PREPRINT
Vol. 00, No. 00, Month 200x, 1–14

RESEARCH ARTICLE

Stochastic volatility models for exchange rates and their estimation using quasi-maximum-likelihood methods: an application to the South African Rand

M.V. Kulikova[a] and D.R. Taylor[b]*

[a] *CEMAT, Instituto Superior Técnico, Technical University of Lisbon, Av. Rovisco Pais, 1049-001 Lisboa, Portugal*; [b] *Dept. of Actuarial Science and the African Collaboration for Quantitative Finance & Risk Research, University of Cape Town, Rondebosch 7701, South Africa*

(*xxx, 2012*)

This paper is concerned with the volatility modeling of a set of South African Rand (ZAR) exchange rates. We investigate the Quasi-Maximum Likelihood (QML) estimator based on the Kalman filter and explore how well a choice of stochastic variance (SV) models fit the data. We note that a data set from a developing country is used. The main results are: (1) the SV model parameter estimates are in line with those reported from the analysis of high-frequency data for developed countries; (2) the SV models we considered, along with their corresponding QML estimators, fit the data well; (3) using the range return instead of the absolute return as a volatility proxy produces QML estimates that are both less biased and less variable; (4) although the log range of the ZAR exchange rates has a distribution that is quite far from normal, the corresponding QML estimator has a superior performance when compared to the log absolute return.

Keywords: exchange rates; quasi-maximum likelihood estimation; Kalman filter; stochastic volatility; adaptive filtering

1. Introduction

The widely used term volatility is a central concept in finance as it is usually regarded as a measurement of risk. It plays an important role in almost all financial applications, from portfolio construction and derivative pricing to risk management and policy decisions. Consequently, it has attracted considerable attention, not least during the crisis periods of the last three decades. From a modeling perspective, it is a particularly difficult problem as there is no reasonably accepted canonical model available. This is exacerbated by the application of the model. In risk management applications, a volatility model will often assume a completely different form to one used in derivative pricing and hedging. Furthermore, most derivative pricing models have either estimated future volatility or market implied volatility as an input. Consequently, derivative markets are generally accepted as volatility trading forums. Explicit trading of volatility may have a significant impact on the stability and trading volumes of financial markets.

*Corresponding author: David.Taylor@uct.ac.za

ISSN: 0266-4763 print/ISSN 1360-0532 online

DOI: 10.1080/0266476YYxxxxxxxx
http://www.tandfonline.com

The purpose of this study is three-fold. First, the data set from a country not previously considered in the literature is used. This, in itself, is not necessarily interesting, except that the data set is for a developing country (South Africa). In addition to this, financial time series in South Africa often lack the property of aggregational Gaussianity [1]. Both of these factors may have a significant impact on the performance and choice of model. Second, the stochastic volatility models considered are estimated (calibrated) using a variety of Quasi-Maximum-Likelihood (QML) methods based on the Kalman filter. Third, the use of the log range-return as a proxy for the volatility is investigated within the context of the chosen models and the QML estimator.

Most financial returns series are empirically heteroskedastic, one of the "stylized facts" of finance [1]. As a consequence, the volatility, or variance, is time-varying. Initial attempts to model this effect in discrete-time were based on autoregressive conditional heteroskedasticity (ARCH), introduced by Engle in [2], and later generalized to GARCH by Bollerslev in [3]. In the class of (G)ARCH models, volatility is observable at time $t-1$; see for instance [4]. A powerful alternative to (G)ARCH-based models is the class of stochastic volatility (SV) models proposed by Taylor in [5]. These have gained popularity due to their superiority of fit [6] and have become an indispensable tool in mathematical finance. SV models treat volatility as an unobservable variable and allow for separate error processes for the conditional mean and variance. The basic univariate stochastic volatility model specifies that the conditional volatility follows a log-normal auto-regressive model with innovations that are assumed to be independent of the innovations in the conditional mean equation [7]. An excellent review on SV models can be found in [8–10] and a comparison study with (G)ARCH in [11].

Despite its remarkable advantages, SV modeling poses serious problems since neither the density nor, consequently, the likelihood function exists in closed form. So, while the maximum likelihood estimation is straightforward for (G)ARCH models, this is not the case for SV models since volatility is modeled as an unobservable variable and the series are not conditionally Gaussian. Initially, these problems were enough to deter empirical application. However, practical estimators do exist that are usually variants of the (generalized) method of moments [12] and the Quasi-Maximum Likelihood (QML) estimator (proposed independently by Nelson in [13] and Harvey et al. in [14]). Other likelihood-based estimation methods evaluate the likelihood function either through numerical integration [15] or Monte Carlo integration using either importance sampling [6, 16, 17] or Markov Chain methods [11, 18]. However, as mentioned in [19], the methods are computationally intensive and rely on assumptions that are hard to check in practice, such as the accuracy of numerical integrals and the convergence of simulated Markov chains to their steady state. The QML estimator based on the Kalman Filter (KF) is popular because it is easy to implement and has good finite-sample properties. For the parameter values often found in the empirical analysis of high-frequency financial time series, the QML estimator usually outperforms estimators based on the Generalized Method of Moments (GMM) in terms of efficiency [20]. Excellent surveys of different estimation techniques developed for the class of SV models can be found in [15, 21–23].

This paper is concerned with the QML estimation of stochastic volatility models of ZAR exchange rates. Our foremost concern is that the effectiveness of the QML estimator depends critically on the distribution properties of the returns. Standard SV models based on log-returns are known to be non-Gaussian. A significant new approach based on the log high-low intraday price range was proposed in [19] and applied to exchange rate data in [24]. Apart from other advantages, the log range has a distribution that is very close to normal. This leads to Gaussian state-space models and the QML estimator reduces to the exact maximum likelihood estimator. As a result, it is an effective and attractive estimator for the class of SV models. Previous empirical studies [25, 26] suggests that speculative price changes and rates of return are well described by a uni-modal symmetric distribution with fatter tails and a higher kurtosis than that of the normal distribution. A cursory study of daily log-returns and log range data for ZAR exchange rates shows similar results. The question then concerns the goodness-of-fit of SV models (with an associated QML estimator) to empirical data from a developing market.

