Parameterisation and Efficient MCMC Estimation of Non-Gaussian State Space Models

Chris M. Strickland*, Gael M. Martin† and Catherine S. Forbes‡

October 4, 2007

Abstract

The impact of parameterisation on the simulation efficiency of Bayesian Markov chain Monte Carlo (MCMC) algorithms for two non-Gaussian state space models is examined. Specifically, focus is given to particular forms of the stochastic conditional duration (SCD) model and the stochastic volatility (SV) model, with four alternative parameterisations of each model considered. A controlled experiment using simulated data reveals that relationships exist between the simulation efficiency of the MCMC sampler, the magnitudes of the population parameters and the particular parameterisation of the state space model. Results of an empirical analysis of two separate transaction data sets for the SCD model, as well as equity and exchange rate data sets for the SV model, are also reported. Both the simulation and empirical results reveal that substantial gains in simulation efficiency can be obtained from simple reparameterisations of both types of non-Gaussian state space models.

Keywords: Bayesian Estimation; Non-Centred Parameterisations; Inefficiency Factor; Stochastic Volatility Model; Stochastic Conditional Duration Model.

JEL codes: C11, C15, C22.

*School of Mathematics, Queensland University of Technology, Queensland, Australia and Department of Econometrics and Business Statistics, Monash University, Victoria, Australia.

†Corresponding Author: Department of Econometrics and Business Statistics, PO Box 11E, Monash University, Vic., 3800, Australia. Email: gael.martin@buseco.monash.edu.au. Phone: 61 3 9905 1189. Fax: 61 3 9905 5474.

‡Department of Econometrics and Business Statistics, Monash University, Victoria, Australia.

1 Introduction

The state space form provides a convenient framework for building time series models for observed phenomena, whereby relatively simple model components are combined to explain potentially complex dependencies in observed data. Linear Gaussian state space models, in particular, have had a long history in both estimation and forecasting applications, with many traditional time series models, such as autoregressive integrated moving average (ARIMA) models, having a linear Gaussian state space representation. This representation is often exploited in computing algorithms, both Bayesian and classical, designed to infer parameter values and produce forecasts.

More recently, non-Gaussian state space models have enjoyed increasing popularity, partly due to developments in inferential simulation techniques. In particular, non-Gaussian state space models have been used to characterise the dynamic features of various financial time series, such as the time between transactions or the volatility of asset returns, with a range of Markov chain Monte Carlo (MCMC) methods having been employed to implement Bayesian analyses of such models. Concurrent with the increased application of MCMC sampling schemes to non-Gaussian state space models, has been the revelation that substantial improvements in the simulation efficiency of MCMC schemes, in a variety of contexts, can sometimes be obtained though simple model reparameterisation. Relevant work includes Gelfand, Sahu and Carlin (1995), Roberts and Sahu (1997), Pitt and Shephard (1999), Robert and Mengersen (1999), Früwirth-Schnatter and Sögner (2003), Papaspiliopoulos, Roberts and Sköld (2003), Früwirth-Schnatter (2004), Roberts, Papaspiliopoulos and Dellaportas (2004) and Bos and Shephard (2006).

This paper contributes to the literature by examining the effect of particular types of reparameterisation in two specific non-Gaussian state space models. Firstly, a form of the stochastic conditional duration (SCD) model of Bauwens and Veredas (2004), based on the assumption of conditionally exponential data, is considered. Variants of such a model, under alternative distributional assumptions, have recently been applied to financial trade durations in Strickland, Forbes and Martin (SFM hereafter) (2006), with an MCMC algorithm developed for one particular parameterisation. In the present paper, in which the focus is on documenting computational performance under a range of scenarios, the use of the exponential distribution serves to minimise the number of parameters involved in the reparameterisations, such that the number of results to be

produced and summarised is manageable. The second model considered is the stochastic volatility (SV) model for financial returns, based on conditional normality; see for example Jacquier, Polson and Rossi (1994), Shephard and Pitt (1997), and Kim, Shephard and Chib (1998), amongst many others. Again, the assumption of conditional normality is chosen for expositional convenience.

Comparisons are first conducted using artificially simulated data, based on multiple parameter settings, with the efficiency of the algorithms measured via inefficiency factors. The parameter settings are chosen to ensure that the simulated data resemble typical trade durations and financial returns data. Empirical evaluation of the alternative parameterisations of the SCD model is then conducted using trade durations data for two Australian listed companies: Broken Hill Proprietary Limited (BHP) and News Corporation (NCP), for the month of August 2001. These two data sets were analysed in SFM, using one particular parameterisation of an SCD model. The SV model is estimated using daily returns on the Morgan Stanley Capital Index (MSCI) between 1989 and 2002, and the pound/dollar daily exchange rate returns between 1981 and 1985. This exchange rate data was previously analysed in Harvey, Ruiz and Shephard (1994), Kim, Shephard and Chib (1998) and Durbin and Koopman (2001), using specific parameterisations and various numerical algorithms that are alternatives to those used in this paper.

Four alternative parameterisations of both non-Gaussian state space models are considered. Using the nomenclature in the literature, the parameterisations are referred to as: 'centred', 'non-centred in location', 'non-centred in scale' and 'non-centred in both location and scale', with all parameterisations augmented to incorporate regressors in the state equation. In the centred parameterisation all of the parameters (persistence, scale and location) appear in the state equation, while the non-centred parameterisations are based on either the location or scale parameter, or both, appearing in the measurement equation. The simulation results reveal clear relationships between the simulation efficiency of the MCMC sampler, the chosen parameterisation and the magnitudes of the population parameters. Overall, both the experimental and empirical results indicate that parameterisations that place parameters in the measurement equation, as opposed to the state equation, tend to perform better. Specifically, substantial improvements in efficiency are associated with relocation of the scale parameter, either on its own, or in conjunction with the location parameter.

An outline of the paper is as follows. Section 2 defines the two non-Gaussian state

space models that are considered, including the four alternative parameterisations of each model. Section 3 provides details of the algorithms used to estimate the different parameterisations. All algorithms are modifications of the hybrid Gibbs/Metropolis-Hastings (MH) MCMC sampling scheme proposed in SFM. Section 4 presents the criteria used to compare the performance of the algorithms and summarises the results of the simulation experiment. Section 5 details the four empirical applications, two using the SCD model and two using the SV model. Section 6 concludes.

2 A Non-Gaussian State Space Framework

2.1 The Centred Parameterisation

Defining $\mathbf{y} = (y_1, y_2, \ldots, y_T)'$ as the $(T \times 1)$ observation vector and assuming conditional independence, the measurement equation of a non-Gaussian state space model may be represented by the following probability relation

$$p(\mathbf{y}|\boldsymbol{\alpha}) = \prod_{t=1}^{T} p(y_t|\alpha_t), \qquad t = 1, 2, ..., T, \tag{1}$$

where $p(y_t|\alpha_t)$ denotes the probability density function (pdf) for y_t, conditional on α_t. The $(t+1)^{th}$ element of the $(T \times 1)$ state vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_T)'$ is assumed to be generated by the following state equation,

$$\alpha_{t+1} = \delta_1 + \mathbf{W}_t'\boldsymbol{\delta}_2 + \phi\alpha_t + \sigma_\eta\eta_t, \qquad t = 1, 2, ..., T-1, \tag{2}$$

where δ_1 is a constant, $\boldsymbol{\delta}_2 = (\delta_2, \delta_3, \ldots, \delta_k)'$ is a $((k-1) \times 1)$ vector of coefficients and $\mathbf{W}_t'$ is the t^{th} row of the $(T \times (k-1))$ matrix of regressors, $\mathbf{W}$. It is further assumed that $\eta_t \sim N(0,1)$ is independent of $y_t|\alpha_t$ and that $|\phi| < 1$ and $\sigma_\eta^2 > 0$. Defining the parameter μ implicitly via $\delta_1 = \mu(1-\phi)$, with $\mu \in \mathbb{R}$, the full vector of unknown parameters is denoted by $\boldsymbol{\theta} = (\mu, \boldsymbol{\delta}_2', \phi, \sigma_\eta)'$. The assumed pdf for the initial state is given by

$$p(\alpha_1|\mathbf{W}_0, \boldsymbol{\theta}) \sim N\left(\mu + \frac{\mathbf{W}_0'\boldsymbol{\delta}_2}{1-\phi}, \frac{\sigma_\eta^2}{1-\phi^2}\right). \tag{3}$$

The specification of the state space model in (1), (2) and (3) is referred to here as the 'centred parameterisation'. Adopting the assumption of a conditional exponential distribution in the case of the SCD model, it follows that

$$p(y_t|\alpha_t) = \exp(-\alpha_t)\exp\left\{-y_t\exp\left(-\alpha_t\right)\right\}, \tag{4}$$

for all $t = 1, 2, ..., T$, with conditional mean $E[y_t|\alpha_t] = \exp(\alpha_t)$. The SV model assumes that the (potentially demeaned) y_t, conditional on α_t, has a normal distribution, with

$$p(y_t|\alpha_t) = (2\pi \exp(\alpha_t))^{-1/2} \exp\left(-\frac{1}{2\exp(\alpha_t)} y_t^2\right), \tag{5}$$

for all $t = 1, 2, ..., T$, with conditional variance $Var(y_t|\alpha_t) = \exp(\alpha_t)$. Note that in this so-called 'centred' parameterisation none of the elements of $\boldsymbol{\theta}$ explicitly enter the measurement equation. This is in contrast to the reparameterisations presented in the following section.

2.2 Reparameterising the Non-Gaussian State Space Model

The centred parameterisation is the most commonly used parameterisation of the non-Gaussian state space model, at least within the SCD and SV literature. See, for example, Jacquier, Polson and Rossi (1994), Kim, Shephard and Chib (1998), Bauwens and Veredas (2004) and SFM. However, it is equally valid to modify the model specification by moving one or both of μ and σ_η into the measurement equation. Although, conditional on $\boldsymbol{\theta}$, the probability distribution of the data remains unchanged, reparameterisation may impact upon the simulation efficiency of the MCMC algorithms used to infer such model parameters from the observed data.

