

A Closer Look at the Relation between GARCH and Stochastic Autoregressive Volatility

JEFF FLEMING

Rice University

CHRIS KIRBY

University of Texas at Dallas

ABSTRACT

We show that, for three common SARV models, fitting a minimum mean square linear filter is equivalent to fitting a GARCH model. This suggests that GARCH models may be useful for filtering, forecasting, and parameter estimation in stochastic volatility settings. To investigate, we use simulations to evaluate how the three SARV models and their associated GARCH filters perform under controlled conditions and then we use daily currency and equity index returns to evaluate how the models perform in a risk management application. Although the GARCH models produce less precise forecasts than the SARV models in the simulations, it is not clear that the performance differences are large enough to be economically meaningful. Consistent with this view, we find that the GARCH and SARV models perform comparably in tests of conditional value-at-risk estimates using the actual data.

KEYWORDS: GARCH, stochastic volatility, volatility forecasting, value-at-risk, particle filter, Markov chain Monte Carlo.

There is no shortage of research on estimating volatility. Nonetheless, there are key aspects of the relation between the two main classes of volatility models — generalized autoregressive conditional heteroscedasticity (GARCH) and stochastic autoregressive volatility (SARV) — that warrant further investigation. Although GARCH models dominate in terms of popularity, this stems more from computational convenience than anything else. Indeed, theory suggests strong arguments for modeling volatility as stochastic and, if these arguments are valid, then the usual approach to estimating GARCH models is fundamentally misspecified. It is natural to ask, therefore, whether we can reconcile the

We thank Barbara Ostdiek for providing many useful comments on an earlier draft as well as the editor (Eric Renault), an associate editor, three anonymous referees, Tim Bollerslev, Neil Shephard, and seminar participants at the Australian Graduate School of Management. Address correspondence to Jeff Fleming, Jones Graduate School of Management, Rice University, P.O. Box 2932, Houston, TX 77252-2932, or e-mail: jfleming@rice.edu.

widespread use of GARCH models with the theoretical benefits of treating volatility as stochastic.

Of course, this question is more interesting if SARV models actually outperform GARCH models in practice. Thus far the empirical research on SARV models tends to support this hypothesis. Kim, Shephard, and Chib (1998), for example, fit a univariate SARV model to stock index returns and find that it outperforms a widely used GARCH specification. Similarly, Danielsson (1998) fits a multivariate version of the model and finds that it delivers higher likelihood values than several multivariate GARCH alternatives for both currency and index returns. These studies, however, like most research in this area, focus on a particular SARV model — the log-linear specification of Taylor (1986) — and employ statistical rather than economic performance criteria. Thus it is unclear whether the empirical advantage of SARV models holds more generally, or whether it yields tangible economic benefits in applications like portfolio optimization and risk management.

In this article we investigate the empirical performance of GARCH and SARV models from both a statistical and economic perspective. We start by showing that three common discrete-time SARV models — a square-root process for the variance, an AR(1) process for the volatility, and an AR(1) process for the log variance — have linear state-space representations. Once we establish this, it is a simple matter to derive minimum mean square (MMS) linear filters for these models. Our main theoretical contribution is to show that the MMS linear filters are closely related to standard GARCH models. In particular, they imply one-step-ahead forecasting rules for the SARV variance, volatility, and log variance models that are identical to the conditional variance, volatility, and log variance functions of a standard GARCH(1,1) model, an absolute-value GARCH(1,1) model, and a multiplicative GARCH(1,1) model, respectively.

Given the parallels between GARCH models and MMS linear filters, we might expect GARCH models to perform well in forecasting stochastic volatility. However, the most common approach for fitting GARCH models is maximum likelihood. Under a SARV data-generating process, the maximum-likelihood estimator of the GARCH parameters is inconsistent, even when fitting what amounts to an MMS linear filter. The problem is that the likelihood function is misspecified because the GARCH model fails to capture the contribution of the unpredictable component of volatility to the forecast errors. Fortunately the least-squares estimator of the GARCH parameters is consistent, and we can implement it in a straightforward fashion. Thus our analysis of the filtering properties of GARCH models suggests looking at three sets of competing forecasts: one set produced by the SARV models, a second set produced by the GARCH models using least-squares estimates, and a third set produced by the GARCH models using maximum-likelihood estimates.

Before confronting the models with real data, we use simulations to evaluate their forecasting performance under controlled conditions. The results reveal that fitting the GARCH models by least squares is an effective strategy for circumventing the problems with the maximum-likelihood estimator. In fact, the least-squares version of the GARCH volatility model does particularly well. It

outperforms all of the other models, producing one-step-ahead forecasts whose mean absolute errors and root mean squared errors are within 1% to 4% of the values for the optimal SARV forecasting rule, regardless of the data-generating process. In contrast, both the least-squares and maximum-likelihood versions of the GARCH log variance model perform poorly, producing mean absolute errors and root mean squared errors 13% to 19% higher than those for the optimal SARV forecasting rule. Although the differences in performance become less pronounced as the forecast horizon gets longer, the overall ranking of the models shows little change.

Overall the simulations suggest that the loss in precision associated with using GARCH forecasts in a SARV framework can be large or small, depending on the model. But this finding is purely statistical in nature. More generally we would like to know whether the differences in performance between the GARCH and SARV forecasts are large enough to be economically meaningful. To provide insights into this issue, we consider a standard application in risk management—estimating value at risk (VaR). We fit each of the models to the daily returns on five currencies and five equity indexes, and we use the one-step-ahead volatility forecasts from each model to construct conditional VaR estimates. We then subject these VaR estimates to specification tests to see how well they perform in both an absolute and a relative sense.

The results indicate that the GARCH and SARV models fit the data comparably for both the currencies and the equity indexes. However, the VaR specification tests generate some interesting findings. We cannot reject any of the SARV models for the currencies, but we reject both the SARV variance and volatility models for three of the equities. The log variance model delivers the best performance of the SARV models, with only one rejection. For the GARCH models, the rejection rate is lower for the least-squares specifications than for the maximum-likelihood specifications. Among the least-squares specifications, the GARCH variance model performs best, with no rejections, but the GARCH volatility model also performs well, with just two. As in the simulations, both specifications of the GARCH log variance model perform the worst.

We also conduct pairwise comparisons of the VaR estimates using a test statistic that is valid for nonnested and potentially misspecified models. Although neither class of models enjoys a clear advantage overall, we can draw some conclusions about the relative performance of specific models for certain assets or asset classes. Among the SARV models, for example, we can only reject the null of equal performance in two cases. Both of these involve currencies, and in both cases we reject the log variance model in favor of the variance and volatility models. In addition, the least-squares GARCH volatility model continues to perform well relative to the other GARCH models. We see only two rejections (for S&P 500 and FTSE) in favor of the maximum-likelihood variance and volatility models.

These results, together with the simulation evidence, support the use of GARCH models for parameter estimation and volatility forecasting, even if we believe volatility follows one of the SARV processes. Not only is our least-squares estimator of the GARCH parameters consistent, we can convert it into a consistent

estimator of the SARV parameters using the transformation implied by our state-space representation. Since this requires only a small fraction of the computational effort required to obtain fully efficient estimates of the SARV parameters, it should encourage wider use of SARV models in applied work. In addition, the forecasts obtained by least-squares estimation of the GARCH volatility model perform well under each of the SARV data-generating processes. Thus GARCH models formulated in terms of absolute return innovations may provide a robust approximation to SARV dynamics in forecasting applications.

Of course, it is important to remember that our findings are based on a limited set of SARV models. Other models might yield different results. Moreover, we impose several assumptions in our empirical analysis that could influence how the GARCH filters perform relative to the SARV models. In the simulations, for example, we assume a Gaussian distribution for the SARV standardized innovations. Perhaps the performance of the GARCH filters would deteriorate if we consider an alternative distribution, such as a Student's t , that allows for conditionally fat tails. Similarly we focus strictly on single-factor SARV models. Perhaps the GARCH filters for multifactor SARV models would not fare as well in forecasting and VaR applications. In any event, our methodology should prove useful in addressing these and related questions, bringing further reconciliation to the debate over the relative merits of GARCH and SARV models.

The remainder of the article is organized as follows. Section 1 introduces the SARV models, illustrates our approach for constructing MMS linear filters, and shows that these filters have a simple GARCH interpretation. Section 2 describes our approach for assessing the forecasting performance of the models. Section 3 uses simulations to evaluate how the models perform under controlled conditions. Section 4 describes the currency and equity index data, presents the model fitting results, and discusses the performance of the models in our risk management application. The final section offers some concluding remarks.

1 FILTERING AND FORECASTING FOR SARV MODELS

Financial economists use two main classes of discrete-time models to describe the dynamics of volatility: stochastic volatility and GARCH. Stochastic volatility models assume that volatility (or some function of it) follows a low-order Markov process. In effect, these models are standard time-series models applied to a latent stochastic process. GARCH models, on the other hand, assume that volatility (or some function of it) can be expressed as a deterministic function of the lagged squared (or absolute) return innovations. These models provide a parsimonious way to model changes in volatility, but they do not allow for randomness that is specific to the volatility process.

Much of the empirical research on discrete-time stochastic volatility focuses on the log-linear model of Taylor (1986). One of the main reasons for the popularity of this model is that we can estimate it using linear state-space methods. We simply square the returns, take logarithms, and the result is a linear state-space representation that can be analyzed using the Kalman filter [see, e.g., Harvey,

Ruiz, and Shephard (1994) and Fleming, Kirby, and Ostdiek (1998)]. In this section we generalize the Kalman filter approach to cover models in which the volatility or variance follows an autoregressive process and then we use the associated state-space framework to study the relation between linear filters, approximate nonlinear filters, optimal nonlinear filters, and some common GARCH models.

1.1 The SARV Models and the State-Space Framework

Let r_t denote the date t return on an equity index, foreign currency, or some other asset. Suppose that r_t has a stochastic volatility representation of the form

$$r_t = \mu + \delta r_{t-1} + \sigma_t z_t, \quad (1)$$

where z_t is i.i.d. with zero mean and unit variance. In the spirit of Andersen (1994), we consider models in which σ_t or a simple function of σ_t follows a first-order Markov process. Specifically we focus on the following three models: a square-root process for the variance,

$$\sigma_{t+1}^2 = \kappa + \phi \sigma_t^2 + \gamma \sigma_t u_t, \quad (2)$$

an AR(1) process for the volatility,

$$\sigma_{t+1} = \kappa + \phi \sigma_t + \gamma u_t, \quad (3)$$

and an AR(1) process for the log variance,

$$\log \sigma_{t+1}^2 = \kappa + \phi \log \sigma_t^2 + \gamma u_t. \quad (4)$$

For all three models we assume that u_t is i.i.d. with zero mean and unit variance. We also assume that z_t is independent of u_τ for all τ , which rules out leverage effects. Strategies for incorporating leverage are discussed briefly in Section 1.6. There is no need to be more specific about the distributions of z_t and u_t for now. Sometimes it is convenient to assume that u_t is drawn from a distribution with bounded lower support to ensure that the variances and volatilities in Equations (2) and (3) are bounded away from zero.¹ In other cases, such as using a discrete-time model to approximate a SARV diffusion, we need to assume u_t is Gaussian and use some other approach to ensure nonnegativity.

Since models like those in Equations (1)–(4) are often used as diffusion approximations, we should briefly comment on the state of research into the relation between discrete- and continuous-time SARV specifications.² A recent study by Meddahi and Renault (2002) is of particular interest because it develops a general class of square-root SARV (SR-SARV) models that is closed under

¹ Suppose, for instance, that $u_t \geq -1$ for all t . We can show that this implies the variance in Equation (2) and the volatility in Equation (3) are strictly nonnegative provided that $\kappa > \gamma^2/4\phi$ and $\kappa > \gamma$ (this presumes $\phi > 0$ in the first case).

² For example, if we specify Gaussian standardized innovations in Equations (1)–(4), then we obtain first-order Euler discretizations of SARV diffusions that have many applications in derivatives pricing and risk management [see, e.g., Melino and Turnbull (1990), Stein and Stein (1991), and Heston (1993)].

temporal aggregation. It shows, in other words, that (i) the exact discretization of an SR-SARV diffusion produces a discrete-time SR-SARV process, and (ii) if single-period returns are generated by a discrete-time SR-SARV process, then multiperiod returns follow the same process, albeit with different parameter values. These results establish a precise relation between discrete- and continuous-time SR-SARV parameterizations. Although we focus on the performance of GARCH models in forecasting stochastic volatility, we also show how to formulate consistent estimators for discrete-time SARV models using standard econometric methods. In light of the Meddahi and Renault (2002) results, it may be possible to extend our approach to SR-SARV diffusions. This would provide a convenient alternative to the generalized method of moments (GMM) estimator proposed in their article.³

We now turn to developing the linear filters for our three SARV models. To facilitate this, let s_t denote the date t variance, volatility, or log variance, and let y_t denote the associated observation, that is,

$$y_t = \begin{cases} e_t^2 & \text{if } s_t = \sigma_t^2, \\ |e_t| & \text{if } s_t = \sigma_t, \\ \log e_t^2 & \text{if } s_t = \log \sigma_t^2, \end{cases} \quad (5)$$

where $e_t = r_t - \mu - \delta r_{t-1}$ denotes the date t return innovation. We will show that each of the three SARV models has a linear state-space representation (LSSR) of the form

$$s_{t+1} = \eta + \phi(s_t - \eta) + v_t, \quad (6)$$

$$y_t = a + b(s_t - \eta) + w_t, \quad (7)$$

where v_t and w_t are zero mean, serially uncorrelated innovations with

$$E(w_t y_\tau) = 0 \quad \text{for all } \tau < t, \quad (8)$$

$$E(v_t s_\tau) = E(v_t y_\tau) = 0 \quad \text{for all } \tau \leq t, \quad (9)$$

$$E(v_t w_\tau) = E(w_t s_\tau) = 0 \quad \text{for all } \tau. \quad (10)$$

Since we can establish this, it is a relatively simple matter to derive MMS linear filters for the models and use them as the basis for estimation and inference.⁴ First, we illustrate the general approach using the LSSR in Equations (6) and (7). Then we specialize our results to each of the three SARV models.

³ Another approach might be to build on the recent work of Barndorff-Nielsen and Shephard (2002). They assume a relatively general SARV diffusion for instantaneous returns and show that realized volatility, as defined by Anderson et al. (2003), has a state-space representation that implies a GARCH-like recursion for the variance forecasts. We can relate their analysis to ours by noting that realized volatility has squared returns as a special case.

