \mu + \phi_1 \ln \sigma_{t-1} + v_t, \tag{C.4.5}$$

where $\{\varepsilon_t\}$ is an IID sequence of $N(0,1)$ random variables and $\{v_t\}$ is an IID sequence of $N\left(0, \sigma_v^2\right)$ random variables which is independent of $\{\varepsilon_t\}$ .The small-sample performance of a GMM estimator for the SV model was explored in depth by Anderson and Sorensen (1996) and we follow their design in this thesis. The 24 moment functions considered by Anderson and Sorensen (1996) are as follows:

$$g_1(\theta) = E\left[|X_t|\right], \tag{C.4.6}$$

$$g_2(\theta) = E\left[X_t^2\right], \tag{C.4.7}$$

$$g_3(\theta) = E\left[|X_t|^3\right], \tag{C.4.8}$$

$$g_4(\theta) = E\left[X_t^4\right], \tag{C.4.9}$$

$$g_{4+i}(\theta) = E\left[|X_t X_{t-i}|\right], \quad i = 1, \ldots, 10, \tag{C.4.10}$$

$$g_{14+i}(\theta) = E\left[X_t^2 X_{t-i}^2\right], \quad i = 1, \ldots, 10. \tag{C.4.11}$$

where $\theta = \left(\mu, \phi_1, \sigma_v^2\right)$. Formulae for each of the $g_i(\theta)$ can be found in Section 1 of Anderson and Sorensen (1996) and in the appendix of Jacquier et al. (1994).

The asymptotic optimal choice of the weight matrix, W^*, is the inverse of the covariance matrix of the moment functions, Hansen (1982), i.e.

$$(W^*)^{-1} = \left(E\left[g_i(\theta)\, g_j(\theta)\right]\right)_{i,j=1,\ldots,k}. \tag{C.4.12}$$

The optimal weight matrix is not known unless the true parameters are known, but can be estimated through a kernel method proposed by Newey and West (1987). For some chosen lag h, the estimate $\left(\widehat{W}^*(\theta)\right)^{-1}$ of $(W^*)^{-1}$ is given by

$$\left(\widehat{W}^*(\theta)\right)^{-1} = n^{-1} \sum_{i=0}^{h} \left(1 - \frac{i}{h+1}\right) \sum_{t=i}^{n} \left(\widehat{g}(t) - g(\theta)\right) \left(\widehat{g}(t-i) - g(\theta)\right)'. \tag{C.4.13}$$

An initial estimate of θ is obtained using an estimate for the weight matrix where the mean of $\widehat{g}(t)$ is substituted for $g(\theta)$ in $(C.4.13)$. A final estimate of θ is obtained using an estimate for the weight matrix obtained by substituting the initial estimate of θ into $(C.4.13)$.

In Hansen (1982) it is shown that asymptotic distribution of $\widehat{\theta}$ is given by

$$\sqrt{n}\left(\widehat{\theta}-\theta_0\right) \xrightarrow{d} D'\left(W^*\right)^{-1} D \tag{C.4.14}$$

where

$$D = \frac{\partial g(\theta)}{\partial \theta} \mid_{\theta=\theta_0} \tag{C.4.15}$$

and θ_0 is the true value of the parameters. The inclusion of additional moment functions tends to decrease the asymptotic variance of the GMM estimator, (Table 4, Anderson and Sorensen (1996)). However, for small samples the use of fewer moment functions can be optimal, (Tables 1 and 2, Anderson and Sorensen (1996)).

Further general information on GMM estimators can be found in Chapter 14 of Hamilton (1994). A comparison of GMM estimation and quasi maximum likelihood estimation of the SV model parameters can be found in Anderson and Sorensen (1997).

Appendix D

Quantile estimation

D.1 Introduction

There is an extensive literature on the statistical properties of the empirical quantile estimators, see for example Cramer (1946), Wilks (1962), Durbin (1973) and Serfling (1980). Other quantile estimators, such as those which involve some form of smoothing, e.g. histograms or kernels (Silverman (1986)), are not considered in this thesis.