2.2.1 Non-centred in location

The first alternative parameterisation of the non-Gaussian state space model considered is based upon a location shift for the state variable. Define $\widetilde{\alpha}_t = \alpha_t - \mu$ and transform the model accordingly. The implied measurement equation is

$$p(\mathbf{y}|\boldsymbol{\alpha}, \mu) = \prod_{t=1}^{T} p(y_t|\widetilde{\alpha}_t, \mu), \quad t = 1, 2, ..., T, \tag{6}$$

with state equation given by

$$\widetilde{\alpha}_{t+1} = \mathbf{W}_t'\boldsymbol{\delta}_2 + \phi\widetilde{\alpha}_t + \sigma_\eta\eta_t, \quad t = 1, 2, ..., T-1. \tag{7}$$

From (3), the implied pdf of the initial state for the model that is non-centred in location is

$$p(\widetilde{\alpha}_1|\mathbf{W}_0, \phi, \sigma_\eta) \sim N\left(\frac{\mathbf{W}_0'\boldsymbol{\delta}_2}{1-\phi}, \frac{\sigma_\eta^2}{1-\phi^2}\right). \tag{8}$$

Under the assumption of a conditional exponential distribution for the SCD model, each component in (6) is given by

$$p(y_t|\widetilde{\alpha}_t, \mu) = \exp(-(\widetilde{\alpha}_t + \mu)) \exp\{-y_t \exp(-(\widetilde{\alpha}_t + \mu))\}, \tag{9}$$

with the conditional mean now specified as $E[y_t|\widetilde{\alpha}_t, \mu] = \exp(\mu + \widetilde{\alpha}_t)$. For the SV model it follows that

$$p(y_t|\widetilde{\alpha}_t, \mu) = \left((2\pi \exp(\widetilde{\alpha}_t + \mu))^{-1/2} \exp\left(-\frac{1}{2\exp(\widetilde{\alpha}_t + \mu)} y_t^2\right)\right), \tag{10}$$

where the conditional variance is now given as $Var(y_t|\widetilde{\alpha}_t, \mu) = \exp(\mu + \widetilde{\alpha}_t)$. Adopting the terminology of Gelfand, Sahu and Carlin (1995), applied in the context of a random effects model, the model in (6) to (9) is referred to here as 'non-centred in location'. In the case where $\boldsymbol{\delta}_2 = \mathbf{0}$, this terminology is used to refer to the fact that the state variable, $\widetilde{\alpha}_{t+1}$, no longer has the unconditional expectation μ, but rather has a zero unconditional mean. From (9) and (10) it is clear that the location parameter, μ, now appears explicitly in the measurement equation.

2.2.2 Non-centred in scale

Next, a scale adjustment to the state variable from the centred parameterisation is considered. Defining the new state variable as $\alpha_t^* = \frac{\alpha_t}{\sigma_\eta}$ produces a parameterisation that is referred to as 'non-centred in scale'. In this case the measurement equation is

$$p(\mathbf{y}|\boldsymbol{\alpha}, \sigma_\eta) = \prod_{t=1}^{T} p(y_t|\alpha_t^*, \sigma_\eta), \quad t = 1, 2, ..., T, \tag{11}$$

and the state equation,

$$\alpha_{t+1}^* = \delta_1^* + \mathbf{W}_t'\boldsymbol{\delta}_2^* + \phi\alpha_t^* + \eta_t, \quad t = 1, 2, ..., T-1, \tag{12}$$

where $\delta_1^* = \frac{\mu(1-\phi)}{\sigma_\eta}$ and $\boldsymbol{\delta}_2^* = \frac{\boldsymbol{\delta}_2}{\sigma_\eta}$. The implied pdf of the initial state is

$$p(\alpha_1^*|\mathbf{W}_0, \phi) \sim N\left(\frac{\delta_1^* + \mathbf{W}_0'\boldsymbol{\delta}_2^*}{(1-\phi)}, \frac{1}{1-\phi^2}\right). \tag{13}$$

Under the assumption of a conditional exponential distribution for the SCD model, each component in (11) is given by

$$p(y_t|\alpha_t^*, \sigma_\eta) = \exp(-(\sigma_\eta\alpha_t^*)) \exp\{-y_t \exp(-(\sigma_\eta\alpha_t^*))\}, \tag{14}$$

with the conditional mean now given by $E\left[y_t|\alpha_t^*,\sigma_\eta\right]=\exp\left(\sigma_\eta\alpha_t^*\right)$. For the SV model it follows that

$$p(y_t|\alpha_t^*,\sigma_\eta)=\left(\left(2\pi\exp\left(\sigma_\eta\alpha_t^*\right)\right)^{-1/2}\exp\left(-\frac{1}{2\exp\left(\sigma_\eta\alpha_t^*\right)}y_t^2\right)\right), \tag{15}$$

where the conditional variance is now $Var\left(y_t|\alpha_t^*,\sigma_\eta\right)=\exp\left(\sigma_\eta\alpha_t^*\right)$. Under this parameterisation the scale parameter, σ_η, appears explicitly in the measurement equation only. It enters the state equation indirectly, via δ_1^* and δ_2^*.

2.2.3 Non-centred in both location and scale

By defining $\alpha_t^{**}=\frac{\alpha_t-\mu}{\sigma_\eta}$, both scale and location adjustments are made to the original centred parameterisation state variable. This parameterisation, referred to as 'non-centred in both location and scale', has measurement equation given by

$$p(\mathbf{y}|\boldsymbol{\alpha},\mu,\sigma_\eta)=\prod_{t=1}^{T}p(y_t|\alpha_t^{**},\mu,\sigma_\eta),\quad t=1,2,...,T, \tag{16}$$

and state equation given by

$$\alpha_{t+1}^{**}=\mathbf{W}_t'\boldsymbol{\delta}_2^*+\phi\alpha_t^{**}+\eta_t,\quad t=1,2,...,T-1, \tag{17}$$

where $\boldsymbol{\delta}_2^*$ is as defined earlier. The implied pdf of the initial state is

$$p\left(\alpha_1^{**}|\mathbf{W}_0,\phi\right)\sim\left(\frac{\mathbf{W}_0'\boldsymbol{\delta}_2^*}{(1-\phi)},\frac{1}{1-\phi^2}\right). \tag{18}$$

Under the assumption of a conditional exponential distribution for the SCD model, each component in (16) is given by

$$p(y_t|\alpha_t^{**},\mu,\sigma_\eta)=\exp(-\left(\sigma_\eta\alpha_t^{**}+\mu\right))\exp\left\{-y_t\exp\left(-\left(\sigma_\eta\alpha_t^{**}+\mu\right)\right)\right\}, \tag{19}$$

with the conditional mean specified as $E\left[y_t|\alpha_t^{**},\mu,\sigma_\eta\right]=\exp\left(\mu+\sigma_\eta\alpha_t^{**}\right)$. For the SV model each component in (16) is given by

$$p(y_t|\alpha_t^{**},\mu,\sigma_\eta)=\left(\left(2\pi\exp\left(\sigma_\eta\alpha_t^{**}+\mu\right)\right)^{-1/2}\exp\left(-\frac{1}{2\exp\left(\sigma_\eta\alpha_t^{**}+\mu\right)}y_t^2\right)\right), \tag{20}$$

with conditional variance, $Var\left(y_t|\alpha_t^{**},\mu,\sigma_\eta\right)=\exp\left(\mu+\sigma_\eta\alpha_t^{**}\right)$. Note that both the location and scale parameters enter the measurement equation, and neither explicitly enters the state equation.

3 Bayesian Estimation

The MCMC sampling scheme of SFM is used as the algorithm for the centred parameterisation, and is then modified to cater for the non-centred parameterisations. As highlighted in SFM, the algorithm is very flexible, being readily applicable beyond both the specific model specifications examined in that paper and the specifications examined in the current paper. In particular, it is more general than the distribution-specific algorithm outlined by Shephard (1994) and Carter and Kohn (1994), and implemented by Kim, Shephard and Chib (1998) for the SV model, whereby the non-Gaussian density in the (linearised) measurement equation is approximated by a mixture of normal densities. Shephard and Pitt (1997) provide a similar algorithm to SFM for non-Gaussian state space models, with focus given to the SV model. In SFM the focus is given to the SCD model using three alternative distributional assumptions, as well as being augmented for regressors in the state equation. The latter authors use the approach of Durbin and Koopman (2000, 2001) to produce a linear Gaussian approximation to the measurement equation, such that alterations to the algorithm required to accommodate different distributional assumptions (and, hence, data types) are relatively straightforward and transparent. In contrast to Durbin and Koopman, who use the approximation as a part of an importance sampling scheme, SFM use the approximating model to construct a candidate distribution in an MH step imbedded in an outer Gibbs chain. SFM argue that the MCMC approach is potentially more efficient than the importance sampling methodology, as the Gibbs sampler allows the high-dimensional latent vector to be broken down into blocks of lower dimension. This has particular relevance to non-Gaussian financial data sets (such as durations and returns) that typically contain a large number of observations. Pitt (2000) also makes note of this drawback of the Durbin and Koopman approach. An alternative MCMC sampler for non-Gaussian state space models is implemented in Jung, Kukuk and Liesenfeld (2006) in a count data context.