⁴ A similar approach is discussed by Barndorff-Nielsen and Shephard (2001) in the context of estimating Ornstein-Uhlenbeck-based diffusions using discretely observed data.

1.2 Linear Filters

Let $\mathcal{I}_t = \{y_1, y_2, \dots, y_t\}$ denote the date t information set associated with the LSSR. The objective in MMS linear filtering is to construct forecasts of y_{t+1} and s_{t+1} that are linear in the elements of \mathcal{I}_t and have the smallest mean squared error (MSE). The MMS linear forecasts, denoted by $y_{t+1|t}$ and $s_{t+1|t}$, are just the linear projections of y_{t+1} and s_{t+1} on \mathcal{I}_t . Once we have the forecasts for date t , we can compute the forecasts for date $t+1$ using a two-step procedure. First, we update our inference about the value of s_t using the formula for updating a linear projection [see, e.g., Hamilton (1994, eq. 4.5.12)],

$$s_{t|t} = s_{t|t-1} + \frac{E((s_t - s_{t|t-1})(y_t - y_{t|t-1}))}{E((y_t - y_{t|t-1})^2)} (y_t - y_{t|t-1}). \quad (11)$$

Next, we note that the linear projections of v_t and w_{t+1} on \mathcal{I}_t are zero [by Equations (8) and (9)] and use Equations (6) and (7) to obtain the MMS linear forecasts for date $t+1$,

$$s_{t+1|t} = \eta + \phi(s_{t|t} - \eta), \quad (12)$$

$$y_{t+1|t} = a + b(s_{t+1|t} - \eta), \quad (13)$$

The recursions begin with $s_{1|0}$ set equal to η .

To make this algorithm operational, we have to evaluate the unconditional expectations in Equation (11). This turns out to be functionally equivalent to implementing the Kalman filter. To see this, let $R = E(w_t^2)$ denote the variance of the innovation to the observation equation. Since $y_t - y_{t|t-1} = b(s_t - s_{t|t-1}) + w_t$, we have

$$E((s_t - s_{t|t-1})(y_t - y_{t|t-1})) = bP_{t|t-1} \quad (14)$$

and

$$E((y_t - y_{t|t-1})^2) = b^2P_{t|t-1} + R, \quad (15)$$

where $P_{t|t-1}$ denotes the MSE of $s_{t|t-1}$. Note that the cross-product terms drop out of Equations (14) and (15) because Equations (8)–(10) imply that $E((s_t - s_{t|t-1})w_t) = 0$. Combining these results with Equations (11)–(13), we obtain the standard one-step-ahead Kalman filter forecasting equation [see, e.g., Hamilton (1994, eq. 13.2.20)],

$$s_{t+1|t} = \eta + \phi(s_{t|t-1} - \eta) + K_t(y_t - a - b(s_{t|t-1} - \eta)), \quad (16)$$

where

$$K_t = \frac{\phi b P_{t|t-1}}{b^2 P_{t|t-1} + R} \quad (17)$$

denotes the filter gain.⁵

⁵ To compute the value of K_t , we find $P_{t|t-1}$ using the formula for updating the MSE of a linear projection [see, e.g., Hamilton (1994, eq. 4.5.13)]. Once again we obtain the standard Kalman filter recursion, that is, $P_{t|t-1} = \phi^2(P_{t-1|t-2} - b^2 P_{t-1|t-2}^2 (b^2 P_{t-1|t-2} + R)^{-1}) + Q$, where $Q = E(v_t^2)$ and $P_{1|0} = Q(1 - \phi^2)^{-1}$.

To apply these general results we need to specialize Equations (16) and (17) to the SARV variance, volatility, and log variance models. We begin by filling in the details of the LSSR for each model. Following this, we consider the one-step-ahead MMS linear forecasts implied by Equations (16) and (17).

1.2.1 Model 1: SARV variance To obtain the state equation for the first model, let $\eta = \kappa/(1 - \phi)$ denote the unconditional mean of σ_{t+1}^2 and express Equation (2) as

$$\sigma_{t+1}^2 = \eta + \phi(\sigma_t^2 - \eta) + \gamma\sigma_t u_t. \quad (18)$$

To obtain the observation equation, use Equation (1) to express e_t^2 as

$$e_t^2 = \eta + (\sigma_t^2 - \eta) + \sigma_t^2(z_t^2 - 1). \quad (19)$$

Equations (18) and (19) are equivalent to Equations (6) and (7), where $s_t = \sigma_t^2$, $v_t = \gamma\sigma_t u_t$, $y_t = e_t^2$, $a = \eta$, $b = 1$, and $w_t = \sigma_t^2(z_t^2 - 1)$. Since z_t and u_t are mutually independent i.i.d. random variables, it is easy to verify that v_t and w_t are zero mean, serially uncorrelated, and satisfy the restrictions in Equations (8)–(10).

Now let $h_{t+1|t} = s_{t+1|t}$ denote the one-step-ahead MMS linear forecast of the date t variance under Equations (18) and (19). Using Equations (16) and (17), we have

$$h_{t+1|t} = \eta + \phi(h_{t|t-1} - \eta) + K_t(e_t^2 - \eta - (h_{t|t-1} - \eta)), \quad (20)$$

where $K_t = \phi P_{t|t-1}(P_{t|t-1} + R)^{-1}$. Since the filter gain converges to a constant K for a covariance stationary process, we can write the steady-state version of Equation (20) as

$$h_{t+1|t} = \omega + \beta h_{t|t-1} + \alpha e_t^2, \quad (21)$$

where $\omega = \kappa$, $\beta = \phi - K$, and $\alpha = K$. This equation is identical to the conditional variance function implied by Bollerslev's (1986) GARCH(1,1) model.

Note that the innovations to the state and observation equations for Model 1 are conditionally heteroscedastic, a feature not usually found in LSSRs. It might seem surprising at first that we do not need to model this feature to apply the Kalman filter. Recall, however, that the Kalman filter is simply a convenient algorithm for recursively updating the coefficients and MSE of a linear projection. Thus, as in linear regression analysis, conditional heteroscedasticity has no effect on how we compute the unconditional population moments that characterize the filter parameters. It does, of course, affect the efficiency of our estimates of these parameters, but this is a separate issue. Also note that the time variation in the filter gain, which reflects the lack of an observed return history when the filter is initialized, is not influenced by conditional heteroscedasticity in any way. As long as the SARV process is covariance stationary, the filter gain will quickly converge to a fixed value as the information set expands and the impact of the initial conditions dies out.

1.2.2 Model 2: SARV volatility To obtain the state equation for the second model, let $\eta = \kappa/(1 - \phi)$ denote the unconditional mean of σ_t and express Equation (3) as

$$\sigma_{t+1} = \eta + \phi(\sigma_t - \eta) + \gamma u_t. \quad (22)$$

To obtain the observation equation, use Equation (1) to express $|e_t|$ as

$$|e_t| = \eta\psi + \psi(\sigma_t - \eta) + \sigma_t(|z_t| - \psi), \quad (23)$$

where $\psi = E(|z_t|)$. Equations (22) and (23) are equivalent to Equations (6) and (7), where $s_t = \sigma_t$, $v_t = \gamma u_t$, $y_t = |e_t|$, $a = \eta\psi$, $b = \psi$, and $w_t = \sigma_t(|z_t| - \psi)$. Since z_t and u_t are mutually independent i.i.d. random variables, it is easy to verify that v_t and w_t are zero mean, serially uncorrelated, and satisfy the restrictions in Equations (8)–(10). The value of ψ depends on the distribution of z_t . For example, if z_t is normal, then $\psi = 0.7979$ [see Abramowitz and Stegun (1970)].

Now let $h_{t+1|t}^{1/2} = s_{t+1|t}$ denote the one-step-ahead MMS linear forecast of the date t volatility under Equations (22) and (23). Using Equations (16) and (17), we have

$$h_{t+1|t}^{1/2} = \eta + \phi(h_{t|t-1}^{1/2} - \eta) + K_t(|e_t| - \eta\psi - \psi(h_{t|t-1}^{1/2} - \eta)), \quad (24)$$

where $K_t = \phi\psi P_{t|t-1}(\psi^2 P_{t|t-1} + R)^{-1}$. Since the filter gain converges to a constant K for a covariance stationary process, we can write the steady-state version of Equation (24) as

$$h_{t+1|t}^{1/2} = \omega + \beta h_{t|t-1}^{1/2} + \alpha |e_t|, \quad (25)$$

where $\omega = \kappa$, $\beta = \phi - K\psi$, and $\alpha = K$. This equation is identical to the conditional volatility function implied by a generalized version of Schwert's (1990) absolute-value ARCH model.

1.2.3 Model 3: SARV log variance To obtain the state equation for the third model, let $\eta = \kappa/(1 - \phi)$ denote the unconditional mean of $\log \sigma_t^2$ and express Equation (4) as

$$\log \sigma_{t+1}^2 = \eta + \phi(\log \sigma_t^2 - \eta) + \gamma u_t. \quad (26)$$

To obtain the observation equation, use Equation (1) to express $\log e_t^2$ as

$$\log e_t^2 = (\psi + \eta) + (\log \sigma_t^2 - \eta) + (\log z_t^2 - \psi), \quad (27)$$

where $\psi = E(\log z_t^2)$. Equations (26) and (27) are equivalent to Equations (6) and (7), where $s_t = \log \sigma_t^2$, $v_t = \gamma u_t$, $y_t = \log e_t^2$, $a = \psi + \eta$, $b = 1$, and $w_t = \log z_t^2 - \psi$. Since z_t and u_t are mutually independent i.i.d. random variables, it is easy to verify that v_t and w_t are zero mean, serially uncorrelated, and satisfy the restrictions in Equations (8)–(10). The value of ψ depends on the distribution of z_t . For example, if z_t is normal, then $\psi = -1.2704$ [see Abramowitz and Stegun (1970)].

Now let $\log h_{t+1|t} = s_{t+1|t}$ denote the one-step-ahead MMS linear forecast of the date t log variance under Equations (26) and (27). Using Equations (16) and (17), we have

$$\log h_{t+1|t} = \eta + \phi(\log h_{t|t-1} - \eta) + K_t(\log e_t^2 - \psi - \eta - (\log h_{t|t-1} - \eta)), \quad (28)$$

where $K_t = \phi P_{t|t-1} (P_{t|t-1} + R)^{-1}$. Since the filter gain converges to a constant K for a covariance stationary process, we can write the steady-state version of Equation (28) as

$$\log h_{t+1|t} = \omega + \beta \log h_{t|t-1} + \alpha \log e_t^2, \quad (29)$$

where $\omega = \kappa - K\psi$, $\beta = \phi - K$, and $\alpha = K$. This equation is identical to the conditional log variance function implied by the GARCH generalization of Geweke's (1986) multiplicative ARCH model.

1.3 GARCH Models as Linear Filters

Our analysis of LSSRs reveals a straightforward relation between discrete-time SARV models and GARCH models. Specifically, since certain GARCH models have the same basic structure as MMS forecasting rules, we might expect these models to deliver reasonably accurate forecasts in SARV applications. This finding has especially interesting implications in the case of Model 1. Square-root SARV diffusions are popular with theorists because they enforce nonnegativity under simple parametric restrictions and often admit closed-form solutions to portfolio optimization and option pricing problems. The GARCH(1,1) model, on the other hand, is the most popular empirical specification because it performs well in a wide range of applications. Our results provide a link between the two models that could potentially be exploited not only for forecasting, but also for estimation and inference.⁶

Additional insights can be obtained by considering estimation of the filter parameters. The simplest approach would be to use the steady-state forecasting rule for each model to construct a least-squares estimate of the parameter vector $\boldsymbol{\vartheta} = (\omega, \beta, \alpha)'$. To implement this approach, we would find the value of $\boldsymbol{\vartheta}$ that minimizes

$$S_{LS} = \sum_{t=1}^T (y_t - y_{t|t-1})^2, \quad (30)$$

where $y_{t|t-1}$ is the forecast of y_t from the steady-state filter. For example, for Model 1, we would minimize the sum of $(e_t^2 - h_{t|t-1})^2$, where $h_{t|t-1}$ is given by Equation (21). This approach would yield consistent estimates of the GARCH parameters for cases in which the data are generated by the corresponding SARV model. In our empirical analysis we will refer to the GARCH models estimated using this approach as least-squares GARCH (LS-GARCH) models.

The main drawback of the least-squares approach is that the forecast errors in Equation (30) are either non-Gaussian (Model 3) or both non-Gaussian and conditionally heteroscedastic (Models 1 and 2), so it may be very inefficient.⁷

⁶ We might, for example, obtain consistent estimates of the GARCH(1,1) parameters via least squares and then infer the values of the diffusion parameters using the exact discretization of the continuous-time process developed in Meddahi and Renault (2002).

⁷ Bollerslev and Rossi (1995) provide a simple calculation, using results from Nelson and Foster (1994), which shows that the efficiency loss for a linear filter relative to an optimal ARCH filter can be substantial.

Presumably, inefficient parameter estimates lead to less accurate volatility forecasts. If we could somehow modify the linear filtering algorithm to take conditional heteroscedasticity into account, then we might obtain more efficient parameter estimates for Models 1 and 2. It turns out that this is roughly equivalent to what the corresponding GARCH models do if we assume a suitable distribution for the standardized returns and then fit the models via maximum likelihood. We illustrate this by considering an approximate approach to nonlinear filtering for the SARV models.

1.4 Approximate Nonlinear Filters

Recall that for Models 1 and 2 the innovation to the observation equation is conditionally heteroscedastic. Specifically we have $w_t = \sigma_t^2(z_t^2 - 1)$ and $w_t = \sigma_t(|z_t| - \psi)$, respectively. Perhaps we could improve the performance of the forecasting rules for these models by replacing the σ_t^2 and σ_t in these expressions with $s_{t|t-1}$ and approximating the conditional innovation variance as $R_{t|t-1} = s_{t|t-1} \text{var}(z_t^2)$ for Model 1 and $R_{t|t-1} = s_{t|t-1} \text{var}(|z_t|)$ for Model 2. We could then substitute $R_{t|t-1}$ for R and perform the updating and prediction steps of the filter as before.⁸ In effect, we would be implementing an extended Kalman filter in which the forecasts depend on the elements of \mathcal{I}_t in a nonlinear fashion.⁹

Approximate nonlinear filters do not have direct GARCH analogs because they imply time-varying filter gains even in steady state. Therefore we do not consider these filters in our empirical analysis. However, we can show that using maximum likelihood to fit the GARCH filters developed in Section 1.2 has a net effect that is very similar to fitting these nonlinear filters. An easy way to see this is to formulate a weighted least-squares (WLS) estimator for the nonlinear filters. We could, for instance, parameterize the filter in terms of $\theta = (\kappa, \phi, \gamma)'$, obtain an initial consistent estimate $\hat{\theta}$ via least squares, substitute $\hat{\theta}$ into the filter recursions to get $\{\hat{P}_{t|t-1}\}_{t=1}^T$ and $\{\hat{R}_{t|t-1}\}_{t=1}^T$, and then reestimate θ by minimizing the weighted sum of squares

$$S_{WLS} = \sum_{t=1}^T \frac{(y_t - y_{t|t-1})^2}{b^2 \hat{P}_{t|t-1} + \hat{R}_{t|t-1}}, \quad (31)$$

where $y_{t|t-1}$ is the forecast of y_t produced by the nonlinear filter and $b^2 \hat{P}_{t|t-1} + \hat{R}_{t|t-1}$ is its estimated conditional MSE.