The following theorems assume that $\{x_j\}_{j=1}^n$ is an independent sample drawn from F. The first theorem shows that the empirical quantile estimator has strong consistency wherever the underlying distribution function is not flat.

Theorem D.1.1 ***(Strong Consistency of $\widehat{\xi}_{n;p}$ - Serfling (1980), Theorem 2.3.1).*** *Let $0 < p < 1$. If ξ_p is the unique solution x of $F(x_-) \leq p \leq F(x)$, then $\widehat{\xi}_{n;p}$ is a strongly consistent estimator of ξ_p.*

The next theorem shows that the empirical quantile estimator is asymptotically normal under some conditions on the underlying distribution function (See also Cramer (1946))

Theorem D.1.2 ***(Asymptotic Normality of Empirical Quantile Estimator - Serfling (1980), Corollary 2.3.3B).*** *For $0 < p < 1$, if F possesses a density f in a neighbourhood of ξ_p and if f is positive and continuous at ξ_p, then*

$$\widehat{\xi}_{n;p} \text{ is } AN\left(\xi_p, \frac{p(1-p)}{f^2(\xi_p)\, n}\right) \tag{D.1.1}$$

or expressed in terms of the quantile density function q

$$\widehat{\xi}_{n;p} \text{ is } AN\left(\xi_p, \frac{p\left(1-p\right)q^2\left(p\right)}{n}\right). \tag{D.1.2}$$

Theorem D.1.2 is extended in the following theorem to cover the estimation of multiple quantiles from a single sample.

Theorem D.1.3 ***(Asymptotic Covariances of Empirical Quantile Estimators - Serfling (1980), Theorem 2.3.3B).*** *Let* $0 < p_1 < \cdots < p_k < 1$.*Suppose that F has a density f in a neighbourhoods of* $\xi_{p_1}, \ldots, \xi_{p_k}$ *and that f is positive and continuous at* $\xi_{p_1}, \ldots, \xi_{p_k}$. *Let* $\widehat{\xi}_n = \left(\widehat{\xi}_{n;p_1}, \ldots, \widehat{\xi}_{n;p_k}\right)'$ *denote the empirical quantiles estimates of* $\xi = \left(\xi_{p_1}, \ldots, \xi_{p_k}\right)'$, *then*

$$\sqrt{n}\left(\widehat{\xi}_n - \xi\right) \xrightarrow{d} N\left(0, V_\xi\right). \tag{D.1.3}$$

The element in the ith row and jth column of V_ξ *is given by*

$$v_{ij} = \begin{cases} \dfrac{p_i\left(1-p_j\right)}{f\left(\xi_{p_i}\right) f\left(\xi_{p_j}\right)} & \text{for } i \leq j \\ \dfrac{p_j\left(1-p_i\right)}{f\left(\xi_{p_i}\right) f\left(\xi_{p_j}\right)} & \text{for } i > j \end{cases}. \tag{D.1.4}$$

Appendix E

Stable distributions

E.1 Introduction

The following definition for stable distributions is from Feller (1970).

Definition E.1.1 ***(Stable Distribution).*** *Let $X, X_1, X_2, \ldots$ be independent random variables with a common distribution F and define S_n by*

$$S_n = X_1 + \cdots + X_n. \tag{E.1.1}$$

The distribution F is said to be stable (in the broad sense) if for each n there exist constants $a_n > 0$ and b_n such that

$$S_n \sim a_n X + b_n \tag{E.1.2}$$

and F is not concentrated at a single point. The distribution F is said to be stable (in the strict sense) if $(E.1.2)$ holds with $b_n = 0$.