3.1 The Joint Posterior

The joint posterior for the full set of unknowns in the non-Gaussian state space model is given by

$$p(\boldsymbol{\alpha}, \boldsymbol{\theta}|\mathbf{y}, \mathbf{W}) \propto p(\mathbf{y}|\boldsymbol{\alpha}, \boldsymbol{\theta}) \times p(\boldsymbol{\alpha}|\mathbf{W}, \boldsymbol{\theta}) \times p(\boldsymbol{\theta}), \tag{21}$$

where $p(\boldsymbol{\alpha}|\mathbf{W}, \boldsymbol{\theta})$ denotes the joint pdf of $\boldsymbol{\alpha}$ conditional on $\boldsymbol{\theta}$ and the observed $\mathbf{W}$, and $p(\boldsymbol{\theta})$ is the prior pdf for $\boldsymbol{\theta}$. The joint pdf $p(\mathbf{y}|\boldsymbol{\alpha}, \boldsymbol{\theta})$ is as defined in (1), (6), (11) or (16),

depending on the chosen parameterisation. The joint density for the state vector is

$$p(\boldsymbol{\alpha}|\mathbf{W},\boldsymbol{\theta}) = \left\{\prod_{t=1}^{T-1} p(\alpha_{t+1}|\mathbf{W}_t,\alpha_t,\boldsymbol{\theta})\right\} \times p(\alpha_1|\mathbf{W}_0,\boldsymbol{\theta}), \tag{22}$$

where $p(\alpha_{t+1}|\mathbf{W}_t,\alpha_t,\boldsymbol{\theta})$ is given by (2), (7), (12) or (17) and $p(\alpha_1|\mathbf{W}_0,\boldsymbol{\theta})$ is given by (3), (8), (13) or (18), once again according to the specified parameterisation.

The following subsections summarise the algorithm used for each of the parameterisations. To simulate values of the parameters in the measurement equation, for all three non-centred parameterisations, the MCMC algorithms use an MH algorithm with a normally distributed candidate. The mean and variance of the candidate distribution are obtained through numerical maximisation of the conditional log likelihood function implied by the relevant measurement equation. Each of the algorithms outlined below requires only minor modifications to cater for different distributional assumptions in the measurement equation. As such, the ease with which the practitioner can modify the associated code, given alternative distributional assumptions, is maintained from SFM.

3.2 The Centred Parameterisation

Estimation of the model in Section 2.1 is essentially described in SFM (Section 3). Note that the modifications required for estimation of the SV model are not explicitly described in SFM. However the necessary modifications can be readily deduced from details provided in SFM and Durbin and Koopman (2001, Chp. 11). The steps of the Gibbs-based sampler are briefly summarised as follows:

1. Initialise $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$. Note that $\boldsymbol{\theta}$ needs to be initialised for the MH algorithm used in Step 2.

2. Sample $\boldsymbol{\theta}|\mathbf{y},\mathbf{W},\boldsymbol{\alpha}$.

3. Sample $\boldsymbol{\alpha}|\mathbf{y},\mathbf{W},\boldsymbol{\theta}$, where $\boldsymbol{\alpha}$ is broken up into blocks of size greater than or equal to one, along the lines of SFM.

4. Repeat steps 2-3 until convergence has been achieved.

Depending on the form of $p(\boldsymbol{\theta})$, there may be no closed-form representation for $p(\boldsymbol{\theta}|\mathbf{y},\mathbf{W},\boldsymbol{\alpha})$. However, standard Bayesian linear regression theory provides a good candidate through which draws from $p(\boldsymbol{\theta}|\mathbf{y},\mathbf{W},\boldsymbol{\alpha})$ can be obtained indirectly using an

MH algorithm. Sampling from $p(\boldsymbol{\alpha}|\mathbf{y}, \mathbf{W}, \boldsymbol{\theta})$ is also accomplished indirectly, as the non-Gaussian measurement equation implies that there is no closed form solution for the conditional posterior for $\boldsymbol{\alpha}$. This is the most complex component of the algorithm, with full details provided in SFM (Section 3.2.3). Crucially, this component of the algorithm is common to all four parameterisations considered in the current paper.

3.3 Non-Centred in Location

The steps required to implement the Gibbs-based sampler for the model in Section 2.2.1 can be summarised as follows:

1. Initialise $\widetilde{\boldsymbol{\alpha}} = (\widetilde{\alpha}_1, \widetilde{\alpha}_2, \ldots, \widetilde{\alpha}_T)'$ and $\boldsymbol{\theta}$. Note that $\boldsymbol{\delta}_2$, ϕ, σ_η and μ need to be initialised, in addition to $\widetilde{\boldsymbol{\alpha}}$, for the MH algorithm used in step 2. The initialised value of μ is used both in step 2 and in the MH algorithm in step 3.

2. Sample $\boldsymbol{\delta}_2, \phi, \sigma_\eta|\mathbf{y}, \mathbf{W}, \widetilde{\boldsymbol{\alpha}}, \mu$.

3. Sample $\mu|\mathbf{y}, \mathbf{W}, \widetilde{\boldsymbol{\alpha}}, \boldsymbol{\delta}_2, \phi, \sigma_\eta$.

4. Sample $\widetilde{\boldsymbol{\alpha}}|y, W, \theta$, where $\widetilde{\boldsymbol{\alpha}}$ is broken up into blocks of size greater than or equal to one. (Sample along the lines of SFM, Section 3.2.3)

5. Repeat steps 2-4 until convergence has been achieved.

The linear Gaussian nature of the state equation once again implies that when sampling from $p(\boldsymbol{\delta}_2, \phi, \sigma_\eta|\mathbf{y}, \mathbf{W}, \widetilde{\boldsymbol{\alpha}}, \mu)$ a good candidate can be obtained using standard Bayesian linear regression theory, with draws from $p(\boldsymbol{\delta}_2, \phi, \sigma_\eta|\mathbf{y}, \mathbf{W}, \widetilde{\boldsymbol{\alpha}}, \mu)$ then obtained indirectly using an MH algorithm. Sampling from $p(\mu|\mathbf{y}, \mathbf{W}, \widetilde{\boldsymbol{\alpha}}, \boldsymbol{\delta}_2, \phi, \sigma_\eta)$ is conducted using an MH algorithm, based on a normal candidate distribution with mean, $\overline{\mu}$, obtained via numerical optimisation of the conditional log likelihood given by the measurement equation in (6). The variance of the candidate distribution is the usual estimator of the asymptotic variance of the conditional maximum likelihood estimator (MLE), that is, the inverse of the second derivative of the negative conditional log likelihood function, evaluated at $\overline{\mu}$.

3.4 Non-Centred in Scale

Estimation of the non-Gaussian state space model for the non-centred in scale parameterisation, defined in (11), (12) and (13), is summarised by the following steps:

1. Initialise $\boldsymbol{\alpha}^* = (\alpha_1^*, \alpha_2^*, \ldots, \alpha_T^*)'$ and $\boldsymbol{\theta}$. Note that μ, $\boldsymbol{\delta}_2$ and ϕ need to be initialised, in addition to $\boldsymbol{\alpha}^*$ and σ_η, for the MH algorithm used in step 2. The initialised value of σ_η is used both in step 2 and in the MH algorithm in step 3.

2. Sample $\mu, \boldsymbol{\delta}_2, \phi | \mathbf{y}, \mathbf{W}, \boldsymbol{\alpha}^*, \sigma_\eta$.

3. Sample $\sigma_\eta | \mathbf{y}, \mathbf{W}, \boldsymbol{\alpha}^*, \mu, \boldsymbol{\delta}_2, \phi$.

4. Sample $\boldsymbol{\alpha}^* | \mathbf{y}, \mathbf{W}, \boldsymbol{\theta}$, where $\boldsymbol{\alpha}^*$ is broken up into blocks of size greater than or equal to one. (Sample along the lines of SFM, Section 3.2.3)

5. Repeat steps 2-4 until convergence has been achieved.

As equation (12) remains linear and Gaussian under this reparameterisation, standard Bayesian linear regression theory again provides an appropriate candidate for sampling from $p(\mu, \boldsymbol{\delta}_2, \phi | \mathbf{y}, \mathbf{W}, \boldsymbol{\alpha}^*, \sigma_\eta)$ using an MH algorithm. For this parameterisation, sampling from the full conditional distribution $p(\sigma_\eta | \mathbf{y}, \mathbf{W}, \boldsymbol{\alpha}^*, \mu, \boldsymbol{\delta}_2, \phi)$ is conducted using an MH algorithm, with the candidate distribution specified comparably to that for μ in Section 3.3. That is, the normal candidate distribution for σ_η is centred around $\overline{\sigma}_\eta$, the conditional MLE for σ_η obtained via numerical optimisation of the conditional log likelihood given by the measurement equation in (11). The variance of the candidate distribution is the inverse of the second derivative of the negative conditional log likelihood function, evaluated at the conditional MLE.

3.5 Non-Centred in Both Location and Scale

Estimation of the non-Gaussian state space model for the non-centred in location and scale parameterisation, defined in (16), (17) and (18) is summarised by the following steps:

1. Initialise $\boldsymbol{\alpha}^{**} = (\alpha_1^{**}, \alpha_2^{**}, \ldots, \alpha_T^{**})'$ and $\boldsymbol{\theta}$. Note that $\boldsymbol{\delta}_2$ and ϕ need to be initialised, in addition to $\boldsymbol{\alpha}^{**}$, μ and σ_η, for the MH algorithm used in step 2. The initialised values of μ and σ_η are used both in step 2 and in the MH algorithms in steps 3 and 4.

2. Sample $\boldsymbol{\delta}_2, \phi | \mathbf{y}, \mathbf{W}, \boldsymbol{\alpha}^{**}, \mu, \sigma_\eta$.

3. Sample $\mu, \sigma_\eta | \mathbf{y}, \mathbf{W}, \boldsymbol{\alpha}^{**}, \phi, \boldsymbol{\delta}_2$.

4. Sample $\boldsymbol{\alpha}^{**}|\mathbf{y}, \mathbf{W}, \boldsymbol{\theta}$, where $\boldsymbol{\alpha}^{**}$ is broken up into blocks of size greater than or equal to one. (Sample along the lines of SFM, Section 3.2.3)

5. Repeat steps 2-4 until convergence has been achieved.

As equation (17) is still linear and Gaussian, a candidate for $p(\boldsymbol{\delta}_2, \phi|\mathbf{y}, \boldsymbol{\alpha}^{**}, \mu, \sigma_\eta)$ can be obtained using Bayesian linear regression theory, and a standard MH algorithm applied. The parameters μ and σ_η are jointly sampled using an MH algorithm, with a bivariate normal candidate distribution. The candidate distribution is centred at the conditional MLE based on the measurement equation in (16), and the variance-covariance matrix is given by the inverse of the negative Hessian matrix, evaluated at the conditional MLE.