Under a maximum-likelihood approach to fitting the GARCH models, the GARCH forecasts should tend to mimic those produced by this WLS procedure.

⁸ Note that in Model 1 the innovation to the state equation is conditionally heteroscedastic as well. Thus, when implementing the nonlinear filter for this model, we would approximate the conditional variance of $v_t = \gamma \sigma_t u_t$ as $Q_{t|t-1} = \gamma^2 s_{t|t-1}$ and modify the MSE recursion accordingly.

⁹ If the innovations in the state-space representation are Gaussian, then these forecasts approximate the conditional expectations of y_{t+1} and s_{t+1} given \mathcal{I}_t . For non-Gaussian innovations, however, this is no longer the case. See Hamilton (1994) for more details.

For example, consider a GARCH(1,1) model in which $e_t \sim N(0, h_{t|t-1})$ with $h_{t|t-1}$ given by Equation (21). The first-order conditions of the maximum-likelihood estimator are

$$\sum_{t=1}^T \frac{e_t^2 - h_{t|t-1}}{2h_{t|t-1}^2} \frac{\partial h_{t|t-1}}{\partial \theta} = 0. \quad (32)$$

Now compare Equation (32) to the first-order conditions implied by Equation (31) for a version of Model 1 in which $z_t \sim \text{NID}(0, 1)$. Since $y_t = e_t^2$, $y_{t|t-1} = h_{t|t-1}$, and $R_{t|t-1} = 2h_{t|t-1}^2$, we have

$$\sum_{t=1}^T \frac{e_t^2 - h_{t|t-1}}{\hat{P}_{t|t-1} + 2\hat{h}_{t|t-1}^2} \frac{\partial h_{t|t-1}}{\partial \theta} = 0, \quad (33)$$

where $h_{t|t-1}$ is our nonlinear forecast of e_t^2 . The parallels are immediately apparent. If $P_{t|t-1}$ is relatively small compared to $h_{t|t-1}$, then the forecasting performance of the GARCH model should be similar to that of the nonlinear filter as long as the variation in the filter gain is not too large. Similar parallels emerge for Model 2 by considering the maximum-likelihood estimator for an absolute value GARCH model in which the standardized return innovations have a generalized error distribution (GED) with tail thickness parameter one.¹⁰

These results provide another reason to suspect that GARCH models might perform well in forecasting stochastic volatility. Not only do the GARCH forecasts have the same recursive structure as those produced by MMS linear filters, the maximum-likelihood estimators for the GARCH models also have some ability to account for conditional heteroscedasticity in the forecast errors. Therefore, as part of our empirical analysis, we compare the forecasts obtained by maximum-likelihood estimation of the GARCH models to those obtained by least squares. We will refer to the models estimated in this fashion as ML-GARCH models.

Of course, unlike the LS-GARCH models, the ML-GARCH models are misspecified under a SARV data-generating process. Thus a maximum-likelihood approach will not yield a consistent estimator of the underlying SARV parameters. Neither, for that matter, will the WLS estimator for the approximate nonlinear filters. In general, optimal nonlinear filtering and fully efficient parameter estimation for SARV models requires Monte Carlo methods. We discuss our general approach to Monte Carlo filtering for SARV models below and postpone further discussion of SARV parameter estimation until Section 3.

¹⁰ Specifically, the first-order conditions for the maximum-likelihood and WLS estimators are

$$\sum_{t=1}^T \frac{|e_t| - 2\lambda h_{t|t-1}^{1/2}}{2\lambda h_{t|t-1}} \frac{\partial h_{t|t-1}^{1/2}}{\partial \theta} = 0 \quad \text{and} \quad \sum_{t=1}^T \frac{|e_t| - \psi h_{t|t-1}^{1/2}}{\psi^2 \hat{P}_{t|t-1} + (1 - \psi^2) \hat{h}_{t|t-1}} \frac{\partial h_{t|t-1}^{1/2}}{\partial \theta} = 0,$$

respectively, where $\lambda = 0.5\sqrt{\Gamma(1)/\Gamma(3)}$. Nelson (1991) provides details on the properties of the GED in the context of GARCH modeling.

1.5 Optimal Nonlinear Filters

Let $\mathcal{F}_t = \{r_1, r_2, \dots, r_t\}$ denote the date t information set for the SARV model. To construct an optimal nonlinear filter we have to implement the exact update and prediction steps implied by Bayes' theorem. In terms of the probability densities, we have

$$p(s_t | \mathcal{F}_t, \theta) \propto p(r_t | s_t, r_{t-1}, \theta) p(s_t | \mathcal{F}_{t-1}, \theta), \quad (34)$$

where

$$p(s_t | \mathcal{F}_{t-1}, \theta) = \int p(s_t | s_{t-1}, \theta) p(s_{t-1} | \mathcal{F}_{t-1}, \theta) ds_{t-1}. \quad (35)$$

Thus, as in linear filtering, the problem comes down to implementing a set of recursions. We start with the prediction density $p(s_t | \mathcal{F}_{t-1}, \theta)$ and obtain the filtering density $p(s_t | \mathcal{F}_t, \theta)$ by evaluating Equation (34). Once we have $p(s_t | \mathcal{F}_t, \theta)$, we obtain the prediction density $p(s_{t+1} | \mathcal{F}_t, \theta)$ by evaluating Equation (35) for date $t + 1$. We use a particle filter method to carry out the required computations [see, e.g., Gordon, Salmond, and Smith (1993), Kitagawa (1996), Pitt and Shephard (1999), and Chib, Nardari, and Shephard (2002)]. Appendix A provides the details.

Given the sequence of prediction and filtering densities, we can use a Monte Carlo approach to compute forecasts that parallel those produced by the MMS linear filters. To illustrate, let $s_{t+k|t} = E(s_{t+k} | \mathcal{F}_t, \theta)$ denote the expected date $t + k$ variance, volatility, or log variance given the returns observed through date t . By definition, this k -step-ahead forecast is given by

$$s_{t+k|t} = \int s_{t+k} p(s_{t+k} | \mathcal{F}_t, \theta) ds_{t+k}. \quad (36)$$

We evaluate Equation (36) using the following procedure. First, we use the particle filter to generate a random sample of N observations $s_t^{(1)}, s_t^{(2)}, \dots, s_t^{(N)}$ from the filtering density $p(s_t | \mathcal{F}_t, \theta)$. Next, we use the SARV model with $s_t = s_t^{(i)}$ to simulate $s_{t+1}^{(i)}, s_{t+2}^{(i)}, \dots, s_{t+k}^{(i)}$ for all $i \leq N$. Finally, we approximate the integral as

$$s_{t+k|t} \approx \frac{1}{N} \sum_{i=1}^N s_{t+k}^{(i)}. \quad (37)$$

The approximation error becomes negligible as $N \rightarrow \infty$. In Section 2.1 we describe how we use the resulting forecasts to benchmark the performance of the MMS linear filters.

1.6 Leverage Effects and GARCH Filters

Our three SARV models assume that z_t is independent of u_τ for all τ . This rules out leverage effects. Incorporating leverage is typically accomplished in the SARV literature by assuming that z_t and u_t are jointly drawn from a symmetric distribution, such as a bivariate normal or bivariate Student's t , with an unknown correlation coefficient [see, e.g., Jacquier, Polson, and Rossi (2001)]. Although we could

easily implement this approach, it would have no effect on the MMS linear filters. With a symmetric distribution, any even function of e_t , such as e_t^2 , $|e_t|$, and $\log e_t^2$, is uncorrelated with v_t regardless of the correlation between z_t and u_t . Since our LSSRs imply that the information set \mathcal{I}_t contains only even functions of e_t , this means that leverage effects would play no role in determining the MMS linear forecasts of the variance, volatility, or log variance.

It might be possible to modify our linear filters to capture leverage effects by expanding \mathcal{I}_t to include additional variables. Harvey and Shephard (1996) show how to do this for Model 3. The idea is to augment \mathcal{I}_τ with the sign of e_t for $t = 1, 2, \dots, \tau$ and condition on this information when implementing the update and prediction steps of the filter. We could probably develop a similar approach for Models 1 and 2. However, the resulting forecasting rules would have time-varying coefficients, even in steady state, which takes us outside the realm of conventional GARCH models. Since our interest lies in evaluating the performance of conventional GARCH models in SARV applications, we do not pursue the issue of leverage effects here.

2 THE EMPIRICAL FRAMEWORK

Although most of the applied research on estimating volatility is based on GARCH models, this is largely a result of computational convenience. Indeed, theories of speculative trading, such as those developed by Tauchen and Pitts (1983) and Andersen (1996), provide strong arguments for modeling volatility as stochastic. It is therefore natural to ask whether GARCH models are capable of accurately approximating a stochastic volatility process in common applications such as forecasting volatility. In this section we discuss the general elements of our approach for addressing this issue.

2.1 Analysis Using Artificial Data

Simulations are integral to our analysis because they allow us to assess the performance of GARCH and SARV models under controlled conditions. In broad terms, our strategy is to generate a sequence of artificial returns, $\{r_t\}_{t=1}^T$, from each of our discrete-time SARV models, fit the GARCH models to the artificial returns, and then evaluate the performance of the resulting GARCH forecasts. We focus in particular on how the GARCH forecasts perform in relation to the optimal non-linear forecasts obtained by fitting the SARV model to the artificial returns and implementing the filter discussed in Section 1.5.

Recall that $s_{t+k|t} = E(s_{t+k} | \mathcal{F}_t, \theta)$ denotes the expected date $t+k$ variance, volatility, or log variance given the returns observed through date t . After fitting the SARV model, we estimate $s_{t+k|t}$ for each t using three different choices of k : $k=0$, $k=1$, and $k=10$. The first choice of k corresponds to using the data through date t to estimate the variance, volatility, or log variance for date t . We call this the “filtered” estimate. The other two choices of k correspond to using the data through date t to estimate the variance, volatility, or log variance for date $t+1$

and for date $t + 10$. We call these the “one-step-ahead” and “ten-step-ahead” forecasts. In addition, we obtain an estimate of $s_{t|T}$ as part of our methodology for fitting the SARV model. We call this the “smoothed” estimate of the variance, volatility, or log variance because it corresponds to estimating s_t using all of the available data.

We evaluate the performance of the forecasts using standard statistical criteria. To illustrate, let $\hat{s}_{t+k|t}$ denote our SARV estimate of $s_{t+k|t}$. We use the mean absolute forecast error (MAE) and root mean-squared forecast error (RMSE) to measure how well the sequence $\{\hat{s}_{t+k|t}\}_{t=1}^{T-k}$ tracks the true variance, volatility, or log variance process. These criteria are given by

$$MAE = \frac{1}{T-k} \sum_{t=1}^{T-k} |\hat{s}_{t+k|t} - s_{t+k}| \quad (38)$$

and

$$RMSE = \left(\frac{1}{T-k} \sum_{t=1}^{T-k} (\hat{s}_{t+k|t} - s_{t+k})^2 \right)^{1/2}, \quad (39)$$

respectively. Similarly we use the difference between the MAE (or RMSE) for the GARCH and SARV forecasts to assess how well the GARCH model approximates the SARV model.

2.2 Analysis Using Real Data

The simulations allow us to evaluate the forecasting performance of GARCH models under controlled conditions. More generally, however, we are interested in determining which class of models—GARCH or SARV—performs better in real data applications. Since theory provides support for modeling volatility as stochastic, we might expect to find in general that the SARV models outperform the GARCH models. We investigate whether this is the case using a standard application in risk management: estimating VaR. By framing the analysis in terms of VaR estimation, we obtain direct evidence on the economic significance of differences in forecasting performance between models.

The basics of our approach are straightforward. Suppose, for example, that we want to assess how the GARCH variance model performs in comparison to the SARV variance model. After fitting the two models to the data, we construct a pair of one-step-ahead volatility forecasts for each date t in the sample. Using these volatility forecasts, we form two sets of one-day VaR estimates and evaluate their performance in both an absolute and relative sense. We make this evaluation by looking at two forms of specification diagnostics, one of which is an information-theoretic criterion that allows us to compare potentially misspecified VaR measures across models.

Our strategy for constructing the VaR estimates follows Christoffersen, Hahn, and Inoue (2001). Let $\mu_{t|t-1}$ and $\varsigma_{t|t-1}$ denote the conditional mean and conditional

volatility of r_t . We begin by assuming that $(r_t - \mu_{t|t-1})/\varsigma_{t|t-1}$ is an i.i.d. random variable. This allows us to express the $\pi\%$ conditional quantile for the return innovation, e_t , as

$$F_{t|t-1}(\mathbf{q}_\pi) = q_{\pi 1} + q_{\pi 2}\varsigma_{t|t-1}, \quad (40)$$

for some $q_{\pi 1}$ and $q_{\pi 2}$. This approach encompasses most of the methods employed in the literature. It is common, for example, to assume that the standardized returns are i.i.d. normal random variables, which implies that the 5% conditional quantile corresponds to $q_{\pi 1} = 0$ and $q_{\pi 2} = -1.645$. More generally, the distribution of standardized returns could be asymmetric and/or fat-tailed relative to the normal, so we leave this distribution unspecified and estimate $q_{\pi 1}$ and $q_{\pi 2}$ from the data.