Stable distributions can also be defined by their characteristic functions. Several different parameterisations for stable characteristic functions have been proposed, see Zolotarev (1986). In this thesis we use the $S^0_\alpha(\beta, \gamma, \delta)$ parameterisation recommended by Nolan (1998), where $\alpha \in (0, 2]$ is the index of stability, $\beta \in [-1, 1]$ is the skewness parameter, $\gamma > 0$ is the scale parameter and δ is the location parameter. If $X \sim S^0_\alpha(\beta, \gamma, \delta)$, then the characteristic

function of X is given by

$$E\left[\exp\left(iuX\right)\right] = \begin{cases} \exp\left\{ \begin{array}{c} -\gamma^{\alpha}\left|u\right|^{\alpha}\left[1-i\beta\left(\tan\frac{\pi\alpha}{2}\right)\operatorname{sign}\left(u\right)\right]+ \\ i\left[\delta-\beta\left(\tan\frac{\pi\alpha}{2}\right)\gamma\right]u \end{array} \right\} & \text{if } \alpha\neq 1 \\ \exp\left\{ \begin{array}{c} -\gamma\left|u\right|\left[1+i\beta\frac{2}{\pi}\operatorname{sign}\left(u\right)\ln\left|u\right|\right]+ \\ i\left[\delta-\beta\frac{2}{\pi}\gamma\ln\gamma\right]u \end{array} \right\} & \text{if } \alpha=1 \end{cases}. \quad \text{(E.1.3)}$$

Note that the $S_{\alpha}^{0}\left(\beta,\gamma,\delta\right)$ parameterisation in Nolan (1998) is the same as the $S\left(\alpha,\beta,\gamma,\delta;0\right)$ in Nolan (2015).

Unfortunately, closed form expressions for a stable density function exist only in the special cases of the Cauchy distribution ($\alpha=1$) and the normal distribution ($\alpha=2$). The S_{α}^{0} parameterisation of a normal distribution is given by,

$$N\left(\mu,\sigma^{2}\right)\sim S_{2}^{0}\left(0,\sigma/\sqrt{2},\mu\right). \quad \text{(E.1.4)}$$

Under the S_{α}^{0} parameterisation, stable distributions exhibit the properties described in the following lemma.

Lemma E.1.1 *The S_{α}^{0} parameterisation has the following properties.*

a) *The parameters γ and δ are true scale and location parameters: if $X\sim S_{\alpha}^{0}\left(\beta,\gamma,\delta\right)$, then for any $a\neq 0, b$*

$$aX+b\sim S_{\alpha}^{0}\left(sign\left(a\right)\beta,\left|a\right|\gamma,a\delta+b\right) \quad \text{(E.1.5)}$$

b) *The characteristic function and hence the density and distribution functions are jointly continuous in all four parameters $\left(\alpha,\beta,\gamma,\delta\right)$.*

c) *If $X_1, X_2, \ldots, X_n$ are pairwise independent and $X_j\sim S_{\alpha}^{0}\left(\beta_j,\gamma_j,\delta_j\right)$ are independent, then $\sum_{j=1}^{n}X_j\sim S_{\alpha}^{0}\left(\beta,\gamma,\delta\right)$ where*

$$\gamma^{\alpha} = \sum_{j=1}^{n}\gamma_j^{\alpha} \quad \text{(E.1.6)}$$

$$\beta = \frac{\sum_{j=1}^{n}\beta_j\gamma_j^{\alpha}}{\sum_{j=1}^{n}\gamma_j^{\alpha}} \quad \text{(E.1.7)}$$

$$\delta = \begin{cases} \sum_{j=1}^{n}\delta_j+\left(\tan\frac{\pi\alpha}{2}\right)\left[\beta\gamma-\sum_{j=1}^{n}\beta_j\gamma_j\right] & \text{if } \alpha\neq 1 \\ \sum_{j=1}^{n}\delta_j+\frac{2}{\pi}\left[\beta\gamma\ln\gamma-\sum_{j=1}^{n}\beta_j\gamma_j\ln\gamma_j\right] & \text{if } \alpha=1 \end{cases}. \quad \text{(E.1.8)}$$

Proof. Statements a) and b) are taken from Lemma 1, Nolan (1998). Statement c) is included for $n = 2$ in Lemma 1, Nolan (1998). The extension of c) to general n is a straightforward induction. ■

We make the following remarks on two of the more important consequences of Lemma E.1.1 for this thesis.