4 Simulation Experiment

A simulation experiment is used to explore the efficiency of the MCMC algorithms under the four different parameterisations of each of the SCD and SV non-Gaussian state space models. To simplify the experiment, $\boldsymbol{\delta}_2$ is set equal to a vector of zeros. The parameters ϕ and σ_η are assigned a range of empirically plausible values for each of the specified models, which leads to a total of 18 simulated data sets, each of which has $T = 5000$ observations. The values of ϕ that are considered for both the SCD and SV models are $\{0.8, 0.9, 0.95\}$. The values of σ_η under consideration are $\{0.1, 0.2, 0.3\}$ and $\{0.2, 0.3, 0.4\}$ for the SCD and the SV model, respectively. Specific references for the range of parameter values considered for the SCD model are Bauwens and Veredas (2004) and SFM, and for the SV model, Jacquier, Polson and Rossi (1994), Kim, Shephard and Chib (1998), Liesenfeld and Richard (2003) and Bos and Shephard (2006). The values also accord with certain of the empirical estimates produced in Section 5.

For the SCD model, the value of δ_1 is specified as

$$\delta_1 = -\frac{\sigma_\eta^2(1-\phi)}{2(1-\phi^2)}, \tag{23}$$

for given values of ϕ and σ_η. Defining δ_1 in this way ensures that the unconditional mean of the durations is always equal to 1. This assumption is consistent with empirical applications of the SCD model, as it is typically applied to a transaction data set that has been de-seasonalised assuming a multiplicative intraday pattern (producing observations that average to about 1); see, for example, Bauwens and Veredas (2004) and SFM. The expression in (23) implies a range of values for μ between -0.14 and -1.53.

For the SV model, δ_1 is also set conditionally on the specified values for σ_η and ϕ, with

$$\delta_1 = (1-\phi)\left[\ln(0.4) - \frac{\sigma_\eta^2}{2(1-\phi^2)}\right]. \tag{24}$$

Equation (24) implies an unconditional variance in the SV model of 0.4, a value comparable to that implied by typical empirical estimates of the SV parameters for daily exchange rate returns data and some stock index returns data, including the two data sets analysed in Section 5.2. Specifically, it corresponds to an expected annualised volatility of approximately 10% for continuously compounded returns.

Given the nature of the experiments it is natural to assume diffuse priors. Specifically, for δ_1 we assume a uniform prior over $\mathbb{R}$ and for ϕ we assume a uniform distribution over the $(-1,1)$ interval. An inverted-gamma distribution is assumed for σ_η, such that $p(\sigma_\eta) \sim IG\left(\frac{\sigma_r}{2}, \frac{S_\sigma}{2}\right)$ (see Zellner, 1996), with the hyper-parameters σ_r and S_σ set to 1.0001 and 0.01 respectively, implying very diffuse prior information on σ_η. As noted in the algorithmic details outlined above, the state vector for each parameterisation is sampled in blocks of size greater than or equal to one. For each iteration of the algorithm, the strictly positive random block size with expected length 36 is generated using one plus a Poisson variate with a mean of 35.

4.1 Simulation Efficiency Comparison

Under each parameterisation the so called inefficiency factor (IF) is used to benchmark simulation efficiency. The IF features prominently in the literature as a measure for comparing the performance of various alternative algorithms; see for example Chib and Greenberg (1996), Shephard and Pitt (1997) and Kim, Shephard and Chib (1998). The IF is calculated using the following formula

$$IF = 1 + 2\frac{B}{B-1}\sum_{i=1}^{B} K_{QS}\left(\frac{i}{B}\right)\widehat{\rho}_i, \tag{25}$$

where $\widehat{\rho}_i$ is the estimate of the correlation at lag i of the MCMC iterates, K_{QS} is the Quadratic Spectral (QS) kernel and B is the bandwidth. The QS kernel is defined as

$$K_{QS}(x) = \frac{25}{12\pi^2 x^2}\left(\frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5)\right). \tag{26}$$

The bandwidth B is selected automatically following Andrews (1991). Implementation of this procedure is summarised in SFM (Appendix B).

Further understanding of the IF comes through its relationship with the (estimated) Monte Carlo standard error (MCSE) for the posterior mean of any given parameter. That is,

$$MCSE^2 = \frac{\sigma^2_{MCMC}}{M} IF, \tag{27}$$

where σ^2_{MCMC} is the sample variance of the MCMC iterates and M denotes the number of iterations. From (27) it is apparent that the IF represents the ratio of the simulation variance (square of the estimated MCSE) to the variance that would arise from a hypothetical sample of M independent draws. For example, a value of 100 for IF implies that, for a given number of iterations M, correlation in the MCMC iterates produces an estimated MCSE for the posterior mean that is ten times greater than that which would be associated with an independent sample. The aim of reparameterisation is to reduce the dependence in the MCMC draws and, hence, the IF, to the greatest extent possible.

In the simulation experiment, the MCMC sampler is run for 25000 iterations with a burn-in period of 10000 iterations. The length of the burn-in period is assessed using the diagnostic of Raftery and Lewis (1992) and is concluded as being acceptable in all cases. IFs are calculated for each sequence of simulated draws of ϕ, σ_η, and δ_1, for each of the four parameterisations and for data simulated under each of the nine different parameter settings (for the SCD and SV models respectively). We say that parameterisation A is preferred to parameterisation B, for a particular parameter set $\boldsymbol{\theta} = (\phi, \sigma_\eta, \delta_1)$, if the largest IF for the elements of $\boldsymbol{\theta}$ under parameterisation A is less than the largest IF for the elements of $\boldsymbol{\theta}$ under parameterisation B. The rational in only comparing the largest IF for each relevant parameter setting is that the largest IF measures the minimum level of accuracy across the parameter set. Parameterisation A is then deemed to be superior to parameterisation B across all parameter settings if A is preferred to B in the majority of parameter settings considered for the model. Further, parameterisations will be ranked from 1 to 4, corresponding to most efficient to least efficient, if there is a clear ordering of pairwise comparisons between the four parameterisations.

4.1.1 The centred parameterisation

Table 1 reports the IFs for the MCMC sampler for the centred (C) parameterisation of both the SCD and SV models. It is apparent that there are large differences in the magnitude of the IFs for different parameter settings. Further, there are systematic patterns in the IFs that are dependent upon the magnitudes of the parameters σ_η and ϕ. For both models, as the true value of σ_η increases there is a corresponding increase

Inefficiency Factors (IF) for the Centred (C) Parameterisation

Panel A : SCD Model					Panel B: SV Model				
True σ_η :		*0.1*	*0.2*	*0.3*	*True* σ_η :		*0.2*	*0.3*	*0.4*
True ϕ		IF			*True* ϕ		IF		
0.95	ϕ	33	11	6	*0.95*	ϕ	23	12	8
	σ_η	**249**	**77**	**49**		σ_η	**167**	**93**	**73**
	δ_1	6	3	3		δ_1	19	10	7
0.9	ϕ	85	26	14	*0.9*	ϕ	77	34	21
	σ_η	**474**	**138**	**71**		σ_η	**445**	**160**	**109**
	δ_1	12	6	4		δ_1	71	29	17
0.8	ϕ	292	79	40	*0.8*	ϕ	562	106	50
	σ_η	**895**	**281**	**120**		σ_η	**937**	**306**	**146**
	δ_1	40	14	10		δ_1	547	99	46
Time (seconds)	7				**Time** (seconds)	8			

Table 1: C parameterisation simulation results for the SCD model (Panel A) and the SV model (Panel B). The top row of each panel contains the true values (in italics) of the scale parameter, while the first column contains the true values (in italics) of the persistence parameter. The second column in each panel contains the parameters to which IFs (reported in the subsequent columns) refer, with the maximum IF for each parameter setting indicated in bold. The average time (across parameter settings) to obtain 1000 iterations is reported in the bottom row of each panel.

in the simulation efficiency (i.e. a reduction in the IFs) of the marginal posterior mean estimates of all three parameters, ϕ, σ_η and δ_1, irrespective of the true value of ϕ. An increase in efficiency associated with all parameters is also apparent as the degree of persistence in the state variable, measured by ϕ, increases, for any given true value of σ_η. For all nine parameter settings, for both models, the maximum IF is associated with the parameter σ_η, indicating that dependence in the MCMC draws is the most marked for this particular parameter; see also Kim, Shephard and Chib (1998).

4.1.2 Non-centred in location

Table 2 reports the IFs for the MCMC sampler for the non-centred in location (NCL) parameterisation for both the SCD and SV models. Based on the benchmarking criterion defined in Section 4.1, the NCL parameterisation is not superior to the C parameterisation for either the SCD or SV models. Improvements in efficiency are seen in only four of the nine parameter sets considered for both models. Moreover, when reductions in the IFs do occur (as reported in Table 2 corresponding to the symbol $\Delta\%$) they are often fairly small.

As is the case with the C parameterisation, the simulation efficiency of the estimates of both ϕ and σ_η increases as the true values of ϕ and σ_η increase. Interestingly, however, the positive relationship between the efficiency of the estimates of δ_1 and the true values of ϕ and σ_η is no longer uniform for the SCD model. Also in common with the C parameterisation, the highest IF values for the NCL parameterisation are associated with σ_η.

4.1.3 Non-centred in scale

Table 3 reports the IFs for the MCMC sampler for the non-centred in scale (NCS) parameterisation of the SCD and SV models. In all nine cases for the SCD model, the NCS parameterisation is superior to both the C and NCL parameterisations. Furthermore, in many cases, the gains in efficiency are substantial, with the maximum IF being up to 88% lower than the corresponding figure for the C parameterisation, and 90% lower than the corresponding figure for the NCL parameterisation. Thus, for the SCD model, there is a clear ranking of the three alternative parameterisations: 1) NCS; 2) C; 3) NCL, but with the substantial efficiency gains corresponding to the move from NCL (or C) to NCS.