To illustrate, let $\iota_t(\mathbf{q}_\pi) = I(e_t < F_{t|t-1}(\mathbf{q}_\pi))$, where $I(\cdot)$ is the indicator function. If a VaR measure is well specified, then it must satisfy

$$E((\iota_t(\mathbf{q}_\pi) - \pi)\mathbf{x}_{t-1}) = \mathbf{0}, \quad (41)$$

where \mathbf{x}_{t-1} is any vector of instruments in the risk manager's information set. Thus our null hypothesis is that Equation (41) holds for either $\varsigma_{t|t-1} = \hat{\sigma}_{t|t-1}$ or $\varsigma_{t|t-1} = \hat{h}_{t|t-1}^{1/2}$, depending on the type of model under consideration. Our estimator of \mathbf{q}_π , which is given by

$$\hat{\mathbf{q}}_\pi = \operatorname{argmax}_{\mathbf{q}_\pi} \min_{c_\pi} \frac{1}{T} \sum_{t=1}^T \exp(c'_\pi((\iota_t(\mathbf{q}_\pi) - \pi)\mathbf{x}_{t-1})), \quad (42)$$

is obtained by minimizing the sample counterpart of the Kullback-Leibler information criterion [see Kitamura and Stutzer (1997)]. It might seem easier to employ a GMM estimator since it would have the same limiting distribution as $\hat{\mathbf{q}}_\pi$. However, like Christoffersen, Hahn, and Inoue (2001), we adopt the information theoretic framework to facilitate comparisons of nonnested and potentially misspecified models.

2.2.1 Specification tests Equation (41) implies that the expected value of $\iota_t(\mathbf{q}_\pi)$ equals π , that it is independently distributed, and that it is uncorrelated with any vector of instruments in the risk manager's information set. We assess whether our VaR estimates violate any of these conditions using two types of specification tests. The first test, which is based on a likelihood ratio, is from Christoffersen (1998). The second, which is an information theoretic analog of the GMM test for overidentifying restrictions, is from Christoffersen, Hahn, and Inoue (2001).

For the first test, we implement the estimator in equation (42) with $\mathbf{x}_{t-1} = (1, \varsigma_{t|t-1})'$, which is analogous to using a quantile regression approach to estimate \mathbf{q}_π . Since \mathbf{x}_{t-1} and \mathbf{q}_π are the same dimension, there are no overidentifying restrictions to test. However, we can construct a likelihood ratio test by noting

that ι_t is the outcome of a binomial process. This allows us to write the likelihood function for the sequence $\{\iota_t(\hat{q}_\pi)\}_{t=1}^T$ as

$$L_1 = (1 - \pi_1)^{T-N_1} \pi_1^{N_1}, \quad (43)$$

where π_1 is the unconditional probability of an exceedence and N_1 is the number of exceedences in a sample of size T . Similarly, if we let $\pi_{1|j} = \Pr(\iota_t(\hat{q}_\pi) = 1 | \iota_{t-1}(\hat{q}_\pi) = j)$, then independence requires that $\pi_{1|0} = \pi_{1|1} = \pi_1$. We test the joint hypothesis that $\pi_1 = \pi$ and $\pi_{1|0} = \pi_{1|1} = \pi_1$ using the likelihood ratio statistic

$$LR = -2 \log(L_1/L_2), \quad (44)$$

where

$$L_2 = (1 - \hat{\pi}_{1|0})^{N_{0|0}} \hat{\pi}_{1|0}^{N_{1|0}} (1 - \hat{\pi}_{1|1})^{N_{0|1}} \hat{\pi}_{1|1}^{N_{1|1}} \quad (45)$$

and $\pi_1 = \pi$. Here $\hat{\pi}_{ij}$ denotes our estimate of $\Pr(\iota_t(\hat{q}_\pi) = i | \iota_{t-1}(\hat{q}_\pi) = j)$ and N_{ij} is the number of times this outcome is observed in our sample. This statistic is asymptotically χ^2_2 under the null.

For the second test, we implement the estimator in Equation (42) using an expanded set of instruments. The vector \mathbf{x}_{t-1} contains a constant, the one-step-ahead forecasts for the three SARV models, and the one-step-ahead forecasts for both the least-squares and maximum-likelihood versions of the three GARCH models. This allows us to use the minimized value of the objective function as the basis for a specification test. In particular, we have

$$-2T \log \left(\frac{1}{T} \sum_{t=1}^T \exp(\hat{c}'_\pi((\iota_t(\hat{q}_\pi) - \pi)\mathbf{x}_{t-1})) \right) \rightarrow \chi^2_8. \quad (46)$$

This criterion, which we call the κ -statistic, is analogous to the J -statistic for over-identified GMM systems. It tests whether we can predict $\iota_t(\hat{q}_\pi) - \pi$ for a given model using the forecasts produced by any of the competing models.

2.2.2 Comparing VaR estimates Even if the specification tests lead to rejections for both the GARCH and SARV models, we might still be interested in whether one model performs significantly better than another. This requires a method for comparing VaR estimates for nonnested and potentially misspecified models. To this end, suppose that

$$M_T^i(\hat{q}_\pi^i, \hat{c}_\pi^i) = \frac{1}{T} \sum_{t=1}^T \exp(\hat{c}_\pi^{i'}((\iota_t(\hat{q}_\pi^i) - \pi)\mathbf{x}_{t-1})) \quad (47)$$

denotes the minimized value of the objective function when we estimate q_π for model i using the larger of our two instrument sets. Under the null that there is no difference in the performance of the VaR estimates for models i and j , Christoffersen, Hahn, and Inoue (2001) show that

$$\sqrt{T}(M_T^i(\hat{q}_\pi^i, \hat{c}_\pi^i) - M_T^j(\hat{q}_\pi^j, \hat{c}_\pi^j)) \rightarrow N(0, \sigma_\infty^2), \quad (48)$$

where

$$\sigma_\infty^2 = \lim_{T \rightarrow \infty} \text{var} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \exp(c_\pi^j((\iota_t(q_\pi^j) - \pi)x_{t-1})) - \exp(c_\pi^j((\iota_t(q_\pi^j) - \pi)x_{t-1})) \right). \quad (49)$$

Thus we can rely on asymptotic t -ratios to evaluate whether a given model performs significantly better or worse than another.¹¹ We refer to these t -ratios as Δ -statistics because they measure differences in performance between models.

3 THE SIMULATIONS

The analysis of Section 1 shows that we can use properly specified GARCH models to obtain MMS linear forecasts of stochastic volatility. By construction, however, these GARCH forecasts will be less precise under an MSE criterion than the forecasts produced by an optimal nonlinear SARV filter. In this section we investigate the performance of the GARCH forecasts using artificial data generated under each of our three SARV models. We begin by describing how we generate the data, how we fit the GARCH and SARV models, and how we construct the volatility forecasts. We follow this with the simulation results.

3.1 Generating the Data

To generate the data, we first need to parameterize the three SARV models. We simplify the analysis by setting $\mu = \delta = 0$ (both when generating the data and when fitting the models). We set the remaining parameters (i.e., κ , ϕ , and γ) equal to the posterior means obtained by fitting the SARV models to NYSE index returns (see below). The next step is to specify the distributions for z_t and u_t . We assume that both variables are standard normal since this is the most common approach in the literature. An obvious concern, however, is that this allows the variance and volatility in Models 1 and 2 to potentially go negative. But this is more of a theoretical than a practical concern given our parameterizations. For Model 2—the more likely of the models to produce negative values—our parameterization implies that $\sigma_t \sim N(0.8, 0.091)$.¹² Consequently we expect only about 0.4% of the realizations of σ_t to be negative, which should have a negligible impact on our results.

Given this setup, we perform 1000 simulation trials for each SARV model. In each trial we create an artificial sample of $T = 2000$ daily returns by (i) generating $\{z_t\}_{t=1}^T$ and $\{u_t\}_{t=0}^T$ from a standard normal distribution, (ii) using Equation (2), (3), or (4) to construct $\{\sigma_t^2\}_{t=1}^T$, $\{\sigma_t\}_{t=1}^T$, or $\{\log \sigma_t^2\}_{t=1}^T$, (iii) converting this sequence into volatilities using the appropriate model-specific transformation, and (iv) using Equation (1) to construct $\{r_t\}_{t=1}^T$. We note for future reference that,

¹¹ We compute the t -ratios using the Newey and West (1987) estimator with a lag length of six to estimate σ_∞^2 .

¹² To see this, note that the unconditional distribution of σ_t is $N(\kappa/(1 - \phi), \gamma^2/(1 - \phi^2))$.

under our parameterization of the three SARV models, the MMS linear filters have ω , β , and α values of 0.015, 0.922, and 0.058 for Model 1; 0.012, 0.921, and 0.080 for Model 2; and 0.049, 0.935, and 0.047 for Model 3.¹³

3.2 Fitting the GARCH Models

For each artificial sample we estimate the three GARCH models corresponding to Equations (21), (25), and (29). As mentioned earlier, we estimate each model twice. First, we fit the models using ordinary least squares as described in Section 1.3 (i.e., the LS-GARCH models). The least-squares approach delivers consistent parameter estimates. Second, we fit the models via maximum-likelihood by specifying a GED for the standardized returns (i.e., the ML-GARCH models). The maximum-likelihood approach entails a specification of the form $e_t = h_{t-1}^{1/2} \epsilon_t$ with

$$p(\epsilon_t) = \frac{\nu \exp(-(1/2)|\epsilon_t/\lambda|^\nu)}{\lambda 2^{(\nu+1)/\nu} \Gamma(1/\nu)}, \quad (50)$$

where $\Gamma(\cdot)$ is the gamma function, ν is a tail-thickness parameter, and λ is given by

$$\lambda = \left(\frac{2^{-2/\nu} \Gamma(1/\nu)}{\Gamma(3/\nu)} \right)^{1/2}. \quad (51)$$

For $\nu = 2$, equation (50) is just a standard normal density. For $\nu < 2$, it has thicker tails than the standard normal. By estimating ν as a free parameter, we accommodate the excess kurtosis displayed by ϵ_t . This excess kurtosis reflects the inability of the GARCH models to capture the unpredictable component of volatility.

We know from Section 1.4 that, in general, the maximum-likelihood estimates will not be consistent for data generated under a SARV process. But this may not necessarily be a serious deficiency in a forecasting context. If the probability limit of the maximum-likelihood estimator is close to the true parameter vector, then the lack of consistency might be more than offset in small samples by the efficiency gains from partially accounting for the conditional heteroscedasticity in the forecast errors. Of course, the opposite could be true as well. The only way to tell for sure is to fit the models by maximum-likelihood and see whether they do better or worse than when we fit them by least squares. This also allows us to assess how the GARCH models perform using estimation methods that are typically employed in practice. Although there are studies, such as Rich, Raymond, and Butler (1991), that use methods other than maximum-likelihood to fit GARCH models, they account for only a tiny fraction of the vast GARCH literature.

¹³ To convert the values of κ , ϕ , and γ into values for ω , α , and β , we first compute Q and R and iterate on the Kalman filter MSE recursion to find the steady-state value of the filter gain. Once we have the steady-state filter gain, the conversion is straightforward (see Section 1.2).

3.3 Fitting the SARV Models

In each simulation trial we also fit the three SARV models. Since a maximum-likelihood approach is too computationally intensive to be practical, we adopt a Bayesian perspective and use Markov chain Monte Carlo (MCMC) methods to carry out the required computations. These methods, which were first applied in a SARV context by Jacquier, Polson, and Rossi (1994), have rapidly become standard tools in Bayesian time-series research because they allow us to sample from analytically intractable posteriors. The idea is to generate candidate values from a tractable proposal density, such as a multivariate normal distribution, and correct for this approximation via an accept/reject step commonly known as the Metropolis-Hastings algorithm. Our approach is built around a computationally efficient sampling scheme for generating the unobserved variances, volatilities, and log variances. We enforce nonnegativity of the variances and volatilities in Models 1 and 2 by directly incorporating this constraint in our sampling scheme. Details are provided in Appendix B.

We specify noninformative priors for all of the parameters to ensure that our inferences are driven by the data. Since the sampling scheme is initialized using an arbitrary set of parameter values, we require a burn-in period to allow the scheme to converge. Following the burn-in period, the iterates produced by the sampling scheme correspond to the correct posterior distribution. After some experimentation, we determined that 2000 iterations were sufficient for convergence to occur. Therefore we fit the models for each simulation trial by performing 12,000 iterations of the sampling scheme and discarding the output from the first 2000. The remaining output is used to conduct posterior inferences. Note that we do not use the optimal nonlinear filter in fitting the SARV models except in the following sense. Under an MCMC approach, we condition on the entire return history when generating the unobserved variances, volatilities, and log variances. Thus, averaging the 10,000 MCMC iterates of s_t is equivalent to producing a smoothed estimate of the date t variance, volatility, or log variance using the optimal nonlinear filter for the model.

3.4 Constructing the Forecasts

The one-step-ahead forecasts for the GARCH models are just the fitted variances, volatilities, and log variances based on the least-squares or maximum-likelihood estimates of the model parameters. To construct the ten-step-ahead forecasts, we simply note that, as the horizon increases, the forecasts decay toward a fixed long-run value, with the level of persistence determined by $\beta + \alpha$ in the case of Models 1 and 3, and $\beta + \psi\alpha$ in the case of Model 2. For the SARV models, we use the Monte Carlo approach discussed in Section 1.5 to obtain the forecasts and the particle filter method discussed in Appendix A to obtain the filtered estimates. To implement the particle filter, we set the model parameters equal to their posterior means and use 5000 and 1000 particles for the first- and second-stage sampling operations, respectively.

3.5 Simulation Results

Table 1 presents the results of fitting the GARCH models to the artificial data. We summarize the results for each data-generating process (DGP) under a separate heading. Since we fit the models using both least-squares and maximum-likelihood methods, we need to perform six different estimations for a given DGP. Each row of the table reports the mean and standard deviation of the 1000 estimates of each parameter, along with the average value of the logarithm of the Bayesian

Table 1 GARCH models fitted to returns generated under stochastic volatility.