The paper is organized as follows. In Section 2 we briefly describe the class of SV models. The corresponding QML estimator is discussed in Section 3 (a practical algorithm is presented in Appendix A). Section 4 presents and discusses the results of the empirical study and Section 5 concludes the paper.

2. Discrete-time SV models

A comprehensive introduction to the statistics of the class of SV models can be found in [19]. We abbreviate their exposition: consider a volatility proxy that is a statistic $f\left(s(t_k, t_{k+1})\right)$ of the continuous sample path $s(t_k, t_{k+1})$ of the log asset price ($s(t) = \ln S(t)$) between times t_k and t_{k+1}. If the statistic is homogeneous in some power γ of volatility, $\sigma(\cdot)$, then

$$f\left(s(t_k, t_{k+1})\right) = \sigma^{\gamma}(t_k) f\left(s^*(t_k, t_{k+1})\right), \tag{1}$$

where $s^*(t_k, t_{k+1})$ denotes the continuous sample path of a *standardized* diffusion generated by the same innovations as $s(t_k, t_{k+1})$, but with $\sigma^*(t_k) = 1$. It follows from (1) that,

$$\ln\left|f\left(s(t_k, t_{k+1})\right)\right| = \gamma \ln \sigma(t_k) + \ln\left|f\left(s^*(t_k, t_{k+1})\right)\right|, \tag{2}$$

from which a linear state-space model can be constructed:

$$\ln\left|f\left(\cdot\right)\right| = \gamma \ln \sigma_{t-1} + \beta + \eta_t, \tag{3}$$

$$\ln \sigma_t = \phi \ln \sigma_{t-1} + \xi_t, \quad |\phi| < 1, \qquad \xi_t \sim \mathcal{N}(0, \sigma_\xi^2), \tag{4}$$

where $\ln\left|f\left(\cdot\right)\right|$ is the log volatility proxy and $\beta = \mathbf{E}\left\{\ln\left|f\left(S^*(t_k, t_{k+1})\right)\right|\right\}$. The projection errors, η_t, are zero mean but not necessarily Gaussian, and $|\phi| < 1$ ensures stationarity ($\phi = 1$ corresponds to a pure random walk).

The standard volatility proxies, the squared or absolute returns, correspond to $\gamma = 2$ or 1 in (2). Since γ merely scales the log volatility proxy in (3), we can, without loss of generality, choose $\gamma = 1$ ($f\left(\cdot\right) = \ln(S_i(t_{k+1})/S_i(t_k))$). An alternative proxy, the daily log range, was proposed in [19],

$$\begin{aligned} \ln\left|f\left(\cdot\right)\right| &= \ln\left[s_{\text{high}}(t_k) - s_{\text{low}}(t_k)\right] \\ &= \ln\left[\ln(S_{\text{high}}(t_k)/S_{\text{low}}(t_k))\right]. \end{aligned}$$

The daily log range can be shown to be a superior proxy to absolute or squared returns. Its distribution is approximately normal (which can be exploited for the efficient estimation of range-based SV models) and it is robust to microstructure noise. Furthermore [19],

> " . . . the variance of the measurement errors associated with daily log range is far less than the variance of the measurement errors associated with daily log absolute or squared returns, due to the intraday sample path information contained in the range."

In this investigation, we are interested in comparing numerical results between absolute returns and this proxy. We may also note that, in contrast to a (G)ARCH approach, the generalization of SV models to multivariate series is not difficult to estimate and interpret; see [14].

3. Quasi-maximum likelihood estimation of multivariate SV models

The QML estimator was proposed independently by Nelson [13] and Harvey et al. in [14] and is based on an application of the Kalman filter (KF).

Let θ denote the vector of unknown system parameters of a basic linear SV model (3) and (4),

$$\theta = (\phi, \sigma_\xi^2, \sigma_\eta^2, \beta).$$

If these parameters are known *a priori*, then the object is to determine an estimate of the latent dynamic state, $\hat{x}_{t|t-1} = \ln \sigma_t$, which minimizes the expected squared estimation error $\left(x_t - \hat{x}_{t|t-1}\right)^T \left(x_t - \hat{x}_{t|t-1}\right)$[1]. The solution to this problem is known as the Kalman filter – a linear estimator. For Gaussian state-space models, the KF reduces to a minimum mean-square estimate (MMSE) rather than a *linear* MMSE of the unobservable state vector, x_t. For this reason, it is convenient to identify a volatility proxy with a distribution that is approximately normal.

The standard way of solving the problem of uncertain parameters is to use *adaptive filters* where the model parameters, θ, are estimated together with the dynamic state, x_t; see for instance [27]. To construct an adaptive filter we may use different estimation criteria. The method of maximum likelihood is a general method for parameter estimation and is often used in system identification. However, the QML estimator is most efficient when the considered state-space model is Gaussian and, hence, the log likelihood function can be written explicitly; see [28, 29]. This is often not the case. The QML estimator has been extensively used in practice because it is easy to implement and has good finite-sample properties [20].

With the introduction of the log range as a volatility proxy, the QML estimator becomes effective for the class of SV models. The use of the log range leads to (almost) Gaussian state-space models and the QML estimator reduces to the true (or exact) maximum likelihood (ML) estimator. In the Appendix we propose a practical adaptive KF algorithm for the QML estimator.

4. Empirical Results

We now fit SV models to ZAR exchange rate data using the QML estimation method. We consider four exchange rates: Euro/Rand (EUR/ZAR), Dollar/Rand (USD/ZAR), Pound/Rand (GBP/ZAR) and Yen/Rand (JPY/ZAR). The observation period is from July 18, 2005 to August 16, 2010 with a total of 1326 daily observations. We compare log absolute with log range returns,

[1]The estimate $\hat{x}_{t|t-1}$ is called the one-step ahead, predicted estimate of the unobservable dynamic state x_t.