For the SV model the results are qualitatively similar, in that the NCS parameterisation is preferred to both the C and NCL parameterisations, according to the benchmarking criterion being applied. From the results reported in Tables 2 and 3 the IFs for the NCS parameterisation are seen to be smaller than those for the C parameterisation in seven of the nine SV settings, and smaller than the IFs for the NCL parameterisation in five cases. Hence, the NCS parameterisation clearly outranks the other two. It is evident, however, that the gains in efficiency associated with the NCS parameterisation for the SV model are substantially less than for the SCD model, with the largest percentage improvement over the C parameterisation ($\Delta\%$ in Table 3) being only 45%, and

the corresponding improvement over the NCL parameterisation, 51%.

As with the C parameterisation, for both models there remains a positive relationship between the simulation efficiency of the estimates of ϕ and δ_1 and the magnitudes of the true values of ϕ and σ_η. For the relocated scale parameter σ_η, however, this relationship no longer holds uniformly. Also in contrast with both the C and NCL parameterisations, the IF values for σ_η are not uniformly the largest for all parameter settings. For the SCD model in particular, the IFs for σ_η are markedly reduced as a result of the relocation of the scale parameter to the measurement equation.

4.1.4 Non-centred in both location and scale

Table 4 records the IFs of the MCMC sampler for both the SCD and SV models, for the parameterisation that is non-centred in both location and scale (NCLS). For the SCD model the NCLS parameterisation outperforms the NCS parameterisation in six of the nine parameter settings and outperforms the NCL and C parameterisations in all nine parameter settings. In summary then, for the SCD model, the ranking of the four alternative parameterisations is: 1) NCSL; 2) NCS; 3) C; 4) NCL.

For the SV model the NCLS parameterisation produces substantial gains in all of the nine cases over all three other parameterisations. For example, the percentage reductions in IFs produced by the NCLS parametersation, relative to the C parameterisation (Δ% in Table 4) range from 55% to 89%. Overall then, the appropriate ranking for the SV model is: 1) NCSL; 2) NCS; 3) C; 4) NCL; the same ranking as for the SCD model, but with much more substantial gains being produced in the SV case by the dual relocation of the location and scale parameters.

For both models, only the positive relationship between the simulation efficiency of the estimator of ϕ and the magnitude of the true values of ϕ and σ_η remains. In the case of the parameters δ_1 and σ_η, the positive relationship between simulation efficiency and the true values of ϕ and σ_η no longer holds uniformly. Interestingly, the IFs for σ_η are, in general, much more in line with those of the other two parameters, for both models.

4.1.5 Summary of simulation efficiency results

The experiments clearly illustrate that substantial gains in simulation efficiency can be achieved through simple reparameterisation of the state space models. For the empirically important regions of the parameter space for the SCD model the main efficiency gains are to be had by re-locating the scale parameter from the state equation into

the measurement equation, with there being substantial reductions in the IF values in virtually all cases. For this model further gains can be obtained by also relocating the location parameter, although the magnitude of the gain overall is not as great as that yielded by the relocation of the scale.

For the SV model there is a clear advantage in relocating both the location and scale parameters. Large gains in simulation efficiency, relative to all other parameterisations, are produced via dual relocation for this model. Finally, relocation of either the location or scale parameters or both into the measurement equation tends to diminish (or eliminate) the systematic relationships that hold between the IFs and the magnitudes of the population parameters in the centred formulation.

Importantly, for both models and for all parameterisations, the acceptance rates for the MH algorithms for the state parameters (whether re-located or not) are similar and consistently high, typically above 95%. That is, the efficiency differences documented above can be attributed to the re-parameterisations themselves and not to differential acceptance rates for the MH algorithms used under the various parameterisations.

An interesting feature of the results for both models is the overall tendency of the IFs to decrease, for fixed signal noise σ_η^2, as the true underlying value of ϕ increases. This result hold uniformly for the draws of all three parameters in the centred parameterisation, and for the draws of ϕ across all parameterisations. This result may seem somewhat surprising, as it contrasts with the theoretical findings of Pitt and Shephard (1999) and with the simulation-based findings of Frühwirth-Schnatter (2004), both of these studies finding the opposite result in the linear Gaussian state space context. Pitt and Shephard (1999) show theoretically that the speed of convergence of a Gibbs sampling algorithm for the state in the centred parameterisation is a function of the level of persistence in the state, ϕ, and the signal to noise ratio only, with the speed of convergence being lower the higher the true value of $|\phi|$, for a fixed signal to noise ratio. It is not transparent, however, how these results carry over to the non-Gaussian state space model. Not only is it not clear that the same formulae would apply, but even if they did, in this context it is evident that the variation in the measurement equation (used to define the signal to noise ratio) is itself a non-linear function of both ϕ and σ_η. That is, the role played by ϕ in the non-Gaussian case is different from the role played in the Gaussian case and, hence, the relationship between efficiency and ϕ is not expected to be the same in the two cases. Further investigation of this type of issue is left for future research.

5 Empirical Evaluation

In this section four empirical data sets are examined, with a view to ascertaining the extent to which the results in the previous section hold when the data are not artificially generated. Regressors are also reintroduced into the analysis in the case of the SCD model.

For the empirical analysis, priors reflecting those used in the literature are specified. Uniform priors over $\mathbb{R}$ and $\mathbb{R}^k$ are assumed for δ_1 and $\boldsymbol{\delta}_2$, respectively. However, in contrast to the uniform prior specified for ϕ in the simulation exercise, the prior for ϕ is now derived from a beta distribution that has been stretched over the $(-1, 1)$ interval, with density,

$$p(\phi) \propto \left\{\frac{1+\phi}{2}\right\}^{(\phi_1-1)} \left\{\frac{1-\phi}{2}\right\}^{(\phi_2-1)},$$

and with hyper-parameters ϕ_1, $\phi_2 > 0.5$. The hyperparameters ϕ_1 and ϕ_2 are set to 15 and 1.5, respectively, implying a prior mean of 0.82 and variance of 0.02 for ϕ. As before, an *a priori* inverted-gamma distribution is assumed for σ_η, such that $p(\sigma_\eta) \sim IG\left(\frac{\sigma_r}{2}, \frac{S_\sigma}{2}\right)$, but here with $\sigma_r = 3$ and $S_\sigma = 0.03$, implying a prior mean of 0.14 and prior variance of 0.01 for σ_η. The overall prior specification is similar to that used in several papers in the high frequency financial data literature, including Kim, Shephard and Chib (1998) and SFM.

For the initial set of empirical results for the SCD and SV models, reported in Section 5.1 and 5.2 respectively, the average block size for the state vector is 36. Results reported in Section 5.3 show that the empirical results are reasonably robust to changes in this value.

5.1 SCD Model

5.1.1 Data

The alternative parameterisations of the SCD model are estimated using trade durations data for two Australian listed companies: Broken Hill Proprietary Limited (BHP) and News Corporation (NCP), for the month of August 2001. Following Engle and Russell (1998) and SFM, only distinct trades between 10:20 a.m. and 4:00 p.m. are used, leaving $T = 27746$ observations for BHP and $T = 13832$ observations for NCP. The intraday pattern is modelled using a cubic smoothing spline, $g(y_t)$, estimated using the 'fields'

package in 'R'. The adjusted durations are calculated as

$$\widehat{y_t} = \frac{y_t}{g\left(y_t\right)}. \tag{28}$$

Following SFM, trading volume is included as a regressor, with coefficient denoted by δ_2, included in the state equation. In addition, and similar to Zhang, Russell and Tsay (2001), an additional regressor is defined, with coefficient denoted by δ_3, and value equal to the number of distinct trades occurring simultaneously. For both regressors, the intraday pattern is removed assuming the same type of relationship as used to adjust the duration series in (28).

5.1.2 Empirical results

Table 5 reports estimates of the marginal posterior mean of each parameter of the SCD model, and associated IF, for each data set: BHP (Panel A) and NCP (Panel B). The algorithm for each parameterisation of the SCD model is run for 100000 iterations with a burn-in period of 20000 iterations. The (average) time taken (in seconds) for 1000 iterations is reported at the bottom of each panel. For each data set, IFs are reported for the C, NCL, NCS and NCLS parameterisations.

For both data sets, the algorithms for the NCS and NCLS parameterisations perform substantially better than the C and NCL parameterisations. For the NCP data set relocation of the scale parameter produces maximum IFs (across parameters) that are approximately half the size of those associated with the alternative parameterisations in which the scale parameter remains in the state equation. For the BHP data set the maximum IFs are approximately one fifth the size of those that prevail for the C/NCL parameterisations. The superior performance of the NCS/NCLS parameterisations, as well as their overall similarity one to the other, mimics the simulation results. This is despite the fact that the empirical results are based on an SCD model with regressors in the state equation, as well as being based on much larger sample sizes.

Figure 1 contains plots of the autocorrelation function (ACF) associated with the MCMC iterates of σ_η, for both the BHP and NCP data sets. The ACFs are calculated from the MCMC output for σ_η and visually confirm the results captured by the IFs.

This figure contains plots of the autocorrelation function (ACF) for the BHP (left) and NCP (right) data sets. The ACF estimates are calculated using the MCMC output for the scale parameter in both cases.