Remark E.1.1 *If* $\{X_t\}$ *is a moving average process with stable innovations, then* $\{X_t\}$ *has a stable distribution and we can calculate the distributional parameters of* $\{X_t\}$ *from the distributional parameters of the innovations and the moving average process parameters.*

Remark E.1.2 *If* $\{X_t\}$ *is an iid process with a stable distribution, then its temporal aggregation* $\left\{S_t^{(r)}\right\}$ *has a stable distribution and we can calculate the distributional parameters of* $\left\{S_t^{(r)}\right\}$ *from the distributional parameters of* $\{X_t\}$.

We can use Lemma E.1.1, to establish a simple, yet useful, relation between the quantiles and log quantile differences of stable random variables.

Lemma E.1.2 *Suppose* $X \sim S_\alpha^0(\beta, \gamma, \delta)$ *and* $Y \sim S_\alpha^0(\beta, |a|\gamma, a\delta + b)$ *for some* $a > 0, b$. *Let* $\xi_{X;p}, \xi_{Y;p}$ *denote respectively the pth quantiles of* X *and* Y, *then*

$$\xi_{Y;p} = a\xi_{X;p} + b \tag{E.1.9}$$

Let $\zeta_{X;p_1,p_2}, \zeta_{Y;p_1,p_2}$ *denote respectively the* (p_1, p_2)*th log quantile differences of* X *and* Y, *then*

$$\zeta_{Y;p_1,p_2} = \ln a + \zeta_{X;p_1,p_2} \tag{E.1.10}$$

Proof. From Lemma E.1.1, we have that

$$Y \sim aX + b. \tag{E.1.11}$$

and therefore

$$\begin{aligned} p &= P\left(X \leq \xi_{X;p}\right) \\ &= P\left(Y \leq a\xi_{X;p} + b\right) \end{aligned} \tag{E.1.12}$$

which proves $(E.1.9)$.

From $(E.1.9)$ we get

$$\begin{aligned}\zeta_{Y;p_1,p_2} &= \ln\left(\left|\xi_{Y;p_2}-\xi_{Y;p_1}\right|\right)\\ &= \ln\left(\left|a\xi_{X;p_2}-a\xi_{X;p_1}\right|\right)\\ &= \ln a+\zeta_{X;p_1,p_2}\end{aligned} \tag{E.1.13}$$

which proves $(E.1.10)$. ■

The scale and location properties of the γ and δ parameters lead to the following useful result on the quantiles of the stable distribution.

Lemma E.1.3 *Let*

$$X \sim S_\alpha^0(\beta,\gamma,\delta), \tag{E.1.14}$$

$$X^* \sim S_\alpha^0(\beta,1,0) \tag{E.1.15}$$

be stable random variables. Let ξ_p and ξ_p^* denote respectively the pth quantile of X and X^*. Then for any $0<p_1,p_2<1$ where $p_1\neq p_2$ we have

$$\gamma=\frac{\xi_{p_2}-\xi_{p_1}}{\xi_{p_2}^*-\xi_{p_1}^*} \tag{E.1.16}$$

and

$$\delta=\xi_{p_1}-\gamma\xi_{p_1}^*. \tag{E.1.17}$$

Proof. From Lemma E.1.1 we have

$$\gamma X^*+\delta\sim X. \tag{E.1.18}$$

It follows that for any $0<p<1$,

$$\xi_p^*=\frac{\xi_p-\delta}{\gamma} \tag{E.1.19}$$

from which $(E.1.16)$ and $(E.1.17)$ follow immediately. ■

A final useful property of stable distributions is the reflection property.