Table 1. Distributions and dynamics of two volatility proxies for four ZAR exchange rates, measured daily from July 18, 2005 to August 16, 2010.

Volatility Proxy	**Unconditional Moments**				**Autocorrelations**				
	Mean	St.Dev.	Skewness	Kurtosis	1st	2nd	5th	10th	20th
			(EUR/ZAR)						
Log range	-4.176	0.465	0.476	3.795	0.511	0.501	0.464	0.384	0.334
Log absolute	-5.374	1.178	-1.148	5.519	0.120	0.109	0.080	0.072	0.041
			(USD/ZAR)						
Log range	-4.095	0.452	0.519	3.507	0.509	0..488	0.457	0.364	0.339
Log absolute	-5.214	1.143	-1.158	5.366	0.085	0.120	0.110	0.089	0.095
			(GBP/ZAR)						
Log range	-4.125	0.447	0.595	4.020	0.503	0.482	0.458	0.374	0.348
Log absolute	-5.326	1.189	-1.107	5.002	0.099	0.089	0.086	0.056	0.055
			(JPY/ZAR)						
Log range	-3.919	0.526	0.848	4.543	0.526	0.474	0.463	0.363	0.324
Log absolute	-4.892	0.935	-1.141	5.538	0.193	0.142	0.152	0.137	0.077

which requires daily closing rates and the daily price range of each exchange rate $\{S_i(t_k), S_i^{\text{high}}(t_k), S_i^{\text{low}}(t_k) : i = 1, \ldots, 4\}$. The log range does not make sense if $S_i^{\text{high}}(t_k) = S_i^{\text{low}}(t_k)$, but this was not detected in our data series.

We present some of the statistics of the data series in Table 1. We reach two initial conclusions: firstly, each returns series exhibits some skewness, and all four samples show moderate to high kurtosis (not excess kurtosis) in comparison to the normal distribution; secondly, we find that the skewness and kurtosis of the log range are less than the same for the absolute returns in all four cases. Finally, the second part of Table 1 illustrates the dynamic of the two proxies. It can be seen that the autocorrelation functions decay slowly. This indicates a relatively slow change in conditional variance.

Thus, it appears that the log range for ZAR exchange rates is not as close to normal as we would have hoped. However, it is clear that the log absolute returns differ substantially from normality for all four samples. This can be confirmed in Fig. 1, where the QQ plots are presented for comparative purposes. For the log range series some outliers can be detected in the tails of the QQ-plots for each exchange rate, with the worst results observed for the Japanese Yen. From Fig. 1 it is clear that the log range is relatively normal in comparison with the log absolute returns.

Next, we calibrate three different SV models using the practical algorithm for the QML estimator presented in the Appendix. All methods were implemented in MatLab with our own code. A Newton-type method was used for optimization purposes with two stopping criteria: $|\mathcal{L}^{(k+1)} - \mathcal{L}^{(k)}| < 10^{-7}$ and $||\theta^{(k+1)} - \theta^{(k)}|| < 10^{-5}$, where $\mathcal{L}$ is the log LF and θ is the vector of unknown system parameters that needs to be estimated. To avoid converging to a local maximum, we followed the standard procedure of initiating the algorithm at several starting points to obtain (as far as possible) a global maximum; see for example [30]. More precisely, for each calibration we used six trials, each starting from different initial values. All 6 trials led to the same final estimates.

Consider the two calibrations performed in Tables 2 and 3. In the first panel (a) of each of the Tables, we have the estimates of the extended 4 parameter AR(1) SV model. We observe that the $\hat{\beta}$ for each exchange rate is very close to the sample mean of the data; compare with the first column of Table 1. This suggests that we should work with a mean-adjusted series and can be confirmed by examining the upper part of Table 4, where the results of the (quasi-)likelihood ratio (LR) test are presented for the log range data. Consider the EUR/ZAR series: the maximum log LF value under the hypothesis of the extended 4 parameter AR(1)