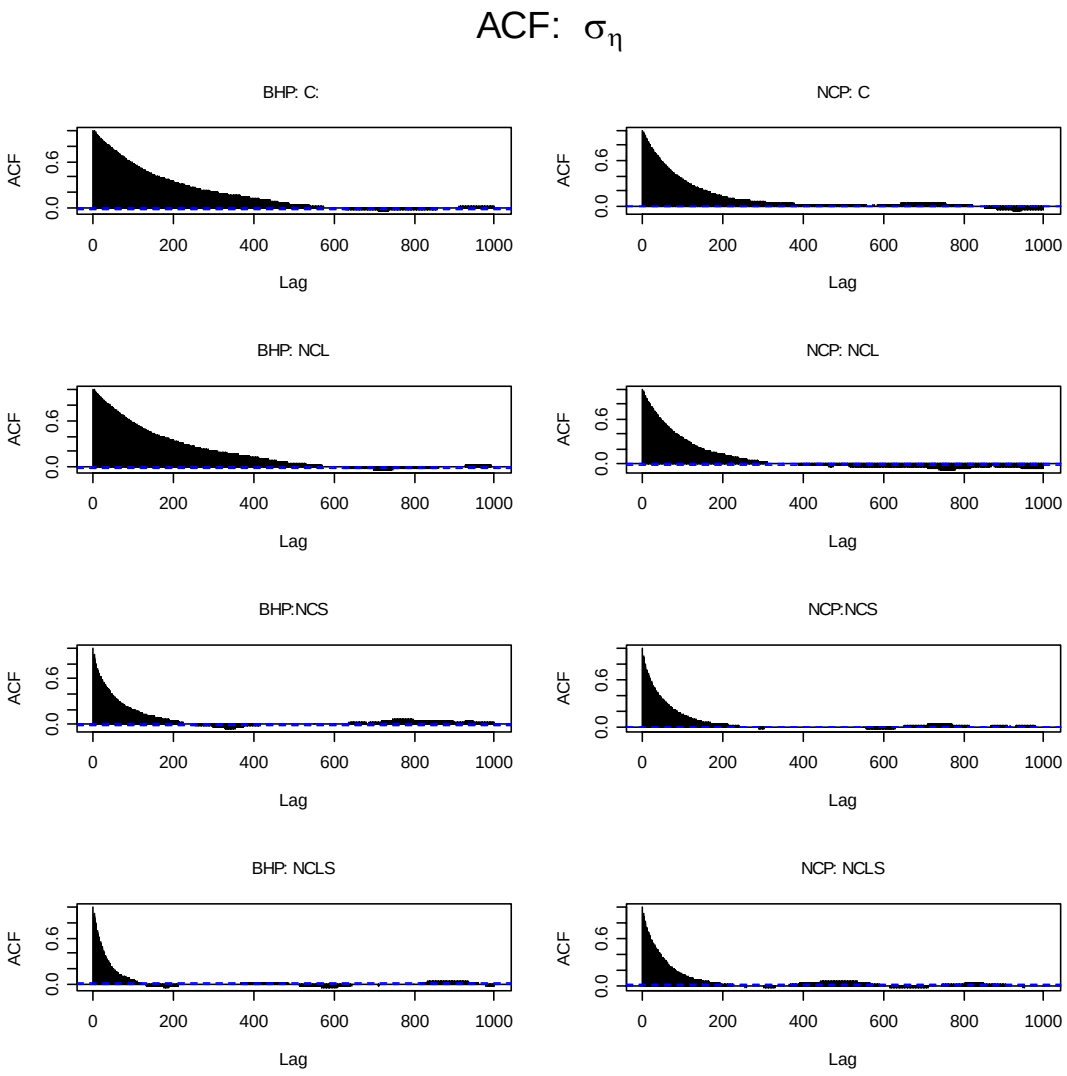

Figure 1:

5.2 SV model

5.2.1 Data

Estimation of the SV model is illustrated using two sets of returns data. The first data set comprises observations on the pound/dollar daily exchange rates from 1 October 1981 to 28 June 1985. This data has been used by Harvey, Ruiz and Shephard (1994), Kim, Shephard and Chib (1998) and Durbin and Koopman (2001) to illustrate their alternative estimation methodologies for the SV model. The second data set comprises observations on the Morgan Stanley Capital Index (MSCI) from 29 December 1989 to 31 May 2002. For the exchange rate series $T = 945$, while $T = 3240$ for the MSCI series.

5.2.2 Empirical results

Table 6 reports estimates of the marginal posterior mean of each parameter of the SV model, and associated IF, for each set of data. Panel A refers to the pound/dollar exchange rate data whilst Panel B contains the output for the MSCI data set. As with the SCD model, the algorithm for each parameterisation of the SV model is run for 100000 iterations with a burn-in period of 20000 iterations. The (average) time taken (in seconds) for 1000 iterations is reported at the bottom of each panel. Results are reported for all four parameterisations for each data set.

As in the simulation study, the NCLS parameterisation clearly produces the most efficient algorithms for estimating the SV model, in both empirical settings. Also consistent with the simulation results, the NCS parameterisation outperforms both the NCL and C parameterisations, whilst the latter produce rather similar results.

Figure 2 contains plots of the ACF for both the Pound and NCP data sets. The ACFs are calculated from the MCMC output for σ_η and, as with the corresponding graphs for the durations data, visually confirm the results captured by the IFs.

An interesting comparison is with the results of Kim, Shephard and Chib (1998), where the latter authors use a model-specific algorithm to estimate the SV model, using the same Pound/Dollar data set as used here. Given the flexible nature of our algorithm, it being deliberately designed to cater for the general state space form specified in (1) to (3), we would anticipate a reduction in simulation efficiency, relative to an algorithm custom-made for a particular model. Comparing our results with the IFs from Table 7 in Kim, Shephard and Chib, it can be seen that the algorithm used in the latter paper is less than four times more efficient than our NCLS parameterisation, compared with being about 16 times more efficient than the C parameterisation. This is a very

This figure contains plots of the sample autocorrelation function (ACF) associated with the MCMC iterates of σ_η, for the SCD model estimated with the Pound\Dollar (left) and MSCI (right) data sets.

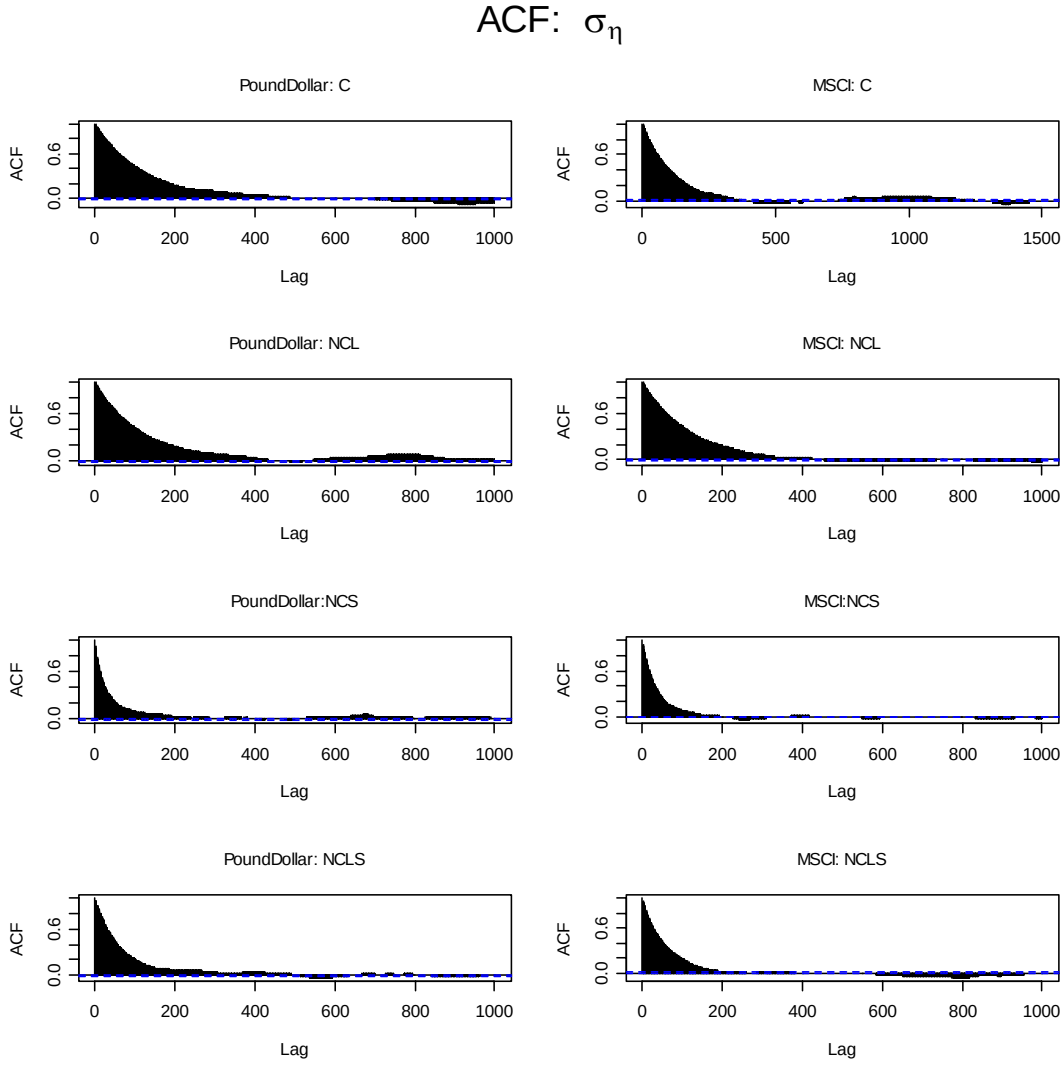

Figure 2:

promising result as the cost, in terms of simulation efficiency, of using a general algorithm in preference to a model specific algorithm, through simple reparameterisation, has been substantially reduced.