Proposition E.1.1 ***(Reflection Property - Nolan (2015), Proposition 1.11).*** *For any α and β, if $X_1\sim S_\alpha^0(\beta,1,\beta\tan\pi\alpha/2)$ and $X_2\sim S_\alpha^0(-\beta,1,\beta\tan\pi\alpha/2)$ then*

$$X_1\overset{d}{=}-X_2. \tag{E.1.20}$$

Bibliography

Adler, R. F., Feldman, R. E., and Gallagher, C. (1998). Analyzing stable time series. In Adler, R. J., Feldman, R. E., and Taqqu, M. S., editors, *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhäuser.

Akgiray, V. and Booth, G. G. (1988). The stable-law model of stock returns. *Journal of Business and Economic Statistics*, 6:51–57.

Amemiya, T. and Wu, R. Y. (1972). The effect of aggregation on prediction in the autoregressive model. *Journal of the American Statistical Association*, 67:628–632.

Anderson, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001a). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61:43–76.

Anderson, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001b). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96:42–55.

Anderson, T. G. and Sorensen, B. E. (1996). GMM estimation of a stochastic volatilty models: A monte carlo study. *Journal of Business and Economic Statistics*, 14:328–352.

Anderson, T. G. and Sorensen, B. E. (1997). GMM and QML asymptotic standard deviations in stochastic volatilty models: Comments on Ruiz (1994). *Journal of Econometrics*, 76:397–403.

Bachelier, L. (1900). Theory of speculation. In Cootner, P., editor, *The random character of stock prices*. MIT Press.

Barker, A. W. (2014). On quantiles of the temporal aggregation of a stable moving average process and their applications. In Nelson, M., Hamilton, T., Jennings, M.,

and Bunder, J., editors, *Proceedings of the 11th Biennial Engineering Mathematics and Applications Conference, EMAC-2013*, volume 55 of *ANZIAM J.*, pages C153–C167. url http://journal.austms.org.au/ojs/index.php/ANZIAMJ/article/view/7826 [May 22, 2014].

Barndorff-Nielsen, O. E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, 24:1–13.

Billingsley, P. (1995). *Probability and Measure*. Wiley, 3rd edition.

Blattberg, R. C. and Gonedes, N. J. (1974). A comparison of the stable and student distributions as statistical models for stock prices. *Journal of Business*, 47:244–280.

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

Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65:1509–1526.

Brewer, K. R. W. (1973). Some consequences of temporal aggregation and systematic sampling for ARMA and ARMAX models. *Journal of Econometrics*, 1:133–154.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. Springer, 2nd edition.

Brorsen, B. W. and Yang, S. R. (1990). Maximum likelihood estimates of symmetric stable distribution parameters. *Communications in Statistics - Simulation and Computation*, 19:1459–1464.

Buonaccorsi, J. P. (2010). *Measurement Error: Models, Methods and Applications*. Chapman and Hall.

Chan, W. S., Cheung, S. H., Zhang, L. X., and Wu, K. H. (2008). Temporal aggregation of equity return time series models. *Mathematics and Computers in Simulation*, 78:172–180.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236.

Cramer, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press.

Dacorogna, M. M., Gencay, R., Muller, U., Olsen, R. B., and Pictet, O. V. (2001). *An Introduction to High-Frequency Finance*. Academic Press.

Davis, R. A. (1996). Gauss-Newton and M-estimation for ARMA processes with infinite variance. *Stochastic Processes and their Applications*, 63:75–95.

Davis, R. A., Knight, K., and Liu, J. (1992). M-estimation for autoregressions with infinite variance. *Stochastic Processes and their Applications*, 40:145–180.

Davis, R. A. and Resnick, S. (1986). Limit theory for the sample covariance and correlation functions of moving averages. *The Annals of Statistics*, 14:533–558.

Dominicy, Y., Hormann, S., Ogata, H., and Veredas, D. (2013). On sample marginal quantiles for stationary processes. *Statistics and Probability Letters*, 83:28–36.