Model	ω	S.E.	α	S.E.	β	S.E.	ν	S.E.	BIC
M₀: SARV variance									−2381.9 ¹²³
LS-GARCH variance	0.018	0.010	0.059	0.013	0.915	0.025	—	—	—
LS-GARCH volatility	0.020	0.008	0.083	0.014	0.909	0.018	—	—	—
LS-GARCH log variance	0.050	0.012	0.052	0.012	0.920	0.021	—	—	—
1. ML-GARCH variance	0.013	0.005	0.089	0.017	0.896	0.019	1.64	0.08	−2389.8 ³
2. ML-GARCH volatility	0.019	0.007	0.095	0.016	0.903	0.017	1.64	0.08	−2389.5 ³
3. ML-GARCH log variance	0.050	0.011	0.045	0.009	0.927	0.018	1.59	0.08	−2406.7
M₀: SARV volatility									−2307.4 ¹²³
LS-GARCH variance	0.014	0.008	0.057	0.013	0.923	0.021	—	—	—
LS-GARCH volatility	0.015	0.006	0.082	0.015	0.916	0.017	—	—	—
LS-GARCH log variance	0.067	0.022	0.064	0.021	0.913	0.025	—	—	—
1. ML-GARCH variance	0.007	0.005	0.094	0.021	0.899	0.019	1.61	0.10	−2324.2
2. ML-GARCH volatility	0.013	0.006	0.106	0.025	0.902	0.019	1.61	0.10	−2321.9 ³
3. ML-GARCH log variance	0.059	0.015	0.051	0.013	0.924	0.018	1.56	0.11	−2339.6
M₀: SARV log variance									−2320.5 ²³
LS-GARCH variance	0.019	0.015	0.073	0.021	0.900	0.040	—	—	—
LS-GARCH volatility	0.018	0.007	0.090	0.016	0.905	0.019	—	—	—
LS-GARCH log variance	0.045	0.012	0.047	0.009	0.930	0.017	—	—	—
1. ML-GARCH variance	0.011	0.005	0.085	0.013	0.900	0.016	1.66	0.08	−2325.1 ³
2. ML-GARCH volatility	0.016	0.006	0.092	0.013	0.909	0.015	1.65	0.08	−2326.8 ³
3. ML-GARCH log variance	0.051	0.011	0.045	0.008	0.933	0.015	1.59	0.08	−2347.8

The table reports the results of fitting GARCH models to returns generated from three different stochastic volatility specifications (M₀): SARV models in variance, volatility, and log variance. For each SARV model we use the parameter estimates obtained from fitting the model to NYSE index returns (see Table 6), we generate an artificial sample of 2000 returns, and we use these returns to estimate LS-GARCH and ML-GARCH models in variance, volatility, and log variance. We repeat this experiment 1000 times. The table reports the average parameter estimates for each GARCH model, where ω denotes the constant term, α and β denote the coefficients on lagged returns and volatility, and ν denotes the GED constant. We also report the standard errors (S.E.) for each parameter estimate and the average log Bayesian information criterion (BIC) for the fitted models. ^{1,2,3} denote log BICs that are greater than those for the ML-GARCH models in (1) variance, (2) volatility, and (3) log variance in 95% of the artificial samples.

information criterion for the maximum-likelihood estimations.¹⁴ To provide a benchmark for comparison, we also report the average log Bayesian information criterion for the SARV model itself. We compute this by replacing the SARV parameters with their posterior means, using the algorithm in Appendix A to find the associated log likelihood, and then averaging the log Bayesian information criterion values across the 1000 simulation trials.

We begin with the least-squares results. For the models that correspond to the MMS linear filters — the variance model for the variance DGP, the volatility model for the volatility DGP, and the log variance model for the log variance DGP — the mean parameter estimates are close to the true parameters for the simulations. The variance model, for example, yields mean parameter estimates of 0.018, 0.915, and 0.059 for ω , β , and α , respectively, compared to the true parameter values of 0.015, 0.922, and 0.058. This is expected since the MMS linear filters are well specified under the least-squares approach. When we pair each model with either of the other two DGPs, the mean parameter estimates need not bear any particular relation to the true parameters. But we would expect them to indicate strong persistence regardless of the DGP, and this is indeed the case.

Turning next to the maximum-likelihood results, it is apparent that our GED approach captures significant excess kurtosis in the conditional distribution of returns, with mean ν estimates in the 1.5 to 1.7 range. We know, however, that under this approach, the MMS linear filters are no longer well specified for the variance and volatility DGPs. Consequently, if we look at the variance and volatility models under these DGPs, we find that the mean estimates of ω , β , and α tend to diverge from the true parameters. The effect is most pronounced for α and β . For example, fitting the variance model yields a mean α estimate of 0.089 and a mean β estimate of 0.896. This translates into using a steady-state filter gain that is too large on average. Fitting the log variance model, on the other hand, produces mean parameter estimates that are very close to those obtained by least squares. But this is not surprising given the similarity between the first-order conditions for the maximum-likelihood and least-squares estimators.¹⁵

Although we are dealing with nonnested models, we can get an idea of how they compare on the goodness-of-fit dimension using the average log Bayesian information criterion values. We report these with superscripts to identify models that fit significantly better than others: “1” represents the GARCH variance model, “2” represents the GARCH volatility model, and “3” represents the GARCH log variance model. For example, fitting the GARCH variance model under the variance DGP produces an average log Bayesian information criterion of -2389.8 .

¹⁴ The log Bayesian information criterion is equal to the log likelihood minus $(N/2)\log T$, where N is the number of parameters and T is the sample size. See Schwarz (1978).

¹⁵ The first-order conditions for the maximum-likelihood estimator of $\boldsymbol{\theta} = (\omega, \beta, \alpha)'$ can be written as

$$\sum_{i=1}^T \left(\frac{\nu}{2\lambda^\nu} \exp\left(\frac{1}{2}\nu(\log e_i^2 - \log h_{i|t-1})\right) - 1 \right) \frac{\partial \log h_{i|t-1}}{\partial \boldsymbol{\theta}} = \mathbf{0}.$$

If we approximate the exponential with the first two terms of its infinite Taylor series expansion about zero, then we end up with first-order conditions that are nearly identical to those of the least-squares estimator of $\boldsymbol{\theta}$.

Table 2 Forecast errors for returns generated under stochastic volatility.

Forecasts	M₀: SARV variance		M₀: SARV volatility		M₀: SARV log variance	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
SARV model						
Smoothed estimates	0.1728	0.2340	0.0970	0.1230	0.2585	0.3242
Filtered estimates	0.2171	0.2913	0.1243	0.1572	0.3278	0.4113
One-step-ahead forecasts	0.2251 ⁰	0.3009 ⁰	0.1294 ⁰	0.1633 ⁰	0.3408 ⁰	0.4274 ⁰
GARCH models						
LS-GARCH variance	0.2319 ²	0.3079 ¹	0.1408 ⁴	0.1763 ⁴	0.3677 ⁴	0.4633 ⁴
LS-GARCH volatility	0.2383 ¹	0.3104 ²	0.1344 ¹	0.1692 ¹	0.3495 ¹	0.4375 ¹
LS-GARCH log variance	0.2545 ⁵	0.3492 ⁶	0.1530 ⁶	0.1943 ⁶	0.3937 ⁵	0.4929 ⁵
ML-GARCH variance	0.2362 ⁴	0.3239 ⁴	0.1376 ²	0.1742 ²	0.3494 ¹	0.4368 ¹
ML-GARCH volatility	0.2336 ³	0.3165 ³	0.1381 ³	0.1745 ³	0.3537 ³	0.4420 ³
ML-GARCH log variance	0.2549 ⁵	0.3389 ⁵	0.1513 ⁵	0.1896 ⁵	0.4027 ⁶	0.5030 ⁶

The table reports summary statistics for the forecast errors for the GARCH and SARV models fitted in Table 1. In each of the artificial samples we construct the smoothed, filtered, and one-step-ahead volatility (or variance or log variance) forecasts under the SARV model and we compute the fitted estimates for each GARCH model. We then calculate the mean absolute error (MAE) and the root mean-squared error (RMSE) for each set of forecasts in the artificial sample. The table reports the average MAEs and RMSEs across 1000 artificial samples. The superscripts indicate the number of alternative models that have averages significantly smaller (at the 99% level) than the reported average. We exclude the smoothed and filtered estimates from these significance tests.

We report this with a “3” superscript to indicate that the log Bayesian information criterion for this model is greater than the log Bayesian information criterion for the GARCH log variance model in at least 95% of the simulation trials. Note that the GARCH variance and volatility models typically produce similar results, and none of the differences in their log Bayesian information criterion values are significant at the 5% level. In contrast, the GARCH log variance model performs significantly worse than these models, even under the log variance DGP. This is likely due to its multiplicative structure, a feature that makes it vulnerable to returns close to zero.

The one unanticipated finding from Table 1 concerns estimator efficiency. Specifically, the maximum-likelihood estimator does not seem to have any clear efficiency advantage over the least-squares estimator. In some cases the standard deviation of the least-squares estimates is larger than that of the maximum-likelihood estimates, and in others the opposite relation holds. For instance, for the subset of results in which the least-squares approach delivers consistent parameter estimates, the standard errors are equal to or lower than those of the maximum-likelihood estimates in four out of nine cases. This finding, along with the adverse impact of specification error documented earlier, suggests that least-squares may be the preferred approach for fitting the MMS linear filters.

Table 2 provides additional evidence on the relative merits of the least-squares and maximum-likelihood methods. The table examines the forecasting performance of the GARCH models in the simulation trials. To construct the table, we generate one-step-ahead forecasts of the variances, volatilities, or log variances and compute their MAEs and RMSEs. In each case we include the MAEs and RMSEs of the optimal SARV forecasts for comparison. The columns are grouped in sets of two, corresponding to the SARV variance, volatility, and log variance DGPs. The superscripts on the MAEs and RMSEs denote the number of competing specifications with averages significantly smaller (at the 99% level) than the averages reported for each model.

The first three rows of Table 2 summarize the results for the SARV model used to generate the data. In addition to the one-step-ahead forecasts, we also examine the performance of the SARV filtered and smoothed estimates. The average MAEs and RMSEs for the one-step-ahead forecasts are around 3% to 4% larger than those for the filtered estimates, which in turn are around 25% to 28% larger than those for the smoothed estimates. Since each successive set of estimates in this progression conditions on a larger information set, these differences in performance are a direct measure of the incremental value of the additional information. This helps place the observed differences in performance between the one-step-ahead SARV and GARCH forecasts in perspective.

The performance of the GARCH forecasts is summarized in the last six rows of the table. All of the MAEs and RMSEs for the GARCH one-step-ahead forecasts are significantly larger than those for the SARV one-step-ahead forecasts at the 1% level. This is not surprising given that the SARV forecasts are based on the optimal nonlinear filter. The more interesting finding concerns the differences in performance between the various GARCH specifications. The best overall performer is the LS-GARCH volatility model. It yields MAEs and RMSEs around 1% to 4% higher than those of the SARV one-step-ahead forecasts. There is only one instance in which another GARCH model does significantly better: when we fit the LS-GARCH variance model under the variance DGP and we use the RMSE criterion to evaluate performance.

In contrast, the GARCH log variance model is at the other end of the performance spectrum. It always ranks worst, regardless of how we estimate the models. Moreover, the associated MAEs and RMSEs are around 13% to 19% higher than those of the SARV one-step-ahead forecasts. To put this in perspective, recall that the MAEs and RMSEs of the SARV one-step-ahead forecasts are around 25% to 28% larger than those for the smoothed estimates. Thus it seems clear that the one-step-ahead forecasts from the GARCH log variance model are much less precise than those from both the SARV model and the LS-GARCH volatility model. Apparently the GARCH log variance model does a poor job of capturing the dynamics of the underlying SARV process.

Two other aspects of the results in Table 2 are also worth noting. First, if we use the RMSE criterion as our performance metric, then the best forecasts for the variance and volatility DGPs are obtained by using the LS-GARCH variance and volatility models. In other words, implementing the MMS linear filters with

consistent parameter estimates produces the best forecasts for the variance and volatility DGPs. This is in line with both the analysis in Section 1 and the model fitting results in Table 1. Second, fitting a model by maximum-likelihood may produce significantly better or worse performance than fitting it by least-squares. Although the explanation for this is not immediately clear, the consistently good performance of the LS-GARCH volatility model suggests that outliers might play a role in explaining the differences across models and estimation methodologies.

Davidian and Carroll (1987), for example, argue that using absolute return innovations to estimate conditional variances is more robust than using squared innovations, because estimators based on the former are less sensitive to outliers. This is essentially the same argument for why the GARCH log variance model performs poorly, since returns close to zero become outliers under a logarithmic transformation. To see if outliers have a significant impact on the forecasting performance of the GARCH models, we take the output from each simulation run and sort the forecasts into three categories, corresponding to the lowest 5%, middle 90%, and highest 5% of the absolute value of the previous day's return innovation. This allows us to assess whether the relative performance of the models changes following either large or small shocks to returns.

Table 3 presents the results of this analysis. In some cases we see a substantial change in the relative performance of a model following a high or low absolute return realization. Consider, for example, the LS-GARCH variance model under the volatility or log variance DGP. There is only one specification—the SARV model—that performs significantly better for lagged absolute returns in the upper 5% tail. But there are at least four specifications that perform significantly better for lagged absolute returns in the middle 90% and lower 5% tail. With minor exceptions, however, the rankings of the models for the middle 90% of lagged absolute returns are the same as those obtained without trimming the tails of the distribution. Therefore it does not appear that we can attribute the results in Table 2 to the impact of outliers.

3.6 Robustness Tests

The evidence thus far points to statistically significant differences in how well the three GARCH models perform as filters in a discrete-time SARV framework. To assess the robustness of this evidence, we extend the simulations along several dimensions. First, we check to see whether the relative forecasting performance of the models is sensitive to the forecast horizon. Second, we investigate whether the Monte Carlo error in our SARV forecasts is large enough to have a meaningful impact on the results. Third, we repeat the simulations using out-of-sample forecasts to see how this affects our conclusions. Finally, we consider the impact of using different parameter values to generate the artificial data.