5.3 Robustness to Block Size

Table 7 reports summary statistics related to IFs for the SCD model, estimated using the BHP data set. The SCD model is estimated with nine different values assigned to the average block size of the state vector, namely: 26, 36, ... 96, 106. The summary statistics reported are, the minimum, mean, maximum and coefficient of variation (CV) of the IFs, calculated from the simulation output based on all nine (average) block sizes. The NCS and NCLS results, for all parameters, are very robust to the choice of block size, with little deviation between the highest and lowest values, and CV values that range from only 0.03 to 0.05. This is in contrast with the C and NCL parameterisations, for which both the range and CV values are often larger, most notably for the scale parameter σ_η. This is an important result, as it is indicates that little effort need be made by the practitioner when setting the average block size (the only tuning parameter in the algorithm) as long as the scale parameter is relocated to the measurement equation prior to implementation of the algorithm. Comparable results for the SV estimates, based on the Pound/Dollar data set, are reported in Table 8. As is consistent with both the simulation results and the empirical results reported earlier, the NCLS parameterisation outperforms the others, in terms of robustness of the largest IFs (associated with σ_η) to the block size.

Importantly, for both models and all four data sets, the basic ranking of the parameterisations appears to be invariant to the block size. For the SCD model the IFs for the NCS and NCLS parameterisations are always substantially lower than those for the C and NCL parameterisations; whilst for the SV model it is the NCLS parameterisation that dominates. It is worth remembering, when using the range of IF values reported in Tables 7 and 8 to rank the parameterisations, that it is only the largest IF for any parameterisation that is relevant. For example, for the C and NCL results in Table 7, it is only the range of IFs for σ_η that are relevant, whilst for the NCS and NCSL results in that table, the range of IF's for both ϕ and σ_η could produce the largest value and need to be considered when establishing the ranking.

6 Conclusion

Previous studies have documented the fact that substantial gains in the simulation efficiency of MCMC algorithms may be obtained through the use of simple reparameterisations. This paper has contributed to this literature by focussing on the impact of reparameterisation in the context of two non-Gaussian state space models that feature in empirical finance, namely the SCD and SV models. A further contribution of the paper is the presentation of the modifications to the general algorithm of SFM required to accommodate the alternative parameterisations of the non-Gaussian state space specification.

Simulated and empirical data are used to explore the impact of parameterisation on simulation efficiency. The parameter settings used in the simulation experiment are representative of parameter estimates appearing in both the existing empirical literature and in the empirical analysis conducted herein. Four different parameterisations are examined in relation to each of the SCD and SV models, with the impact of reparameterisation measured using inefficiency factors.

For all parameterisations, the simulation results indicate that systematic patterns exist between the efficiency of simulation estimators and the true value of certain of the parameters, with these patterns being most marked for the centred parameterisation. Most importantly, the experimental results reveal that for the two models considered, and for the empirically relevant parameter ranges explored, the most substantial gains in simulation efficiency are produced by moving either the scale parameter or both the location and scale parameters from the state equation to the measurement equation. For the SCD model in particular, relocation of the scale parameter produces a marked increase in efficiency, with small extra gains being produced, in several cases, by the additional relocation of the location parameter. For the SV model, dual relocation of the location and scale parameters produces substantial gains for all parameter settings. The empirical results mimic the simulation results, with the ranking of the parameterisations for the empirical data being robust to the average block size used to draw the state values. Certainly, the main conclusion to be drawn from both the simulation and empirical results is that the dual relocation parameterisation (NCLS) is the safest choice for practitioners applying these particular non-Gaussian models to typical data sets.

Acknowledgements

This research has been supported by Australian Research Council Discovery Grant No. DP0664121. The authors would like to thank a co-editor and two referees for very constructive comments on an earlier draft of the paper. Note that most of the numerical results in the paper are produced using C++ run on a Pentium core 2 duo, 2.4GHz computer. The C++ code also makes use of the Template Numerical Toolkit which can be found at http://math.nist.gov/tnt/index.html. Software for the random number generator used can be found at http://www.agner.org /random/randomc.htm. The ACF plots and the Raftery and Lewis (1992) diagnostic test were computed using the 'R' software package.

References

[1] Andrews, D. (1991), "Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817-858.

[2] Bauwens, L. and Veredas, D. (2004), "The Stochastic Conditional Duration Model: A Latent Variable Model for the Analysis of Financial Durations," *Journal of Econometrics,* 199, 381-412.

[3] Bos, C. and Shephard, N. (2006), "Inference for Adaptive Time Series Models: Stochastic Volatility and Conditionally Gaussian State Space Form," *Econometric Reviews*, 25, 279-244.

[4] Carter, C. and Kohn, R. (1994), "On Gibbs Sampling for State Space Models," *Biometrika*, 81, 541-553.

[5] Chib, S. and Greenberg, E. (1996), "Markov Chain Monte Carlo Simulation Methods in Econometrics," *Econometric Theory*, 12, 407-431.

[6] Durbin, J. and Koopman, S. (2000), "Time Series Analysis of Non-Gaussian Observations Based on State Space Models from both Classical and Bayesian Perspectives," *Journal of the Royal Statistical Society Ser. B*, 62, 3-56.

[7] Durbin, J. and Koopman, S. (2001), *Time Series Analysis by State Space Methods*, Oxford University Press.

[8] Engle, R. and Russell, J. (1998), "Autoregressive Conditional Duration: A New Approach for Irregularly Spaced Transaction Data," *Econometrica*, 66, 987-1007.

[9] Frühwirth-Schnatter, S. (2004), "Efficient Bayesian Parameter Estimation for State Space Models Based on Reparameterisations," *State Space and Unobserved Component Models: Theory and Applications*, Cambridge University Press.

[10] Frühwirth-Schnatter, S. and Sögner, L. (2003), "Bayesian Estimation of the Heston Stochastic Volatility Model," *Working paper*, Johannes Kepler University Linz.

[11] Gelfland, A., Sahu, S. and Carlin, B. (1995), "Efficient Parameterisations for Normal Linear Mixed Models," *Biometrika*, 479-488.

[12] Harvey, A., Ruiz, E. and Shephard, N. (1994), "Multivariate Stochastic Variance Models," *Review of Economic Studies*, 61, 247-264.

[13] Jacquier, E., Polson, N. and Rossi, P. (1994), "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business and Economic Statistics,* 12, 69-87.

[14] Jung, R., Kukuk, M. and Liesenfeld, R. (2006), "Time Series of Count Data: Modeling, Estimation and Diagnostics," *Computational Statistics and Data Analysis*, 51, 2350-2364.

[15] Kim, S., Shephard, N. and Chib, S. (1998), "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, 65, 361-393.

[16] Liesenfeld, R. and Richard, J-F. (2003), "Univariate and Multivariate Stochastic Volatility Models: Estimation and Diagnostics," *Journal of Empirical Finance*, 10, 505-531.

[17] Papaspiliopoulos, O., Roberts, G. and Sköld, M. (2003), "Non-centred Parameterisations for Hierarchical Models and Data Augmentation," *Bayesian Statistics*, 7, 307-326. Oxford University Press.

[18] Pitt, M. (2000), "Discussion of Durbin and Koopman (2000)," *Journal of the Royal Statistical Society Ser. B*, 62, 30-32.

[19] Pitt, M. and Shephard, N. (1999), "Analytic Convergence Rates and Parameterisation Issues for the Gibbs Sampler Applied to State Space Models," *Journal of Time Series Analysis*, 20, 63-85.

[20] Raftery, A., L. and Lewis, S. (1992), "How Many Iterations in the Gibbs Sampler?" *In Bayesian Statistics 4, (ed. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith)*, Oxford University Press, Oxford.

[21] Robert, C. and Mengersen, K. (1999), "Reparameterisation Issues in Mixture Modelling and Their Bearing on MCMC Algorithms," *Computational Statistics and Data Analysis*, 29, 325-343.

[22] Roberts, G., Papaspiliopoulos, O. and Dellaportas, P. (2004), "Bayesian Inference for Non-Gaussian Ornstein-Uhlenbeck Stochastic Volatility Processes," *Journal of the Royal Statistical Society Ser. B*, 66, 369-393.

[23] Roberts, G. and Sahu, S. (1997), "Updating Schemes, Correlation Structure, Blocking and Parameterisation for the Gibbs Sampler," *Journal of the Royal Statistical Society Ser. B,* 59, 291-317.

[24] Shephard, N. (1994), "Partial Non-Gaussian State Space," *Biometrika*, 81, 115-131.

[25] Shephard, N. and Pitt, M. (1997), "Likelihood Analysis of Non-Gaussian Measurement Time Series," *Biometrika*, 84, 653-667.

[26] Strickland, C., Forbes, C. and Martin, G. (2006), "Bayesian Analysis of the Stochastic Conditional Duration Model," *Computational Statistics and Data Analysis*, *Special Issue on Statistical Signal Extraction and Filtering,* 50, 2247-2267.

[27] Zellner, A. (1996), "An Introduction to Bayesian Inference in Econometrics," *John Wiley and Sons Inc.*, New York.

[28] Zhang, M., Russell, J. and Tsay, R. (2001), "A Nonlinear Autoregressive Conditional Duration Model with Applications to Financial Transaction Data," *Journal of Econometrics*, 104, 179-207.

Inefficiency Factors (IF) for the Non-Centred in Location (NCL) Parameterisation

Panel A: SCD Model					Panel B: SV Model				
True σ_η :		*0.1*	*0.2*	*0.3*	*True* σ_η :		*0.2*	*0.3*	*0.4*
True ϕ		IF			*True* ϕ		IF		
0.95	ϕ	39	13	7	*0.95*	ϕ	27	15	9
	σ_η	**273**	**76**	**47**		σ_η	**175**	**93**	**60**
	δ_1	25	31	29		δ_1	29	19	16
	$(\Delta\%)^{(a)}$	*10*	*-1*	*-4*		$(\Delta\%)$	*5*	*0*	*-18*
0.9	ϕ	104	30	16	*0.9*	ϕ	107	36	21
	σ_η	**572**	**120**	**65**		σ_η	**502**	**152**	**89**
	δ_1	19	23	29		δ_1	105	37	25
	$(\Delta\%)$	*21*	*-13*	*-8*		$(\Delta\%)$	*13*	*-5*	*-18*
0.8	ϕ	467	86	45	*0.8*	ϕ	407	129	60
	σ_η	**893**	**321**	**127**		σ_η	**583**	**390**	**183**
	δ_1	21	20	25		δ_1	330	120	59
	$(\Delta\%)$	*0*	*14*	*6*		$(\Delta\%)$	*-38*	*27*	*25*
Time (seconds)	8				**Time** (seconds)	9			

(a) $\Delta\%$ denotes the percentage change in the maximum IF for a particular parameter setting in the NCL parameterisation, relative to the C parameterisation.