Drost, F. C. and Nijman, T. E. (1993). Temporal aggregation of GARCH processes. *Econometrica*, 61:909–927.

DuMouchel, W. H. (1975). Stable distributions in statistical inference: 2. information from stably distributed samples. *Journal of the American Statistical Association*, 70:386–393.

Durbin, J. (1973). *Distribution Theory for Tests Based on the Sample Distribution Function*. SIAM.

Durbin, J. and Koopman, S. J. (2001). *Time Series Analysis by State Space Methods*. Oxford University Press.

Dutta, K. and Sen, P. K. (1971). On the Bahadur representation of sample quantiles in some stationary multivariate autoregressive processes. *Journal of Multivariate Analysis*, 1:186–198.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United kingdom inflation. *Econometrica*, 50:987–1007.

Engle, R. F. and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, 5:1–50.

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

Fama, E. F. and Roll, R. (1971). Parameter estimates for symmetric stable distributions. *Journal of the American Statistical Association*, 66:331–338.

Fan, J. and Yao, Q. (2005). *Nonlinear Time Series: Nonparametrix and Parametric Methods*. Springer.

Fan, Z. (2006). Parameter estimation of stable distributions. *Communications in Statistics - Theory and Methods*, 35:245–255.

Feller, W. (1970). *An Introduction to Probability Theory and its Applications*, volume II. Wiley, 2nd edition.

Fielitz, B. D. and Rozelle, J. P. (1983). Stable distributions and mixtures of distribution hypotheses for common stock returns. *Journal of the American Statistical Association*, 78:28–36.

Fielitz, B. D. and Smith, E. W. (1972). Asymmetric stable distribution of stock price changes. *Journal of the American Statistical Association*, 67:813–814.

French, K. R., Schwert, G. W., and Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, 19:3–29.

Fuller, W. A. (1996). *Introduction to Statistical Time Series*. Wiley, 2nd edition.

Gander, M. P. S. and Stephens, D. A. (2007). Stochastic volatility modelling in continuous time with general marginal distributions: Inference, prediction and model selection. *Journal of Statistical Panning and Inference*, 137:3068–3081.

Garcia, R., Renault, E., and Veredas, D. (2011). Estimation of stable distributions by indirect inference. *Journal of Econometrics*, 161:325–337.

Gourieroux, C. (1997). *ARCH Models and Financial Applications*. Springer.

Grabchak, M. and Samorodnitsky, G. (2010). Do financial returns have finite or infinite variance? A paradox and an explanation. *Quantitative Finance*, 10:883–893.

Granger, C. W. J. and Morris, M. J. (1976). Time series modelling and interpretation. *Journal of the Royal Statistical Society, Series A*, 13:246–257.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

Hannan, E. J. and Kanter, M. (1977). Autoregressive processes with infinite variance. *Journal of Applied Probability*, 14:411–415.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029–1054.

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

Hautsch, N. (2012). *Econometrics of Financial High-Frequency Data*. Springer.

Jacquier, E., Polson, N. G., and Rossi, P. E. (1994). Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics*, 12:371–389.

Kitigawa, G. (2010). *Introduction to Time Series Modelling*. Chapman and Hall.

Kogon, S. M. and Williams, D. B. (1998). Characteristic function based estimation of stable distribution parameters. In Adler, R. J., Feldman, R. E., and Taqqu, M. S., editors, *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhäuser.

Lee, J. H. and Shin, D. W. (1997). Maximum likelihood estimation for ARMA models in the presence of ARMA errors. *Communications in Statistics - Theory and Methods*, 26:1057–1072.

Lin, J.-W. and McLeod, A. I. (2008). Portmanteau tests for ARMA models with infinite variance. *Journal of Time Series Analysis*, 29:600–617.

Ling, S. (2005). Self-weighted least absolute deviation estimation for infinite variance autoregressive models. *Journal of the Royal Statistical Society, Series B*, 67:381–393.