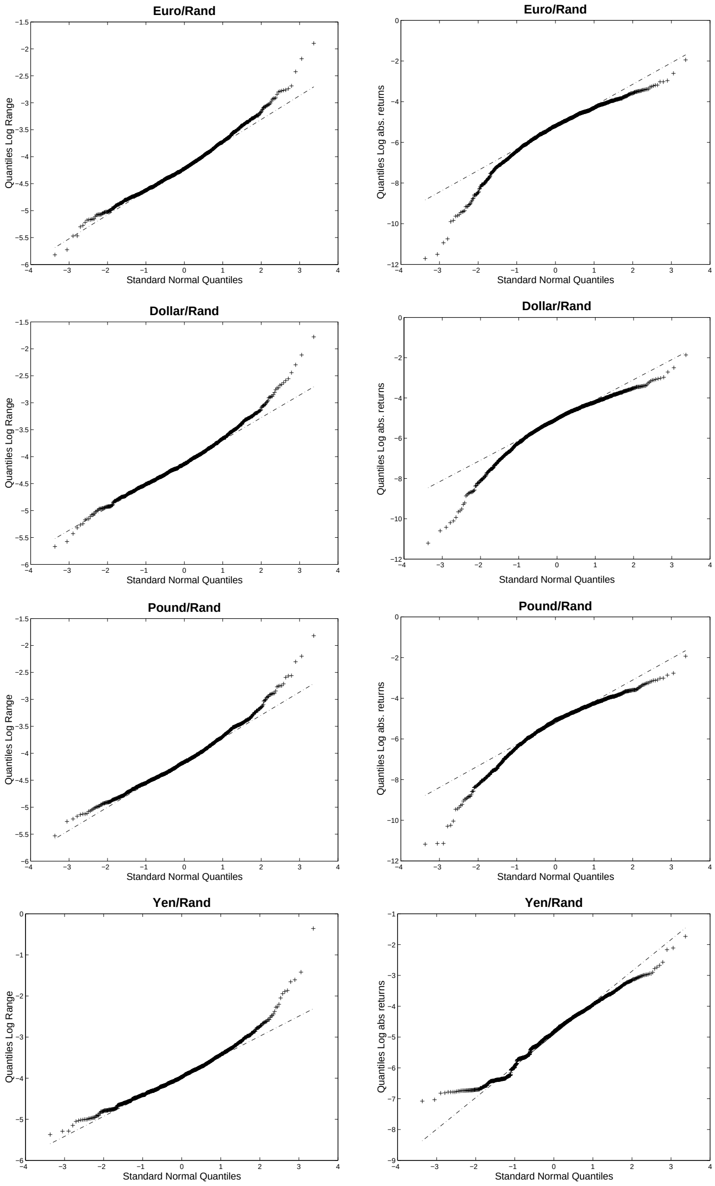

Figure 1. The corresponding QQ plots for the log range (left) and log absolute returns (right) of the four exchange rates.

SV model is 685.3407. This is only fractionally higher than the corresponding maximum log LF value of the 3 parameter AR(1) model, i.e. 685.3404. The number of additional system parameters is 1. As a consequence, the (quasi-)LR statistic $-2(685.3404 - 685.3407) = 0.0006$ should be compared with the $\chi^2_{1,0.99} = 6.63$ value, i.e. the 1% critical value for one degree of freedom, from which we conclude that the 3 parameter AR(1) SV model is no more restrictive than the 4.

Next, we note that in many empirical studies, the estimated values for $\hat{\sigma}^2_\eta$ are usually between 0.01 and 2.77; see the results in [8, 14]. This is in line with our estimates. Also, the parameter ϕ is often very close to or exactly one with very-high-frequency financial time series data. From Tables 2 and 3 we note that our estimates of $\hat{\phi}$ are all close to one. This should imply that a random walk specification would fit almost as well, e.g. see [14]. To confirm this, we estimate an SV model with a random walk specification. The results are summarized in panel (c) of Tables 2 and 3. We then perform a (quasi-)LR test to evaluate whether or not the AR(1)

Table 2. "QML with log absolute returns" – estimation results for three univariate SV models: (a) extended 4 parameter AR(1); (b) AR(1); (c) random walk.

		EUR/ZAR	USD/ZAR	GBP/ZAR	JPY/ZAR
(a)	$\hat{\phi}$	0.9759	0.9923	0.9783	0.9792
	$\hat{\sigma}_\xi^2$	0.0049	0.0021	0.0042	0.0052
	$\hat{\sigma}_\eta^2$	1.2767	1.1731	1.3021	0.7446
	$\hat{\beta}$	-5.3558	-5.2145	-5.3050	-4.8785
	Log $\mathcal{L}$	-858.1433	-797.1912	-868.9903	-516.4142
(b)	$\hat{\phi}$	0.9751	0.9923	0.9781	0.9791
	$\hat{\sigma}_\xi^2$	0.0051	0.0020	0.0042	0.0052
	$\hat{\sigma}_\eta^2$	1.2754	1.1732	1.3020	0.7444
	Log $\mathcal{L}$	-858.1654	-797.1912	-869.0196	-516.4240
(c)	$\hat{\sigma}_\xi^2$	0.0009	0.0013	0.0016	0.0023
	$\hat{\sigma}_\eta^2$	1.3043	1.1787	1.3179	0.7578
	Log $\mathcal{L}$	-862.2617	-799.1592	-873.9565	-521.4376

Table 3. "QML with log range returns" – estimation results for three univariate SV models: (a) extended 4 parameter AR(1); (b) AR(1); (c) random walk.

		EUR/ZAR	USD/ZAR	GBP/ZAR	JPY/ZAR
(a)	$\hat{\phi}$	0.9736	0.9728	0.9790	0.9722
	$\hat{\sigma}_\xi^2$	0.0059	0.0057	0.0042	0.0079
	$\hat{\sigma}_\eta^2$	0.1047	0.0996	0.1000	0.1332
	$\hat{\beta}$	-4.1782	-4.1035	-4.1323	-3.9183
	Log $\mathcal{L}$	685.3407	717.0765	733.0793	521.9541
(b)	$\hat{\phi}$	0.9735	0.9728	0.9790	0.9722
	$\hat{\sigma}_\xi^2$	0.0059	0.0057	0.0042	0.0079
	$\hat{\sigma}_\eta^2$	0.1046	0.0996	0.1000	0.1332
	Log $\mathcal{L}$	685.3404	717.0705	733.0758	521.9540
(c)	$\hat{\sigma}_\xi^2$	0.0043	0.0043	0.0033	0.0060
	$\hat{\sigma}_\eta^2$	0.1073	0.1018	0.1017	0.1362
	Log $\mathcal{L}$	677.8918	709.3750	727.0988	513.9062

model should be rejected in favor of a random walk specification; see the lower part of Table 4. The number of additional system parameters in the AR(1) model is 1. Consequently, the values taken by the (quasi-)LR statistic should be compared with the chi-squared 1% critical value for one degree of freedom, 6.63. From Table 4, it is clear that they are highly significant, and we conclude that the random walk specification is too restrictive for the log range of the exchange rates in this study when compared with the AR(1) SV model. Superiority of the SV model over the random walk (amongst others) is one of the conclusions in [31]. Finally, we note that Ruiz in [20] suggests that with this range of parameter values, there is little doubt about the superior performance of the QML estimator when compared to some GMM estimators.

One of the premier uses for stochastic volatility models is in forecasting. A number of comparative studies have been written on this topic. In [32], the authors write, "This paper investigates the forecasting ability of four different GARCH models and the Kalman filter method (SV models). Measures of forecast errors overwhelmingly support the Kalman filter approach" and conclude that, "Overall, the Kalman filter approach is the best model when forecasted returns are compared with real values. It dominates GARCH models in most cases for different forecast samples." A similar conclusion is reached in Brooks et al. [33] and Faff et al. [34].