Table 4 illustrates how changing the forecast horizon affects our results. The table replicates the analysis of Table 2 using the ten-step-ahead forecasts. Although all the MAEs and RMSEs increase with the horizon, there is little change

Table 3 Forecast errors conditioned on lagged absolute returns

Forecasts	M ₀ : SARV variance		M ₀ : SARV volatility		M ₀ : SARV log variance	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
Lowest 5% absolute returns						
SARV model	0.1997 ⁰	0.2696 ⁰	0.1247 ⁰	0.1584 ⁰	0.3534 ⁰	0.4417 ⁰
LS-GARCH variance	0.2163 ⁴	0.2803 ³	0.1727 ⁶	0.2121 ⁵	0.3996 ⁴	0.5005 ⁴
LS-GARCH volatility	0.2038 ²	0.2803 ³	0.1433 ³	0.1774 ³	0.3611 ²	0.4505 ²
LS-GARCH log variance	0.2465 ⁶	0.3549 ⁶	0.1674 ⁵	0.2158 ⁶	0.4340 ⁶	0.5423 ⁶
ML-GARCH variance	0.2050 ³	0.2782 ²	0.1387 ²	0.1726 ²	0.3629 ³	0.4519 ²
ML-GARCH volatility	0.2018 ¹	0.2766 ¹	0.1357 ¹	0.1695 ¹	0.3600 ¹	0.4486 ¹
ML-GARCH log variance	0.2291 ⁵	0.3228 ⁵	0.1546 ⁴	0.1953 ⁴	0.4091 ⁵	0.5104 ⁵
Middle 90% absolute returns						
SARV model	0.2223 ⁰	0.2971 ⁰	0.1296 ⁰	0.1636 ⁰	0.3421 ⁰	0.4289 ⁰
LS-GARCH variance	0.2282 ²	0.3025 ¹	0.1393 ⁴	0.1743 ⁴	0.3690 ⁴	0.4637 ⁴
LS-GARCH volatility	0.2245 ¹	0.3046 ²	0.1335 ¹	0.1682 ¹	0.3509 ²	0.4392 ²
LS-GARCH log variance	0.2482 ⁵	0.3390 ⁶	0.1510 ⁵	0.1914 ⁶	0.3908 ⁵	0.4891 ⁵
ML-GARCH variance	0.2289 ³	0.3115 ⁴	0.1352 ²	0.1710 ²	0.3499 ¹	0.4374 ¹
ML-GARCH volatility	0.2286 ²	0.3082 ³	0.1364 ³	0.1722 ³	0.3554 ³	0.4440 ³
ML-GARCH log variance	0.2510 ⁶	0.3318 ⁵	0.1506 ⁵	0.1886 ⁵	0.4044 ⁶	0.5048 ⁶
Highest 5% absolute returns						
SARV model	0.3016 ⁰	0.3840 ⁰	0.1301 ⁰	0.1627 ⁰	0.3057 ⁰	0.3813 ⁰
LS-GARCH variance	0.3150 ¹	0.4076 ¹	0.1359 ¹	0.1702 ¹	0.3121 ¹	0.3882 ¹
LS-GARCH volatility	0.3228 ²	0.4188 ²	0.1419 ²	0.1780 ²	0.3132 ¹	0.3907 ²
LS-GARCH log variance	0.3782 ⁵	0.4897 ⁵	0.1756 ⁵	0.2200 ⁵	0.4074 ⁶	0.5061 ⁶
ML-GARCH variance	0.3983 ⁶	0.5160 ⁶	0.1776 ⁵	0.2209 ⁵	0.3271 ⁴	0.4061 ⁴
ML-GARCH volatility	0.3557 ³	0.4601 ³	0.1703 ⁴	0.2126 ⁴	0.3176 ³	0.3956 ³
ML-GARCH log variance	0.3521 ³	0.4558 ³	0.1602 ³	0.2006 ³	0.3670 ⁵	0.4578 ⁵

The table reports summary statistics for the forecast errors for the GARCH and SARV models fitted in Table 1 conditioned on the level of lagged returns. In each of the artificial samples we separate the volatility (or variance or log variance) forecasts into three groups based on whether the previous day's absolute return lies in the lowest 5%, middle 90%, or highest 5% of the distribution. We then compute the MAE and RMSE for each group. The table reports the average MAEs and RMSEs across 1000 artificial samples. The superscripts indicate the number of alternative models within each group that have averages significantly smaller (at the 99% level) than the reported average.

in the relative performance across models. Again the LS-GARCH volatility model delivers the best overall performance. Looking at the percentage differences in performance between models, the percentages are smaller than they were for the one-step-ahead forecasts, but this is expected given the mean-reverting nature of the underlying process. As the forecast horizon increases, the forecasts approach a constant that represents the estimated value of the long-run variance, volatility, or

Table 4 Ten-step-ahead forecast errors for returns generated under stochastic volatility.

Forecasts	M₀: SARV variance		M₀: SARV volatility		M₀: SARV log variance	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
SARV model						
Ten-step-ahead forecasts	0.2775 ¹	0.3636 ⁰	0.1630 ⁰	0.2043 ⁰	0.4260 ⁰	0.5330 ⁰
GARCH models						
LS-GARCH variance	0.2813 ²	0.3671 ¹	0.1711 ²	0.2137 ²	0.4620 ⁵	0.5796 ⁵
LS-GARCH volatility	0.2764 ⁰	0.3687 ²	0.1652 ¹	0.2069 ¹	0.4343 ¹	0.5432 ¹
LS-GARCH log variance	0.2934 ⁵	0.3989 ⁶	0.1772 ⁵	0.2230 ⁶	0.4566 ⁴	0.5708 ⁴
ML-GARCH variance	0.2895 ⁴	0.3851 ⁴	0.1718 ⁴	0.2160 ⁴	0.4430 ²	0.5533 ²
ML-GARCH volatility	0.2833 ³	0.3726 ³	0.1709 ²	0.2145 ³	0.4434 ²	0.5541 ³
ML-GARCH log variance	0.2949 ⁶	0.3847 ⁴	0.1769 ⁵	0.2213 ⁵	0.4724 ⁶	0.5904 ⁶

The table reports summary statistics for the ten-step-ahead forecast errors for the GARCH and SARV models fitted in Table 1. In each of the artificial samples we use a particle filter to construct ten-step-ahead volatility (or variance or log variance) forecasts under the SARV model and we use the fitted parameter estimates to construct ten-step-ahead forecasts for the six GARCH models. We then calculate the MAE and RMSE for each set of forecasts in the artificial sample. The table reports the average MAEs and RMSEs across 1000 artificial samples. The superscripts indicate the number of alternative models that have averages significantly smaller (at the 99% level) than the reported average.

log variance. Thus the differences in performance should decline with the horizon if the models provide reasonable estimates of the long-run means.

The only surprise in Table 4 is that the LS-GARCH volatility model yields a significantly lower MAE than the SARV variance model for data generated under the SARV variance process. In general, we expect the SARV forecasts to outperform all of the GARCH forecasts since the SARV forecasts correspond on the optimal nonlinear filter. But this is not necessarily true, given the way in which we implement this filter. To make the computations feasible, we replace the true parameters with their posterior means (as opposed to integrating over their joint posterior) and we keep the number of particles in our Monte Carlo filtering algorithm relatively small. This second issue is probably the greater concern because the nonlinearities in the filter appear to be mild.

We evaluate the extent of the Monte Carlo error by running a limited set of simulations using a larger number of particles. Specifically, we know that doubling the number of particles should reduce the standard deviation of the Monte Carlo error by 29.3%. When we do this, we find that the MAEs and RMSEs for the one-step-ahead SARV forecasts decrease by 0.2% to 0.3%. Thus the Monte Carlo error is large enough to explain the failure of the SARV variance model to outperform the LS-GARCH volatility model. At the same time, however, the error does not seem large enough to alter the basic message of the simulation results.

Table 5 Out-of-sample forecast errors for returns generated under stochastic volatility.

Forecasts	M ₀ : SARV variance		M ₀ : SARV volatility		M ₀ : SARV log variance	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
SARV model						
Filtered estimates	0.2279	0.3043	0.1343	0.1688	0.3443	0.4310
One-step-ahead forecasts	0.2368 ⁰	0.3149 ⁰	0.1403 ⁰	0.1759 ⁰	0.3589 ⁰	0.4491 ⁰
GARCH models						
LS-GARCH variance	0.2440 ³	0.3216 ¹	0.1507 ⁴	0.1878 ⁴	0.3898 ⁴	0.4881 ⁴
LS-GARCH volatility	0.2385 ¹	0.3216 ¹	0.1421 ¹	0.1779 ¹	0.3653 ¹	0.4564 ¹
LS-GARCH log variance	0.2697 ⁵	0.3668 ⁶	0.1635 ⁵	0.2059 ⁶	0.4174 ⁵	0.5207 ⁵
ML-GARCH variance	0.2426 ³	0.3278 ⁴	0.1446 ²	0.1819 ²	0.3652 ¹	0.4561 ¹
ML-GARCH volatility	0.2414 ²	0.3240 ³	0.1445 ²	0.1814 ²	0.3675 ³	0.4586 ³
ML-GARCH log variance	0.2716 ⁵	0.3602 ⁵	0.1623 ⁵	0.2024 ⁵	0.4255 ⁶	0.5298 ⁶

The table reports summary statistics for the out-of-sample forecast errors for GARCH and SARV models fitted to returns generated from SARV models in variance, volatility, and log variance. For each SARV model we generate an artificial sample of 2000 returns as described in Table 1. We use the first 1000 of these returns to estimate the parameters of the SARV model and the parameters of the six GARCH models and we use the last 1000 returns to evaluate the out-of-sample forecast performance. We construct filtered and one-step-ahead volatility (or variance or log variance) forecasts under the SARV model and we generate one-step-ahead forecasts based on the parameter estimates for each GARCH model. We then compute the MAE and RMSE for each set of forecasts over the last 1000 returns. The table reports the average MAEs and RMSEs across 1000 artificial samples. The superscripts indicate the number of alternative models that have averages significantly smaller (at the 99% level) than the reported average. We exclude the filtered estimates from the significance tests.

Even if we could set the error to zero, we would only expect the MAEs and RMSEs for the one-step-ahead SARV forecasts to be around 1% lower than the values reported in Table 2.

Next, we look at how the models perform in an out-of-sample setting. Since updating the parameter estimates on a daily basis is too computationally intensive to be practical, we simply split the artificial samples in half and use the first 1000 observations to fit the models and the second 1000 observations to evaluate their forecasting performance. This should be conservative from the standpoint of evaluating the impact of parameter uncertainty, since daily updating should yield more accurate parameter estimates on average. Moreover, this approach should highlight any advantage that the SARV models have in small samples since it corresponds to using roughly four years of daily data to fit the models, which is much less than is typically used in practice.

Table 5 presents the results for the one-step-ahead forecasts. As expected, the MAEs and RMSEs are higher than those for the in-sample analysis. But in all other respects Table 5 looks similar to Table 2. We still find, for example, that the

LS-GARCH volatility model performs well across the board, with MAEs and RMSEs around 0.5% to 2.5% higher than those of the SARV model used to generate the data. Some of the rankings of the models change a bit from those in Table 2, but this is most likely due to increased sampling variation in the performance criteria. The bottom line is that the out-of-sample results mirror those from the in-sample analysis.

As a final robustness check, we replicate Tables 1 and 2 with the parameters of the DGPs set equal to the posterior means obtained by fitting the SARV models to returns on the British pound (see Table 6). To conserve space we simply summarize the main results. Once again, the average log Bayesian information criterion values for the GARCH variance and volatility models are similar, both of these models fit significantly better than the GARCH log variance model, and the rankings of the models with respect to forecasting performance are similar to those in Table 2. Thus we find nothing in the results that alters our conclusions from before.

Overall, the simulations suggest that GARCH models can play a useful role in the econometric analysis of SARV specifications. If we fit the appropriate GARCH model via least-squares, then we obtain a consistent estimator of the parameter vector of the MMS linear filter for the corresponding SARV model. In addition, the LS-GARCH volatility model delivers forecasts that have MAEs and RMSEs within 1% to 4% of those for the optimal nonlinear filter for each of the SARV DGPs. This bolsters the view that forecasts based on absolute returns display robustness, and it calls into question whether the advantage of the SARV models is sufficient to warrant the effort required to implement them. Of course, any or all of these findings could change when we confront the models with the real data. This is the subject to which we now turn.

4. THE RISK MANAGEMENT APPLICATION

The simulations reveal statistically significant differences in the performance of the GARCH and SARV forecasts. In this section we use a standard risk management application—estimating VaR—to gain perspective on the relative performance of these forecasts in practice. Specifically, we use the GARCH and SARV volatility forecasts to construct competing sets of conditional VaR estimates for a cross section of equity indexes and currencies, and then we evaluate which set of estimates is more consistent with the observed data. Our objective is to see whether differences like those observed in the simulations are apparent when analyzing real data and to assess whether any differences that emerge are large enough to be economically meaningful. We begin by describing the data and the model fitting results. Then we describe the results of the VaR analysis.

4.1 Data

The equity index data consists of daily prices for the S&P 500, NYSE Composite, NASDAQ Composite, FTSE All-Share (U.K.), and TOPIX (all First Section listed

Table 6 SARV model fitting results.

Asset/model	κ	95% HPD	ϕ	95% HPD	γ	95% HPD	BIC
Panel A: Currencies							
British pound							
SARV variance	0.004	(0.003, 0.006)	0.988	(0.983, 0.993)	0.092	(0.083, 0.101)	−5372.5
SARV volatility	0.014	(0.010, 0.019)	0.975	(0.967, 0.982)	0.054	(0.048, 0.061)	−5351.4
SARV log variance	−0.058	(−0.076, −0.041)	0.961	(0.950, 0.971)	0.328	(0.291, 0.369)	−5347.4
Canadian dollar							
SARV variance	0.002	(0.001, 0.002)	0.976	(0.969, 0.983)	0.043	(0.038, 0.047)	238.2
SARV volatility	0.007	(0.005, 0.009)	0.973	(0.964, 0.981)	0.022	(0.019, 0.025)	242.7
SARV log variance	−0.104	(−0.139, −0.073)	0.966	(0.955, 0.976)	0.233	(0.200, 0.267)	277.7
Deutsche mark							
SARV variance	0.010	(0.008, 0.013)	0.976	(0.968, 0.983)	0.110	(0.098, 0.122)	−6240.1
SARV volatility	0.022	(0.016, 0.029)	0.964	(0.953, 0.974)	0.062	(0.054, 0.070)	−6229.3
SARV log variance	−0.053	(−0.072, −0.038)	0.955	(0.941, 0.967)	0.271	(0.232, 0.311)	−6220.7
Japanese yen							
SARV variance	0.004	(0.003, 0.005)	0.990	(0.985, 0.994)	0.099	(0.091, 0.108)	−5703.4
SARV volatility	0.015	(0.011, 0.020)	0.974	(0.966, 0.981)	0.061	(0.055, 0.067)	−5651.0
SARV log variance	−0.089	(−0.112, −0.067)	0.940	(0.927, 0.952)	0.470	(0.427, 0.516)	−5594.3
Swiss franc							
SARV variance	0.014	(0.010, 0.018)	0.975	(0.967, 0.983)	0.122	(0.107, 0.138)	−7195.7
SARV volatility	0.022	(0.016, 0.030)	0.968	(0.957, 0.977)	0.064	(0.056, 0.074)	−7190.1
SARV log variance	−0.034	(−0.048, −0.023)	0.961	(0.948, 0.972)	0.229	(0.196, 0.266)	−7184.2
Panel B: Equities							
NASDAQ							
SARV variance	0.012	(0.009, 0.016)	0.987	(0.982, 0.991)	0.127	(0.113, 0.142)	−8669.5
SARV volatility	0.006	(0.003, 0.010)	0.992	(0.989, 0.996)	0.059	(0.053, 0.066)	−8677.7
SARV log variance	−0.012	(−0.019, −0.006)	0.981	(0.975, 0.987)	0.194	(0.171, 0.219)	−8564.5
NYSE							
SARV variance	0.015	(0.011, 0.019)	0.980	(0.973, 0.986)	0.114	(0.099, 0.130)	−8744.2
SARV volatility	0.012	(0.007, 0.017)	0.985	(0.979, 0.990)	0.052	(0.045, 0.060)	−8755.3
SARV log variance	−0.011	(−0.017, −0.006)	0.982	(0.974, 0.988)	0.142	(0.122, 0.164)	−8678.7
S&P 500							
SARV variance	0.016	(0.011, 0.021)	0.982	(0.975, 0.987)	0.118	(0.102, 0.135)	−9404.7
SARV volatility	0.012	(0.007, 0.017)	0.987	(0.981, 0.992)	0.054	(0.046, 0.062)	−9399.7
SARV log variance	−0.007	(−0.012, −0.003)	0.983	(0.976, 0.989)	0.135	(0.117, 0.156)	−9320.2
FTSE							
SARV variance	0.014	(0.010, 0.019)	0.984	(0.978, 0.990)	0.114	(0.098, 0.131)	−9488.7
SARV volatility	0.009	(0.005, 0.013)	0.990	(0.986, 0.994)	0.052	(0.045, 0.059)	−9486.9
SARV log variance	−0.007	(−0.011, −0.003)	0.983	(0.977, 0.989)	0.138	(0.119, 0.160)	−9436.7

continued

Table 6 (*continued*) SARV model fitting results.