Liu, J. and David, H. A. (1989). Quantiles of sums and expected values of ordered sums. *Australian Journal of Statistics*, 31:469–474.

Lütkepohl, H. (1986). Forecasting temporally aggregated vector ARMA processes. *Journal of Forecasting*, 5:85–95.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36:394–419.

Marcellino, M. (1999). Some consequences of temporal aggregation in empirical analysis. *Journal of Business and Econimc Statistics*, 17:129–136.

Matsui, M. and Takemura, A. (2006). Some improvements in numerical evaluation of symmetric stable density and its derivatives. *Communications in Statistics - Theory and Methods*, 35:149–172.

McCulloch, J. H. (1986). Simple consistent estimators of stable distribution parameters. *Communications in Statistics - Simulation and Computation*, 15:1109–1136.

McKenzie, E. (1988). The distributional structure of finite moving average processes. *Journal of Applied Probability*, 25:313–321.

Meddahi, N. and Renault, E. (2005). Temporal aggregation of volatility models. *Journal of Econometrics*, 119:355–379.

Melino, A. and Turnbull, S. M. (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics*, 45:239–265.

Miazaki, E. S. and Dorea, C. C. Y. (1993). Estimation of the parameters of a time series subject to the error of rotation sampling. *Communications in Statistics - Theory and Methods*, 22:805–825.

Mikosch, T., Gadrich, T., Kluppelberg, C., and Adler, R. J. (1995). Parameters estimation for ARMA models with infinite variance innovations. *The Annals of Statistics*, 23:305–326.

Müller, U. A., Dacorogna, M. M., Davé, R. D., Olsen, R. B., Pictet, O. V., and Weizsäcker, J. E. (1997). Volatility of different time resolutions - analysing the dynamics of market components. *Journal of Empirical Finance*, 4:213–239.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, 7:308–313.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59:347–370.

Newey, W. K. and West, K. D. (1987). A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55:703–708.

Nolan, J. P. (1998). Parameterizations and modes of stable distributions. *Statistics and Probability Letters*, 38:187–195.

Nolan, J. P. (2001). Maximum likelihood estimation of stable parameters. In Barndorff-Nielsen, O. E., Mikosch, T., and Resnick, I., editors, *Levy Processes: Theory and Application*. Birkhäuser.

Nolan, J. P. (2014). Private email correspondence.

Nolan, J. P. (2015). *Stable Distributions - Models for Heavy Tailed Data*. Birkhäuser, Boston. In progress, Chapter 1 online at academic2.american.edu/∼jpnolan.

Oberhofer, W. and Haupt, H. (2005). The asymptotic distribution of the unconditional quantile estimator under dependence. *Statistics and Probability Letters*, 73:243–250.

Osborne, M. F. M. (1959). Brownian motion in the stock market. *Operations Research*, 7:145–173.

Pagan, A. (1996). The econometrics of financial markets. *Journal of Empirical Finance*, 3:15–102.

Pagano, M. (1974). Estimation of models of autoregressive signal plus white noise. *The Annals of Statistics*, 1:99–108.

Pan, J., Wang, H., and Yao, Q. (2007). Weighted least absolute deviations estimation for ARMA models with infinite variance. *Econometric Theory*, 23:852–879.

Paulson, A. S., Holcomb, E. W., and Leitch, R. A. (1975). The estimation of the parameters of the stable laws. *Biometrika*, 62:163–170.

Peters, G. W. and Shevchenko, P. V. (2015). *Advances in heavy tailed risk modelling: A handbook of operational risk*. Wiley.

Press, S. J. (1972). Estimation in univariate and multivariate stable distributions. *Journal of the American Statistical Association*, 67:842–846.

Rachev, S. and Mittnik, S. (2000). *Stable Paretian Models in Finance*. Wiley.

Rosadi, D. (2007). Identification of moving average process with infinite variance. *Statistics and Probability Letters*, 77:1490–1496.