Next, we would like to compare the performance of the "QML with log range returns" and the "QML with log absolute returns" estimators. In order to judge the quality of each estimator, we conduct the following set of numerical experiments.

Table 4. Likelihood Ratio (LR) test statistics. The "QML with log range returns" estimator was used to calibrate the models.

	Log LF, model specification			**LR test statistics**	
	4 parameter AR(1)	3 parameter AR(1)	random walk	LR	$\chi^2_{1,0.99}$
EUR/ZAR	685.3407	685.3404	–	0.0006	6.63
USD/ZAR	717.0765	717.0705	–	0.0120	6.63
GBP/ZAR	733.0793	733.0758	–	0.0070	6.63
JPY/ZAR	521.9541	521.9540	–	0.0002	6.63
EUR/ZAR	–	685.3404	677.8918	14.89	6.63
USD/ZAR	–	717.0705	709.3750	15.39	6.63
GBP/ZAR	–	733.0758	727.0988	11.95	6.63
JPY/ZAR	–	521.9540	513.9062	16.09	6.63

Table 5. Performance profile of the two QML estimators.

		QML with absolute returns			**QML with range returns**		
		Mean	RMSE	%MRE	Mean	RMSE	%MRE
EUR/ZAR	ϕ	0.9750	0.0013	0.038	0.9736	0.0008	0.027
	σ^2_ξ	0.0050	0.0018	27.19	0.0058	0.0010	12.90
	σ^2_η	1.2769	0.0508	3.634	0.1046	0.0048	3.276
USD/ZAR	ϕ	0.9923	0.0004	0.012	0.9727	0.0010	0.030
	σ^2_ξ	0.0019	0.0008	30.83	0.0056	0.0009	13.01
	σ^2_η	1.1732	0.0467	3.624	0.0996	0.0045	3.146
GBP/ZAR	ϕ	0.9784	0.0008	0.024	0.9790	0.0008	0.023
	σ^2_ξ	0.0040	0.0015	29.42	0.0042	0.0007	14.68
	σ^2_η	1.3028	0.0524	3.627	0.1001	0.0046	3.211
JPY/ZAR	ϕ	0.9790	0.0011	0.028	0.9722	0.0011	0.027
	σ^2_ξ	0.0051	0.0015	22.22	0.0078	0.0013	13.00
	σ^2_η	0.7439	0.0303	3.636	0.1332	0.0061	3.304

Given the *"true"* value of the system parameters for the AR(1) SV models (taken from panel (b) of Tables 2 and 3), the system is simulated for 1300 samples. Then we use the generated data to solve the inverse problem, i.e. to compute the QML estimates by the two different approaches. We perform 500 Monte Carlo simulations for $T = 1300$ daily observations of the two volatility proxies and report the posterior means for ϕ, $\hat{\sigma}^2_\xi$ and $\hat{\sigma}^2_\eta$. Additionally, the root mean squared error (RMSE) and the percentage relative error (%MRE)[1] are computed by averaging over 1300 observations of the 500 samples. All the results are summarized in Table 5.

Having carefully analyzed the results presented in Table 5, we conclude that the posterior means are all close to the "true" values for both QML estimators and that the RMSE are all small. Hence, these SV models fit the data well. However, using the absolute return as the volatility proxy is less accurate than the alternative range return. For example, for the EUR/ZAR series, the mean square error for the ϕ estimates are 0.0013 and 0.0008, respectively. Overall, the "QML with log absolute returns" estimator performs markedly worse for all four exchange rate samples. Clearly, using the range return as a volatility proxy produces QML estimates that are both less biased and less variable.

5. Conclusion

Our empirical study suggests that, although daily ZAR exchange rates are highly volatile, the QML estimators perform effectively and the SV models we considered

[1]The %MRE is given by $100|\hat{\theta} - \theta^*|/|\theta^*|$, where $\hat{\theta}$ is the calculated estimate for a parameter θ and θ^* is the "true" value.

fit the market data well. There are allied studies of exchange rate volatility for major currencies in Harvey et al. [14], Ruiz [20] and Tims and Mahieu [24]. We found that our SV model parameter estimates are in line with those previously reported, and that are commonly found in the empirical analysis of high-frequency financial time series data for developed countries [14], [20]. Using the range return instead of the absolute return as the volatility proxy produces QML estimates that are both less biased and less variable. Thus, the log range return QML estimator is superior in performance to the log absolute return QML estimator.

It is tempting to infer that these SV models may now be regarded as appropriate for data from developing countries, where currency volatility is often much higher than in developed countries. However, idiosyncrasies of the South African market, such as its proxy as "African exposure", may hinder this. It would be appropriate then to choose a set of data from another developing country, and to test these results. It is clear, though, that the use of the log range return estimator is crucial. What we infer from this is that using the log range return as a proxy for volatility leads to accurate and stable SV models in our context.

Acknowledgements

The first author offers thanks for the support of Portuguese National Fund (*Fundação para a Ciência e a Tecnologia*) within the scope of projects PEst-OE/MAT/UI0822/2011 and SFRH/BPD/64397/2009. The authors gratefully acknowledge the recommendations of two anonymous referees.

References

[1] R. Cont, *Empirical properties of asset returns: Stylized facts and statistical issues*, Quantitative Finance 1 (2001), pp. 223–236.

[2] R. Engle, *Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation*, Econometrica 50 (1982), pp. 987–1008.

[3] T. Bollerslev, *Generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics 31 (1986), pp. 307–327.

[4] T. Bollerslev, R.Y. Chou, and K.F. Kroner, *ARCH modeling in finance: A review of the theory and empirical evidence*, Journal of Econometrics 52 (1992), pp. 5–59.

[5] S.J. Taylor, *Modeling financial time series*, Wiley, Chichester, 1986.