Table 2: NCL parameterisation simulation results for the SCD model (Panel A) and the SV model (Panel B). The top row of each panel contains the true values (in italics) of the scale parameter, while the first column contains the true values (in italics) of the persistence parameter. The second column in each panel contains the parameters to which IFs (reported in the subsequent columns) refer, with the maximum IF for each parameter setting indicated in bold. In addition, the percentage change in maximum IF for each parameter setting, relative to the C parameterisation, is given (in italics) in the row immediately below. The average time (across parameter settings) to obtain 1000 iterations is reported in the bottom row of each panel.

Inefficiency Factors (IF) for the Non-Centred in Scale (NCS) Parameterisation

Panel A: SCD Model					Panel B: SV Model				
True σ_η :		*0.1*	*0.2*	*0.3*	*True* σ_η :		*0.2*	*0.3*	*0.4*
True ϕ		IF			*True* ϕ		IF		
0.95	ϕ	30	14	8	*0.95*	ϕ	27	15	10
	σ_η	**37**	**37**	**45**		σ_η	**129**	**97**	**116**
	δ_1	7	4	3		δ_1	22	11	8
	$(\Delta\%)^{(a)}$	*-85*	*-52*	*-8*		$(\Delta\%)$	*-23*	*4*	*58*
0.9	ϕ	**59**	26	17	*0.9*	ϕ	90	36	23
	σ_η	50	**34**	**32**		σ_η	**245**	**112**	**96**
	δ_1	14	7	5		δ_1	78	29	18
	$(\Delta\%)$	*-88*	*-75*	*-55*		$(\Delta\%)$	*-45*	*-30*	*-12*
0.8	ϕ	**368**	**57**	35	*0.8*	ϕ	**633**	131	59
	σ_η	96	50	**36**		σ_η	600	**230**	**119**
	δ_1	55	17	11		δ_1	511	111	49
	$(\Delta\%)$	*-59*	*-80*	*-70*		$(\Delta\%)$	*-32*	*-25*	*-18*
Time (seconds)		8			**Time** (seconds)		9		

(a) $\Delta\%$ denotes the percentage change in the maximum IF for a particular parameter setting in the NCS parameterisation, relative to the C parameterisation.

Table 3: NCS parameterisation simulation results for the SCD model (Panel A) and the SV model (Panel B). The top row of each panel contains the true values (in italics) of the scale parameter, while the first column contains the true values (in italics) of the persistence parameter. The second column in each panel contains the parameters to which IFs (reported in the subsequent columns) refer, with the maximum IF for each parameter setting indicated in bold. In addition, the percentage change in maximum IF for each parameter setting, relative to the C parameterisation, is given (in italics) in the row immediately below. The average time (across parameter settings) to obtain 1000 iterations is reported in the bottom row of each panel.

Inefficiency Factors (IF) of the Non-Centred in Both Location and Scale (NCLS) Parameterisation

Panel A: SCD Model					Panel B: SV Model				
True σ_η :		*0.1*	*0.2*	*0.3*	*True* σ_η :		*0.2*	*0.3*	*0.4*
True ϕ		IF			*True* ϕ		IF		
0.95	ϕ	**56**	12	8	*0.95*	ϕ	19	12	8
	σ_η	31	**28**	**35**		σ_η	**29**	**31**	**33**
	δ_1	18	27	26		δ_1	19	15	13
	$(\Delta\%)^{(a)}$	*-78*	*-64*	*-29*		$(\Delta\%)$	*-83*	*-67*	*-55*
0.9	ϕ	**62**	23	15	*0.9*	ϕ	**48**	26	18
	σ_η	48	**29**	**28**		σ_η	40	**32**	**30**
	δ_1	17	18	24		δ_1	47	26	19
	$(\Delta\%)$	*-87*	*-79*	*-61*		$(\Delta\%)$	*-89*	*-80*	*-72*
0.8	ϕ	**213**	**59**	32	*0.8*	ϕ	**249**	**69**	**37**
	σ_η	86	47	**33**		σ_η	54	39	31
	δ_1	18	19	19		δ_1	247	65	35
	$(\Delta\%)$	*-76*	*-79*	*-73*		$(\Delta\%)$	*-73*	*-77*	*-75*
Time (seconds)		8			**Time** (seconds)		9		

(a) $\Delta\%$ denotes the percentage change in the maximum IF for a particular parameter setting in the NCLS parameterisation, relative to the C parameterisation.

Table 4: NCLS parameterisation simulation results for the SCD model (Panel A) and the SV model (Panel B). The top row of each panel contains the true values (in italics) of the scale parameter, while the first column contains the true values (in italics) of the persistence parameter. The second column in each panel contains the parameters to which IFs (reported in the subsequent columns) refer, with the maximum IF for each parameter setting indicated in bold. In addition, the percentage change in maximum IF for each parameter setting, relative to the C parameterisation, is given (in italics) in the row immediately below. The average time (across parameter settings) to obtain 1000 iterations is reported in the bottom row of each panel.

Empirical Application of the SCD Model

Panel A: BHP

		IF			
Parameter	Mean	C	NCL	NCS	NCLS
ϕ	0.88	105	104	48	**74**
σ_η	0.18	**375**	**321**	**52**	72
δ_1	0.02	11	26	10	19
δ_2	0.00	26	25	26	27
δ_3	-0.03	69	67	27	30
Time (secs)		42	46	45	45

Panel B: NCP

		IF			
Parameter	Mean	C	NCL	NCS	NCLS
ϕ	0.66	88	84	**72**	**70**
σ_η	0.57	**157**	**150**	70	65
δ_1	-0.05	19	63	18	59
δ_2	-0.01	11	18	11	18
δ_3	-0.04	13	18	13	17
Time (secs)		21	24	24	24

Table 5: Panel A contains the SCD results for BHP, whilst Panel B contains the results for NCP. For each panel, column 1 reports the relevant parameters, column 2 reports estimated marginal posterior means, and columns 3 - 6 report IFs for the four alternative parameterisations. The maximum IF under each parameterisation is indicated in bold. The average time to obtain 1000 iterations, for each parameterisation, is reported in the bottom row of each panel.

Empirical Application of the SV Model

Panel A: Pound/Dollar

		IF			
Parameter	Mean	C	NCL	NCS	NCLS
ϕ	0.98	31	33	27	28
σ_η	0.16	**223**	**216**	**114**	**55**
δ_1	-0.02	23	45	20	36
Time (seconds)		1.5	1.5	1.5	1.5

Panel B: MSCI

		IF			
Parameter	Mean	C	NCL	NCS	NCLS
ϕ	0.97	32	33	28	28
σ_η	0.21	**195**	**210**	**101**	**64**
δ_1	-0.03	19	40	17	36
Time (seconds)		5	6	6	6

Table 6: Panel A contains the results for the Pound/Dollar exchange rate data, whilst Panel B corresponds to the MSCI data. For each panel, column 1 reports the relevant parameters, column 2 reports estimated marginal posterior means, and columns 3 - 6 report IFs for the four alternative parameterisations. The maximum IF under each parameterisation is indicated in bold. The average time to obtain 1000 iterations, for each parameterisation, is reported in the bottom row of each panel. For the Pound/Dollar data set the result is reported to the nearest half a second rather than to the nearest second, as it was for all other datasets.

Robustness Check Based on Average Block Size

SCD: BHP trade duration data set

Parameter		IF C	IF NCL	IF NCS	IF NCLS
ϕ	Min	81	104	48	63
	Mean	95	112	50	70
	Max	105	123	54	75
	CV	0.09	0.05	0.04	0.05
σ_η	Min	273	293	50	63
	Mean	332	366	53	70
	Max	397	538	58	74
	CV	0.13	0.23	0.04	0.05
δ_1	Min	10	26	10	18
	Mean	11	28	10	19
	Max	11	29	11	20
	CV	0.05	0.03	0.03	0.04
δ_2	Min	25	23	24	24
	Mean	27	26	27	27
	Max	29	28	29	29
	CV	0.05	0.06	0.05	0.05
δ_3	Min	55	67	27	26
	Mean	65	74	29	28
	Max	72	82	30	30
	CV	0.09	0.07	0.03	0.05

Table 7: Summary statistics related to IFs for the SCD model estimated using the BHP data set. The estimation algorithm is repeated under nine different average block size values (26, 36, 46, 56, 66, 76, 86, 96, 106) for the generated state vector. Reported statistics are the minimum (Min), mean (Mean), maximum (Max) and coefficient of variation (CV) of the IFs, calculated from the simulation output based on all nine (average) block size settings.

Robustness Check Based on Average Block Size

SV: Pound/dollar exchange rate data set

Parameter		IF			
		C	NCL	NCS	NCLS
ϕ	Min	29	28	26	27
	Mean	31	33	28	29
	Max	34	37	33	32
	CV	0.05	0.09	0.08	0.07
σ_η	Min	207	176	94	50
	Mean	228	218	114	54
	Max	270	263	140	58
	CV	0.09	0.11	0.12	0.04
δ	Min	21	39	19	35
	Mean	23	44	21	38
	Max	26	51	24	42
	CV	0.07	0.08	0.08	0.06

Table 8: Summary statistics related to IFs for the SV model estimated using the Pound/Dollar data set. The estimation algorithm is repeated under nine different average block size values (26, 36, 46, 56, 66, 76, 86, 96, 106) for the generated state vector. Reported statistics are the minimum (Min), mean (Mean), maximum (Max) and coefficient of variation (CV) of the IFs, calculated from the simulation output based on all nine (average) block size settings.