Asset/model	κ	95% HPD	ϕ	95% HPD	γ	95% HPD	BIC
TOPIX							
SARV variance	0.017	(0.013, 0.020)	0.981	(0.976, 0.987)	0.156	(0.142, 0.170)	−8680.5
SARV volatility	0.015	(0.010, 0.020)	0.982	(0.977, 0.988)	0.081	(0.073, 0.089)	−8670.5
SARV log variance	−0.020	(−0.029, −0.012)	0.970	(0.961, 0.978)	0.266	(0.237, 0.299)	−8563.6

The table reports the results of fitting SARV models in variance, volatility, and log variance to daily returns for various currencies (panel A) and equity indexes (panel B). The models are fit using MCMC methods. For each model we report the mean of the posterior distribution for the constant term (κ), the AR(1) coefficient (ϕ), and the residual variance coefficient (γ), as well as the 95% highest posterior density (HPD) region for each estimate (i.e., the shortest interval that contains 95% of the posterior distribution) and the log Bayesian information criterion (BIC) evaluated at the mean parameter estimates. The sample period for the currencies is April 1973 through September 2000 (6904 returns for each currency) and the sample period for the equities is February 1971 through September 2000 (7318 to 7492 returns, depending on the index).

shares) indexes.¹⁶ The sample period for the equity data begins on February 5, 1971, the date that the last of these indexes (NASDAQ) was introduced, and extends through September 2000 (7318 to 7492 returns depending on the index). The currency data are daily U.S. dollar spot prices for British pounds (BP), Canadian dollars (CD), Deutsche marks (DM), Japanese yen (JY), and Swiss francs (SF).¹⁷ These currencies have historically had the most actively traded currency futures contracts at the Chicago Mercantile Exchange. The currency sample begins in April 1973, after the collapse of the Bretton Woods monetary system, and runs through September 2000 (6904 returns).

4.2 Model Fitting Results and VaR Estimates

We fit the models to the equity index and currency returns using a two-step procedure that is common in the empirical literature. First, we remove any serial correlation in the returns by fitting an AR(1) model for each asset. This yields estimates of μ and δ . Then we fit the SARV and GARCH models to the residuals using the econometric methods described in Sections 3.2 and 3.3. The SARV model fitting results are shown in Table 6. The results for the GARCH models are shown in Table 7. In each case, the currencies are in panel A and the equity indexes are in panel B.

¹⁶ The NYSE and NASDAQ data are from exchange websites (www.marketdata.nasdaq.com/mr4b.html and www.nyse.com/marketinfo/marketinfo.html), and the S&P 500, FTSE, and TOPIX data are from Datastream. For the S&P 500, NYSE, and NASDAQ indexes, we use prices measured in U.S. dollars, while the FTSE and TOPIX index prices are measured in British pounds and Japanese yen, respectively.

¹⁷ These data are from the Federal Reserve's website (www.federalreserve.gov/releases/H10/hist/). We use the "noon buying rates in New York for cable transfers payable in foreign currencies," as reported in the weekly H.10 release. The prices for Deutsche marks after January 1, 1999, are implied from the spot prices for euros using the irrevocable DM/euro conversion rate.

Table 7 GARCH model fitting results.

Asset/model	ω	S.E.	α	S.E.	β	S.E.	ν	S.E.	BIC
Panel A: Currencies									
British pound									
LS-GARCH variance	0.007	0.019	0.053	0.007	0.929	0.026	—	—	—
LS-GARCH volatility	0.007	0.017	0.078	0.012	0.924	0.028	—	—	—
LS-GARCH log variance	0.054	0.053	0.056	0.012	0.934	0.030	—	—	—
ML-GARCH variance	0.001	0.001	0.068	0.010	0.930	0.010	1.129	0.015	−5358.5
ML-GARCH volatility	0.007	0.006	0.100	0.013	0.914	0.014	1.125	0.018	−5333.1
ML-GARCH log variance	0.059	0.035	0.050	0.008	0.928	0.021	1.092	0.019	−5361.1
Canadian dollar									
LS-GARCH variance	0.003	0.004	0.082	0.004	0.874	0.023	—	—	—
LS-GARCH volatility	0.007	0.008	0.107	0.010	0.885	0.029	—	—	—
LS-GARCH log variance	0.014	0.131	0.053	0.012	0.930	0.043	—	—	—
ML-GARCH variance	0.000	0.000	0.065	0.009	0.935	0.020	1.374	0.023	227.4
ML-GARCH volatility	0.005	0.006	0.114	0.014	0.894	0.028	1.367	0.025	252.2
ML-GARCH log variance	0.001	0.093	0.055	0.007	0.916	0.032	1.293	0.024	163.4
Deutsche mark									
LS-GARCH variance	0.023	0.030	0.073	0.006	0.874	0.046	—	—	—
LS-GARCH volatility	0.016	0.026	0.092	0.012	0.900	0.038	—	—	—
LS-GARCH log variance	0.049	0.062	0.057	0.012	0.924	0.040	—	—	—
ML-GARCH variance	0.004	0.007	0.083	0.015	0.911	0.025	1.272	0.027	−6222.5
ML-GARCH volatility	0.012	0.014	0.100	0.016	0.906	0.029	1.267	0.027	−6213.9
ML-GARCH log variance	0.049	0.039	0.049	0.008	0.922	0.034	1.236	0.025	−6261.2
Japanese yen									
LS-GARCH variance	0.016	0.032	0.061	0.003	0.903	0.036	—	—	—
LS-GARCH volatility	0.008	0.021	0.077	0.010	0.924	0.031	—	—	—
LS-GARCH log variance	0.062	0.062	0.062	0.011	0.927	0.029	—	—	—
ML-GARCH variance	0.000	0.001	0.045	0.002	0.955	0.025	1.032	0.016	−5611.2
ML-GARCH volatility	0.010	0.006	0.130	0.014	0.889	0.015	1.027	0.016	−5594.0
ML-GARCH log variance	0.080	0.040	0.061	0.008	0.909	0.023	0.953	0.015	−5664.9
Swiss franc									
LS-GARCH variance	0.042	0.032	0.095	0.006	0.830	0.028	—	—	—
LS-GARCH volatility	0.018	0.028	0.088	0.012	0.903	0.038	—	—	—
LS-GARCH log variance	0.045	0.054	0.048	0.012	0.936	0.045	—	—	—
ML-GARCH variance	0.006	0.011	0.073	0.015	0.918	0.032	1.319	0.028	−7178.5
ML-GARCH volatility	0.013	0.018	0.088	0.016	0.915	0.030	1.311	0.028	−7175.4
ML-GARCH log variance	0.048	0.031	0.042	0.008	0.936	0.034	1.266	0.026	−7230.1
Panel B: Equities									
NASDAQ									
LS-GARCH variance	0.087	0.074	0.232	0.002	0.682	0.004	—	—	—
LS-GARCH volatility	0.019	0.017	0.168	0.005	0.843	0.009	—	—	—

continued

Table 7 (continued) GARCH model fitting results.

Asset/model	ω	S.E.	α	S.E.	β	S.E.	ν	S.E.	BIC
LS-GARCH log variance	0.047	0.033	0.040	0.011	0.954	0.035	—	—	—
ML-GARCH variance	0.012	0.013	0.106	0.012	0.880	0.037	1.401	0.024	-8566.5
ML-GARCH volatility	0.014	0.018	0.112	0.011	0.898	0.029	1.386	0.023	-8584.3
ML-GARCH log variance	0.076	0.022	0.055	0.006	0.932	0.023	1.290	0.020	-8711.3
NYSE									
LS-GARCH variance	0.140	0.201	0.131	0.003	0.694	0.012	—	—	—
LS-GARCH volatility	0.021	0.017	0.108	0.006	0.886	0.014	—	—	—
LS-GARCH log variance	0.032	0.048	0.029	0.011	0.964	0.049	—	—	—
ML-GARCH variance	0.008	0.019	0.058	0.008	0.932	0.043	1.359	0.020	-8652.0
ML-GARCH volatility	0.010	0.023	0.068	0.010	0.936	0.034	1.348	0.019	-8655.9
ML-GARCH log variance	0.048	0.022	0.035	0.008	0.952	0.032	1.274	0.015	-8732.9
S&P 500									
LS-GARCH variance	0.181	0.240	0.128	0.003	0.680	0.015	—	—	—
LS-GARCH volatility	0.019	0.019	0.099	0.006	0.898	0.016	—	—	—
LS-GARCH log variance	0.033	0.042	0.029	0.012	0.965	0.047	—	—	—
ML-GARCH variance	0.008	0.022	0.056	0.008	0.935	0.041	1.381	0.020	-9304.6
ML-GARCH volatility	0.010	0.025	0.066	0.010	0.938	0.033	1.367	0.019	-9310.4
ML-GARCH log variance	0.046	0.019	0.033	0.007	0.957	0.031	1.286	0.014	-9390.9
FTSE									
LS-GARCH variance	0.187	0.055	0.343	0.001	0.469	0.004	—	—	—
LS-GARCH volatility	0.028	0.018	0.159	0.005	0.841	0.013	—	—	—
LS-GARCH log variance	0.047	0.038	0.040	0.011	0.949	0.044	—	—	—
ML-GARCH variance	0.015	0.022	0.086	0.013	0.898	0.039	1.629	0.014	-9446.4
ML-GARCH volatility	0.018	0.025	0.101	0.008	0.901	0.032	1.623	0.015	-9460.8
ML-GARCH log variance	0.075	0.020	0.059	0.006	0.916	0.029	1.479	0.016	-9613.1
TOPIX									
LS-GARCH variance	0.625	0.081	0.313	0.002	0.063	0.006	—	—	—
LS-GARCH volatility	0.034	0.024	0.168	0.005	0.825	0.018	—	—	—
LS-GARCH log variance	0.061	0.029	0.054	0.012	0.936	0.035	—	—	—
ML-GARCH variance	0.012	0.010	0.136	0.014	0.858	0.021	1.283	0.017	-8601.8
ML-GARCH volatility	0.016	0.014	0.144	0.013	0.873	0.021	1.278	0.017	-8589.9
ML-GARCH log variance	0.102	0.020	0.071	0.008	0.911	0.023	1.205	0.016	-8696.4

The table reports the results of fitting LS-GARCH and ML-GARCH models in variance, volatility, and log variance to daily returns for various currencies (panel A) and equity indexes (panel B). ω is the constant term in the specifications, α is the coefficient on e_{t-1}^2 , $|e_t|$, or $\log e_t^2$, β is the coefficient on the lagged variance, volatility, or log variance, and ν is the GED constant. We also report standard errors (S.E.) for these estimates and the log Bayesian information criterion (BIC) for the fitted models. The sample period for the currencies is April 1973 through September 2000 (6904 returns for each currency) and the sample period for the equities is February 1971 through September 2000 (7318 to 7492 returns, depending on the index).

First consider the SARV results. Table 6 reports the posterior mean of each parameter, the 95% highest posterior density (HPD) interval for each parameter (i.e., the shortest interval containing 95% of the posterior), and the log Bayesian information criterion for the model. If we use the log Bayesian information criterion values to rank the models on an asset-by-asset basis, we get some preliminary insights regarding their performance. The log variance model, which is the most popular SARV specification in the empirical literature, delivers the best fit for all five currencies and all five equity indexes. Next comes the volatility model, which fits better than the square-root model for all assets except the NASDAQ and NYSE indexes. The implications of the models concerning the volatility of volatility probably explain at least some of these differences. The log variance model, for example, implies that $\text{var}(\sigma_{t+1}^2 | \sigma_t^2) \propto \sigma_t^{4\phi}$, whereas the square-root model implies that $\text{var}(\sigma_{t+1}^2 | \sigma_t^2) \propto \sigma_t^2$.

Not surprisingly, all of the SARV models suggest that volatility displays strong persistence. The mean of ϕ is 0.94 or greater in every case and shows little variation across assets for a given model. There is more variation in the posterior mean of γ across assets. It suggests that the volatility of volatility is somewhat smaller for the CD than for the other currencies, and somewhat larger for NASDAQ and TOPIX than for the other equity indexes. In general, however, the posteriors for a given model are similar across assets, and the HPD intervals indicate that the data are highly informative about the parameters.

Now consider the GARCH results. Table 7 reports the estimate of each parameter, the standard error of this estimate, and the log Bayesian information criterion for the model in the case of the maximum-likelihood specifications. Every pair of α and β estimates indicates a high level of persistence, regardless of the model and asset, although the maximum-likelihood estimates for the variance and volatility models tend to suggest more persistence than the least-squares estimates. Similarly the least-squares estimates of α for these models tend to be larger than the maximum-likelihood estimates. This implies that the least-squares forecasts will exhibit a larger response to a given return shock than the maximum-likelihood forecasts.