[6] J. Danielsson, *Stochastic volatility in asset prices: Estimation with simulated maximum likelihood*, Journal of Econometrics 64 (1994), p. 375400.

[7] E. Jacquier, N.G. Polson, and P.E. Rossi, *Bayesian analysis of stochastic volatility models with fat-tails and correlated errors*, Journal of Econometrics 122 (2004), pp. 185–212.

[8] S.J. Taylor, *Modeling stochastic volatility: A review and comparative study*, Mathematical Finance 4 (1994), pp. 183–204.

[9] T.G. Andersen and L. Benzoni, *Stochastic volatility*, Encyclopedia of Complexity and Systems Science 19 (2009), pp. 8783–8815.

[10] J. Yu and R. Meyer, *Multivariate stochastic volatility models: Bayesian estimation and model comparison*, Econometric Reviews 25 (2006), pp. 361–384.

[11] S. Kim, N. Shepherd, and S. Chib, *Stochastic volatility: Likelihood inference and comparison with ARCH models*, Review of Economic Studies 65 (1998), pp. 361–393.

[12] T.G. Andersen and B.E. Sorensen, *GMM estimation of a stochastic volatility*

model: A Monte Carlo Study, Journal of Business & Economic Statistics 14 (1996), pp. 328–352.

[13] D.B. Nelson, *The time series behavior of stock market volatility and return. Ph.D. dissertation.*, Massachusetts Institute of Technology, Cambridge, MA, 1988.

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

[15] M. Fridman and L. Harris, *A maximum likelihood approach for non-Gaussian stochastic volatility models*, Journal of Business & Economic Statistics 16 (1998), pp. 284–291.

[16] M.J. Pitt and N. Shephard, *Likelihood analysis of non-Gaussian measurement time series*, Biometrika (1997), pp. 653–667.

[17] J. Durbin and S. Koopman, *Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives*, Journal of the Royal Statistical Society B 62 (2000), pp. 3–56.

[18] E. Jacquier, N.G. Polson, and P.E. Rossi, *Bayesian analysis of stochastic volatility models*, Journal of Business and Economic Statistics (1994), pp. 413–417.

[19] S. Alizadeh, M.W. Brandt, and F.X. Diebold, *Range-based estimation of stochastic volatility models*, Journal of Finance 57 (2002), pp. 1047–1091.

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

[21] C. Broto and E. Ruiz, *Estimation methods for stochastic volatility models: A survey*, Journal of Economic Surveys 18 (2004), pp. 613–649.

[22] E. Capobianco, *State-space stochastic volatility models: A review of estimation algorithms*, Applied Stochastic Models and Data Analysis 12 (1996), pp. 265–279.

[23] B. Jungbacker and S.J. Koopman, *Parameter estimation and practical aspects of modeling stochastic volatility*, Handbook of Financial Time Series (2009), pp. 313–344.

[24] B. Tims and R. Mahieu, *A range-based multivariate stochastic volatility model for exchange rates*, Econometric Reviews 25 (2006), pp. 409–424.

[25] T. Bollerslev, *A conditional heteroskedastic time series model for speculative prices and rates of return*, Review of Economics and Statistics 69 (1987), pp. 542–547.

[26] M.V. Kulikova and D.R. Taylor, *A conditionally heteroskedastic time series model for certain South African stock price returns*, Investment Analysts Journal 72 (2010), pp. 43–52.

[27] S. Särkkä and A. Nummenmaa, *Recursive noise adaptive Kalman filtering by variational Bayesian*, .

[28] F. Schweppe, *Evaluation of likelihood functions for Gaussian signals*, IEEE Transactions on Information Theory 11 (1965), pp. 61–70.

[29] A. Harvey, *Forecasting Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge, 1989.

[30] D. Karlis, *An EM type algorithm for maximum likelihood estimation of the normal-inverse Gaussian distribution*, Statistics & Probability Letters 57 (2002), pp. 43–52.

[31] J. Yu, *Forecasting volatility in the New Zealand stock market*, Applied Financial Economics 12 (2002), pp. 193–202.

[32] T. Choudhry and H. Wu, *Forecasting the weekly time-varying beta of UK firms: GARCH models vs. Kalman filter method*, The European Journal of Finance 15 (2009), pp. 437–444.

[33] R. Brooks, R. Faff, and M. McKenzie, *Time-varying beta risk of Australian industry portfolios: A comparison of modelling techniques*, Australian Journal of Management 23 (1998), pp. 1–22.
[34] R. Faff, D. Hillier, and J. Hillier, *Time varying beta risk: An analysis of alternative modelling techniques*, Journal of Business Finance and Accounting 27 (2000), pp. 523–545.
[35] T. Kailath, A.H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, New Jersey, 2000.
[36] N.K. Gupta and R.K. Mehra, *Computational aspects of maximum likelihood estimation and reduction in sensitivity function calculations*, IEEE Trans. Automat. Contr. AC-19 (1974), pp. 774–783.
[37] G.J. Bierman, M.R. Belzer, J.S. Vandergraft, and D.W. Porter, *Maximum likelihood estimation using square root information filters*, IEEE Trans. Automat. Contr. 35 (1990), pp. 1293–1298.
[38] M.V. Kulikova, *Likelihood gradient evaluation using square-root covariance filters*, IEEE Trans. Automat. Contr. 54 (2009), pp. 646–651.
[39] M.V. Kulikova, *Maximum likelihood estimation via the extended covariance and combined square-root filters*, Mathematics and Computers in Simulation 79 (2009), pp. 1641–1657.

Appendix A. An adaptive Kalman filter for the QML estimation of the state-space models

The QML estimation requires the maximization of the log LF. It leads to implementation of the KF (and its derivatives with respect to unknown system parameters), which is known to be numerically unstable. More recently, array square-root KF algorithms have been found to be particularly efficient because they are more numerically stable than the conventional KF. They improve the robustness of computations with respect to round off errors and are better suited to parallel implementations [35]. We present an efficient and practical adaptive KF algorithm for the QML estimator.