Rosenfeld, G. (1976). Identification of time series with infinite variance. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 25:147–153.

Rudin, W. (1987). *Real and Complex Analysis*. McGraw-Hill, 3rd edition.

Runde, R. (1997). The asymptotic null distribution of the box-pierce q-statistic for random variables with infinite variance. an application to german stock returns. *Journal of Econometrics*, 78:205–216.

Salazar, E. and Fereira, M. A. R. (2011). Temporal aggregation of lognormal AR processes. *Journal of Time Series Analysis*, 32:661–671.

Sbrana, G. and Silvestrini, A. (2013). Aggregation of exponential smoothing processes with an application to portfolio risk evaluation. *Journal of Banking and Finance*, 37:1437–1450.

Sen, P. K. (1968). Asymptotic normality of sample quantiles for m-dependent processes. *The Annals of Mathematical Statistics*, 39:1724–1730.

Sen, P. K. (1972). On the Bahadur representation of sample quantiles for sequences of ϕ -mixing random variables. *Journal of Multivariate Analysis*, 2:77–95.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.

Silvestrini, A. and Veredas, D. (2008). Temporal aggregation of univariate and multivariate time series models: A survey. *Journal of Economic Surveys*, 22:458–497.

Staudenmayer, J. and Buonaccorsi, J. P. (2005). Measurement error in linear autoregressive models. *Journal of the American Statistical Association*, 100:841–852.

Stram, D. O. and Wei, W. W. S. (1986). Temporal aggregation in the ARIMA process. *Journal of Time Series Analysis*, 7:279–292.

Subba Rao, T. and Gabr, M. M. (1984). *An Introduction to Bispectral Analysis and Bilinear Time Series Models*. Springer.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes - a study of daily sugar prices 1961-79. In Anderson, O. D., editor, *Time Series Analysis: Theory and Practice*. Amsterdam: North Holland.

Taylor, S. J. (1986). *Modelling Financial Time Series*. Wiley.

Teichmoeller, J. (1971). A note on the distribution of stock price changes. *Journal of the American Statistical Association*, 66:282–284.

Tiao, G. C. (1972). Asymptotic behaviour of temporal aggregates of time series. *Biometrika*, 59:525–531.

Tsai, H. and Chan, K. S. (2005). Temporal aggregation of stationary and non-stationary discrete-time processes. *Journal of Time Series Analysis*, 26:613–624.

Tsay, R. S. (2010). *Analysis of Financial Time Series*. Wiley, 3rd edition.

Walker, A. M. (1960). Some consequences of superimposed error in time series analysis. *Biometrika*, 47:33–43.

Watson, R. and Gordon, I. (1986). On quantiles of sums. *Australian Journal of Statistics*, 28:192–199.

Weiss, A. A. (1984). Systematic sampling and temporal aggregation in time series models. *Journal of Econometrics*, 26:271–281.

Wilcox, R. R. (2005). *Introduction to robust estimation and hypothesis testing.* Elsevier/Academic Press.

Wilks, S. S. (1962). *Mathematical Statistics.* Wiley.

Wilson, G. (1969). Factorization of the covariance generating function of a pure moving average process. *SIAM Journal of Numerical Analysis*, 6:1–7.

Wise, J. (1956). Stationarity conditions for stochastic processes of the autoregressive and moving-average type. *Biometrika*, 43:215–219.

Xekalaki, E. and Degiannakis, S. (2010). *ARCH models for financial applications.* Wiley.

Zhu, K. and Ling, S. (2012). The global weighted LAD estimators for finite/infinite variance ARMA(p,q) models. *Econometric Theory*, 28:1065–1086.

Zolotarev, V. M. (1986). *One-dimensional Stable Distributions*, volume 65 of *Translations of Mathematical Monographs.* American Mathematical Society.

Zumbach, G. (2011). Characterizing heteroskedasticity. *Quantitative Finance*, 11:1357–1369.