The generalization of SV models to multivariate series is neither difficult to estimate nor interpret. Thus, we consider a multivariate discrete-time linear stochastic system,

$$x_t = F(\theta)x_{t-1} + G(\theta)\xi_t, \quad t = 1, \dots, T, \tag{A1}$$

$$z_t = H(\theta)x_t + \eta_t, \tag{A2}$$

where $x_t \in \mathbb{R}^n$ and $z_t \in \mathbb{R}^m$ are, respectively, the unknown state and the available measurement vectors. For SV models of the form (3) and (4), the vector $x_t = \ln \sigma_t$ is the unobserved log-volatility series and the vector $z_t = \ln |f(\cdot)|$ is the log volatility proxy. The process noise, $\{\xi_t\}$, and the measurement noise, $\{\eta_t\}$, are independent Gaussian white-noise processes, with covariance matrices $Q(\theta) \geq 0$ and $R(\theta) > 0$, respectively. All random variables have known mean values, which we can take, without loss of generality, to be zero. The initial state $x_0 \sim \mathcal{N}(0, \Pi_0(\theta))$. Additionally, system (A1), (A2) is parameterized by a vector of unknown system parameters $\theta \in \mathbb{R}^p$, which needs to be estimated. This means that the entries of $F \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^{n\times q}$, $H \in \mathbb{R}^{m\times n}$, $Q \in \mathbb{R}^{q\times q}$, $R \in \mathbb{R}^{m\times m}$ and $\Pi_0 \in \mathbb{R}^{n\times n}$ are functions of $\theta \in \mathbb{R}^p$.

Solving the problem of parameter estimation by the method of maximum likelihood requires the maximization of a log LF that is often done by using a gradient approach or Newton-type methods [36]. For instance, the scoring equation could be used [37],

$$\hat{\theta}^{(k+1)} = \hat{\theta}^{(k)} - \left(\mathcal{F}|_{\hat{\theta}^{(k)}}\right)^{-1} \left(\left.\frac{\partial \mathcal{L}_\theta}{\partial \theta}\right|_{\hat{\theta}^{(k)}} \right) \tag{A3}$$

where $\mathcal{F}$ is the Fisher information matrix (FIM), $\mathcal{L}_\theta$ is the log LF, $\partial \mathcal{L}_\theta / \partial \theta$ is the gradient vector (score) and $\hat{\theta}^{(k)}$ denotes the value of θ after k iterations of algorithm (A3).

Below we explain how the next cycle for computing $\hat{\theta}^{(k+1)}$ can be obtained by using the scoring equation (A3) and the current approximation, $\hat{\theta}^{(k)}$:

Step 1. Given the current value $\hat{\theta}^{(k)}$, compute $\hat{F} = F\left(\hat{\theta}^{(k)}\right)$, $\hat{G} = G\left(\hat{\theta}^{(k)}\right)$, $\hat{H} = H\left(\hat{\theta}^{(k)}\right)$, $\hat{Q} = Q\left(\hat{\theta}^{(k)}\right)$, $\hat{R} = R\left(\hat{\theta}^{(k)}\right)$ and $\hat{\Pi}_0 = \Pi_0\left(\hat{\theta}^{(k)}\right)$.

Step 2. Use Cholesky decomposition to find $\hat{\Pi}_0^{1/2}$, $\hat{Q}^{1/2}$, $\hat{R}^{1/2}$, which are upper triangular matrices with positive diagonal entries.

Step 3. Set $P_{0|-1}^{1/2} = \Pi_0^{1/2}$ and $P_{0|-1}^{-T/2} \hat{x}_{0|-1} = 0$.

Step 4. For $t = 1, \ldots, T$ do

(1) Given $\hat{P}_{t|t-1}^{1/2}$ and $\hat{P}_{t|t-1}^{-T/2} \hat{x}_{t|t-1}$, recursively update $\hat{P}_{t+1|t}^{1/2}$ and $\hat{P}_{t+1|t}^{-T/2} \hat{x}_{t+1|t}$ as follows,

$$Q_t \left[\begin{array}{cc|c} \hat{R}_t^{1/2} & 0 & -\hat{R}_t^{-T/2} z_t \\ \hat{P}_{t|t-1}^{1/2} \hat{H}_t^T & \hat{P}_{t|t-1}^{1/2} \hat{F}_t^T & \hat{P}_{t|t-1}^{-T/2} \hat{x}_{t|t-1} \\ 0 & \hat{Q}_t^{1/2} \hat{G}_t^T & 0 \end{array}\right] = \left[\begin{array}{cc|c} R_{e,t}^{1/2} & \bar{K}_{p,t}^T & -\bar{e}_t \\ 0 & \hat{P}_{t+1|t}^{1/2} & \hat{P}_{t+1|t}^{-T/2} \hat{x}_{t+1|t} \\ 0 & 0 & \gamma_t \end{array}\right] \tag{A4}$$

where Q_t is any orthogonal rotation that upper-triangularizes the first two (block) columns of the matrix on the left-hand side of (A4) and $\bar{e}_t = R_{e,t}^{-T/2} e_t$ are the normalized innovations of the KF.

(2) Compute the one-step predicted estimate, $\hat{x}_{t+1|t}$, as follows,

$$\hat{x}_{t+1|t} = \left(\hat{P}_{t+1|t}^{1/2}\right)^T \left(\hat{P}_{t+1|t}^{-T/2} \hat{x}_{t+1|t}\right). \tag{A5}$$

We note that no matrices need to be inverted in calculating the state vector. The parentheses are used to indicate the quantities that can be directly read off from (A4).

(3) For each $\theta_i : i = 1, \ldots, p$, apply the orthogonal rotation Q_t from (A4) to the following matrices

$$Q_t \left[\begin{array}{cc|c} \dfrac{\partial \hat{R}_t^{1/2}}{\partial \theta_i} & 0 & \dfrac{\partial\left(-\hat{R}_t^{-T/2} z_t\right)}{\partial \theta_i} \\ \dfrac{\partial\left(\hat{P}_{t|t-1}^{1/2} \hat{H}_t^T\right)}{\partial \theta_i} & \dfrac{\partial\left(\hat{P}_{t|t-1}^{1/2} \hat{F}_t^T\right)}{\partial \theta_i} & \dfrac{\partial\left(\hat{P}_{t|t-1}^{-T/2} \hat{x}_{t|t-1}\right)}{\partial \theta_i} \\ 0 & \dfrac{\partial\left(\hat{Q}_t^{1/2} \hat{G}_t^T\right)}{\partial \theta_i} & 0 \end{array}\right] = \left[\begin{array}{cc|c} X_i & Y_i & M_i \\ N_i & V_i & W_i \\ B_i & K_i & T_i \end{array}\right].$$

During this step, generate and save values as the right-hand side matrix in the equation above.

(4) Calculate for each $\theta_i : i = 1, \dots, p$,

$$\begin{bmatrix} \dfrac{\partial R_{e,t}^{1/2}}{\partial \theta_i} & \dfrac{\partial \bar{K}_{p,t}^T}{\partial \theta_i} \\ 0 & \dfrac{\partial \hat{P}_{t+1|t}^{1/2}}{\partial \theta_i} \end{bmatrix} = \left[\bar{L}_i^T + D_i + \bar{U}_i \right] \begin{bmatrix} R_{e,t}^{1/2} & \bar{K}_{p,t}^T \\ 0 & \hat{P}_{t+1|t}^{1/2} \end{bmatrix}, \tag{A6}$$

$$\begin{bmatrix} -\dfrac{\partial \bar{e}_t}{\partial \theta_i} \\ \dfrac{\partial \left(\hat{P}_{t+1|t}^{-T/2} \hat{x}_{t+1|t} \right)}{\partial \theta_i} \end{bmatrix} = \left[\bar{L}_i^T - \bar{L}_i \right] \begin{bmatrix} -\bar{e}_t \\ \hat{P}_{t+1|t}^{-T/2} \hat{x}_{t+1|t} \end{bmatrix} + \begin{bmatrix} R_{e,t}^{1/2} & \bar{K}_{p,t}^T \\ 0 & \hat{P}_{t+1|t}^{1/2} \end{bmatrix}^{-T} \begin{bmatrix} B_i \\ K_i \end{bmatrix} \gamma_t + \begin{bmatrix} M_i \\ W_i \end{bmatrix} \tag{A7}$$

where $\bar{L}_i$, D_i and $\bar{U}_i$ are strictly lower triangular, diagonal and strictly upper triangular parts of the following matrix product,

$$\begin{bmatrix} X_i & Y_i \\ N_i & V_i \end{bmatrix} \begin{bmatrix} R_{e,t}^{1/2} & \bar{K}_{p,t}^T \\ 0 & \hat{P}_{t+1|t}^{1/2} \end{bmatrix}^{-1} = \bar{L}_i + D_i + \bar{U}_i. \tag{A8}$$

Step 5. Having determined $\bar{e}_t$ and $R_{e,t}^{1/2}$ for each $t = 1, \dots, T$, compute the negative log LF as follows [38],

$$\mathcal{L}_\theta \left(Z_1^T \right) = \frac{1}{2} \sum_{t=1}^{T} \left\{ m \ln(2\pi) + 2 \ln \left(\det R_{e,t}^{1/2} \right) + \bar{e}_t^T \bar{e}_t \right\}, \tag{A9}$$

where $Z_1^T = \{z_1, \dots, z_T\}$ is the T-step measurement history.

Step 6. Having computed $\partial \bar{e}_t / \partial \theta_i$, $\partial R_{e,t}^{1/2} / \partial \theta_i : t = 1, \dots, T$, find the Log LF gradient [38],

$$\frac{\partial \mathcal{L}_\theta \left(Z_1^T \right)}{\partial \theta_i} = \sum_{t=1}^{T} \left\{ \mathbf{tr} \left[R_{e,t}^{-1/2} \cdot \frac{\partial R_{e,t}^{1/2}}{\partial \theta_i} \right] + \bar{e}_t^T \frac{\partial \bar{e}_t}{\partial \theta_i} \right\}, \quad i = 1, \dots, p. \tag{A10}$$

Step 7. The entries of the FIM obey the following equation [37],

$$\mathcal{F}_{i,j} = \sum_{t=1}^{T} \left\{ \mathbf{tr} \left[\mathbf{E} \left\{ \bar{e}_t \frac{\partial \bar{e}_t^T}{\partial \theta_i} \right\} \mathbf{E} \left\{ \bar{e}_t \frac{\partial \bar{e}_t^T}{\partial \theta_j} \right\} \right] + \mathbf{tr}\, \mathbf{E} \left\{ \frac{\partial \bar{e}_t}{\partial \theta_i} \frac{\partial \bar{e}_t^T}{\partial \theta_j} \right\} \right\}, \tag{A11}$$

where $i, j = 1, \dots, p$ and $\mathcal{F}_{i,j}$ denotes the (i, j)-th element of the matrix. The expectation, $\mathbf{E}\{\cdot\}$, is taken over the whole sample space Z_1^T. As mentioned in [37], equation (A11) could also be used by replacing the expected values with sample values, as it is usually done in practice.

Step 8. Having determined the log LF (A9), its gradient (A10) and the elements of the FIM (A11), find the next approximation $\theta^{(k+1)}$ by using the method of scoring (A3).

Remark A1 The detailed derivation of the method presented above, the discussion of its numerical properties and the comparison with the conventional KF approach can be found in [38]. The two-stage method that allows the computation of both one-step ahead predicted estimate $\hat{x}_{t|t-1}$ and the filtered estimate $\hat{x}_{t|t}$ for each $t = 1, \ldots, T$ was proposed in [39]. It could also be effectively used for QML estimation of the state-space models.