All of the ν estimates in Table 7 are substantially less than two. They range from 0.953 for the JY log variance model to 1.629 for the FTSE variance model. This is indicative of conditional return distributions with much heavier tails than a normal, which is characteristic of data generated by a SARV process. However, it is unlikely that the currency and index returns are generated by one of the SARV models considered here. This follows by noting that the ν estimates for NYSE are significantly smaller than the average ν estimates reported in Table 1.

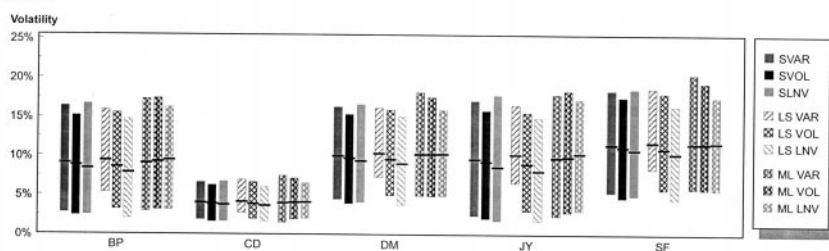
On the whole the ML-GARCH models fit about as well the SARV models. Typically the GARCH variance and volatility models produce higher log Bayesian information criterion values than the SARV variance and volatility models, but the opposite is true for the log variance models. The only exception is for the CD, where the GARCH variance model fits worse than the SARV variance model. Overall the SARV log variance model and the ML-GARCH volatility model each produce the highest log Bayesian information criterion for four of the assets, and

the ML-GARCH variance model produces the highest log Bayesian information criterion for the other two. In making these comparisons we should note that the log Bayesian information criterion values for the SARV models are based on the posterior means of the MCMC iterates rather than maximum-likelihood estimates of the parameters. Consequently they probably exhibit a small downward bias.

If we compare goodness-of-fit across the ML-GARCH models, we find that the results differ with asset class. For the currencies, the volatility model produces the highest log Bayesian information criterion value, followed by the variance model and then the log variance model. For the equity indexes, the ordering of the variance model and the volatility model is reversed in every case except for the TOPIX. Although the relatively poor performance of the GARCH log variance model stands in sharp contrast with the good performance of the SARV log variance model, this is consistent with our findings in the simulations.

Before turning to the VaR analysis, we take a brief look at the properties of the one-step-ahead volatility forecasts produced by the fitted models. Figure 1 summarizes the distribution of these forecasts. Specifically, the bars in the figure

Panel A: Currencies



Panel B: Equities

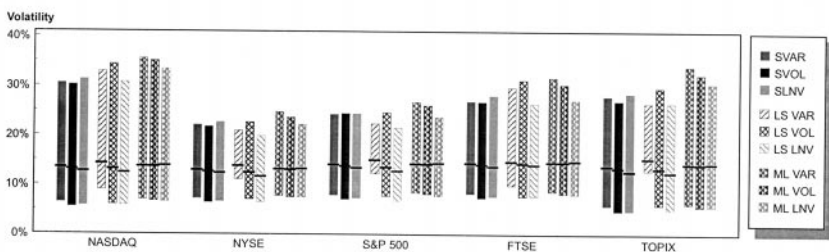


Figure 1 Summary statistics for the daily volatility forecasts. The figure shows summary statistics for the daily volatility forecasts under the fitted SARV, LS-GARCH, and ML-GARCH models. Panel A shows the results for the currencies: British pounds (BP), Canadian dollars (CD), Deutsche marks (DM), Japanese yen (JY), and Swiss francs (SF). Panel B shows the results for the equity indexes: NASDAQ, NYSE, S&P 500, FTSE, and TOPIX. SVAR, SVOL, and SLNV denote the forecasts from the SARV variance, volatility, and log variance models; LS VAR, LS VOL, and LS LNV denote the forecasts from the LS-GARCH models, and ML VAR, ML VOL, and ML LNV denote the forecasts from the ML-GARCH models. The variance, volatility, or log variance forecasts under each model are converted to volatilities and annualized on the basis of 252 trading days per year. In the graph, the bars for each asset show the range of the middle 95% of the distribution of the daily forecasts under each model and the line through each bar shows the mean.

represent the middle 95% range of forecasts, and the line through each bar indicates the mean. The nine sets of forecasts for each asset are considered in the following order: SARV in variance, volatility, and log variance, LS-GARCH in variance, volatility, and log variance, and ML-GARCH in variance, volatility, and log variance.

Among the SARV forecasts, those produced by the log variance model usually have the largest range. This suggests that the superior fit of this model could to some extent reflect its ability to track the extremes in volatility better than the other two SARV models. We also see a distinct pattern in the means. The variance model produces forecasts with the highest mean, followed by the volatility model, and then the log variance model. A similar pattern in the means is evident for the LS-GARCH models, which is consistent with their role as MMS linear filters for the SARV models. However, the volatility model — not the log variance model — tends to produce forecasts with the largest range. The results for the ML-GARCH models are quite different. The means of the three sets of forecasts are nearly the same, but the variance model produces the largest range, followed by the volatility model, and then the log variance model.

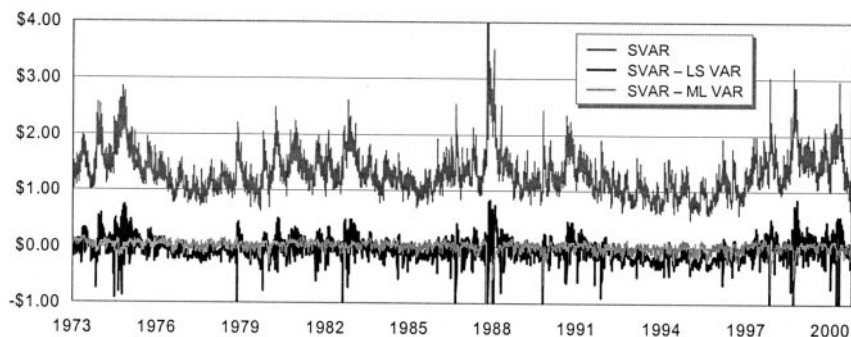
To form our conditional VaR estimates we use the one-step-ahead volatility forecasts to find the quantile regression parameters that minimize the sample counterpart of the Kullback-Leibler information criterion. These VaR estimates are for a long position for the case of $\pi = 5\%$. Figures 2 and 3 show the estimates produced by the SARV models for a representative equity index (NYSE) and a representative currency (BP). We assume a constant \$100 position in each asset for the purposes of illustration. Panel A plots the estimates for the variance model, and panels B and C plot the estimates for the volatility and log variance models. For comparison we also plot the differences between the SARV and GARCH estimates for each model.

It is clear that the estimated VaRs produced by the SARV models display quite a bit of variation over the sample period. For the NYSE, they range from a low of 33 cents in 1995 to a high of \$7.57 (not shown in figure) in 1987. For the BP, the range is from 21 cents in 1977 to \$2.60 in 1985. In contrast, the estimated VaRs based on a constant variance model are about \$1.39 and \$1.01 for the NYSE and the BP, respectively. It is also apparent that the differences between the SARV and GARCH estimates are most dramatic for the log variance models, which is exactly what we would predict based on the model fitting results.

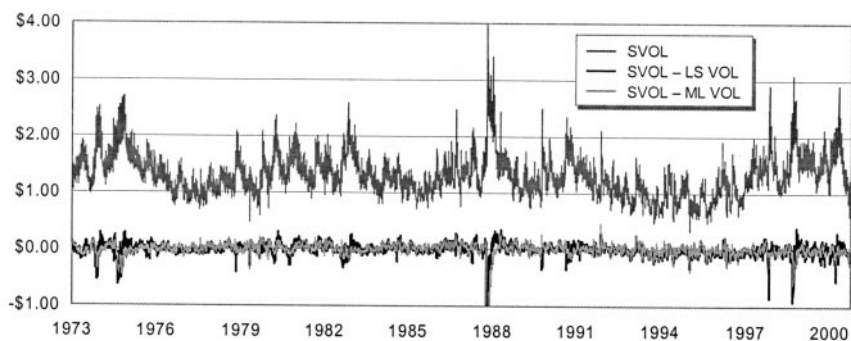
4.3 VaR Specification Tests and Comparisons

Table 8 presents the results of our first specification test. It reports the estimated intercept and slope coefficients for the quantile regression, the estimated probability of an exceedence given that an exceedence did not occur the previous day, and the estimated probability of an exceedence given that an exceedence did occur the previous day. The table also reports the likelihood ratio statistic for the joint hypothesis that the unconditional probability of an exceedence equals π and that the exceedences are independent through time. For comparison we also include results for the unconditional VaR estimates produced by a constant variance

Panel A: Variance Models



Panel B: Volatility Models



Panel C: Log Variance Models

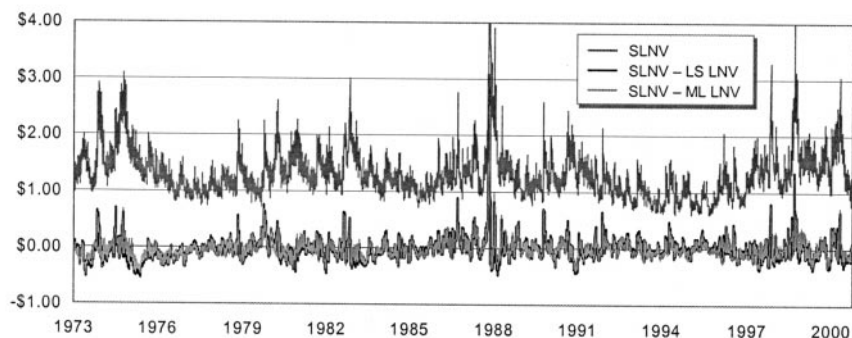
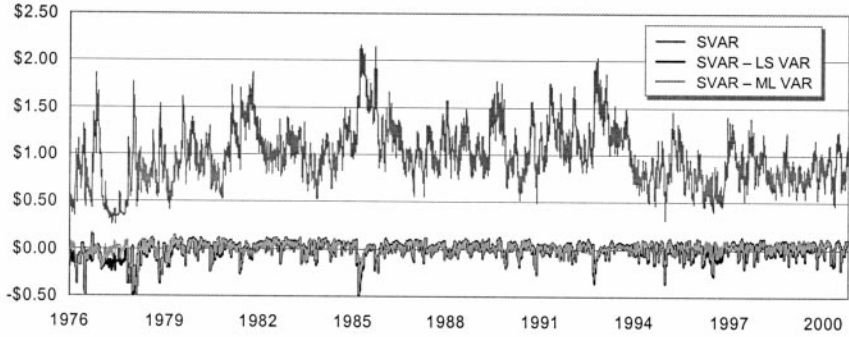
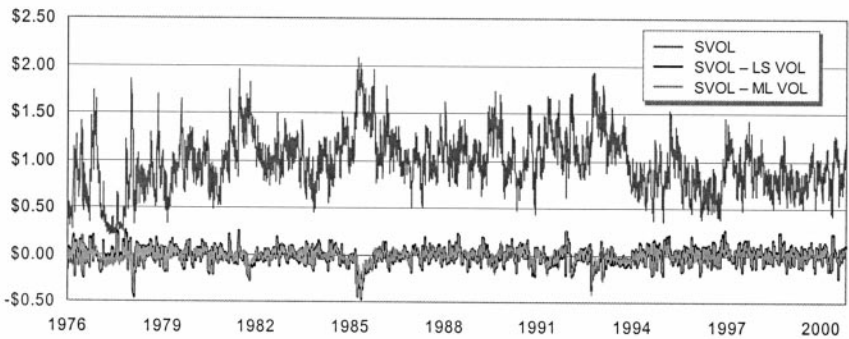


Figure 2 Daily value-at-risk estimates for the NYSE index. The figure shows the daily 5% value-at-risk estimates under the quantile regressions in Table 8 for \$100 invested in the NYSE index. Panel A shows the estimates for the SARV variance model (SVAR) and the differences between the SARV estimates and the estimates for the LS-GARCH variance (LS VAR) and ML-GARCH variance (ML VAR) models. Panel B shows the estimates for the SARV, LS-GARCH, and ML-GARCH volatility models (SVOL, LS VOL, and ML VOL). Panel C shows the estimates for the SARV, LS-GARCH, and ML-GARCH log variance models (SLNV, LS LNV, and ML LNV).

Panel A: Variance Models



Panel B: Volatility Models



Panel C: Log Variance Models

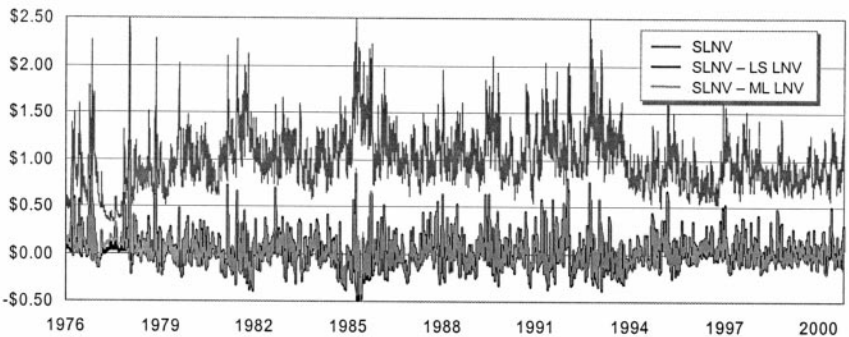


Figure 3 Daily value-at-risk estimates for British pounds. The figure shows the daily 5% value-at-risk estimates under the quantile regressions in Table 8 for \$100 invested in British pounds. Panel A shows the estimates for the SARV variance model (SVAR) and the differences between the SVAR estimates and the estimates for the LS-GARCH variance (LS VAR) and ML-GARCH variance (ML VAR) models. Panel B shows the estimates for the SARV, LS-GARCH, and ML-GARCH volatility models (SVOL, LS VOL, and ML VOL). Panel C shows the estimates for the SARV, LS-GARCH, and ML-GARCH log variance models (SLNV, LS LNV, and ML LNV).

Table 8 Value-at-risk quantile regression results.

Asset/model	SARV models					LS-GARCH models					ML-GARCH models				
	$q_0 \times 100$	q_1	$\hat{\pi}_{1 0}$	$\hat{\pi}_{1 1}$	LR	$q_0 \times 100$	q_1	$\hat{\pi}_{1 0}$	$\hat{\pi}_{1 1}$	LR	$q_0 \times 100$	q_1	$\hat{\pi}_{1 0}$	$\hat{\pi}_{1 1}$	LR
Panel A: Currencies															
British Pound	−1.01	—	0.047	0.109	19.3**										
Variance	−0.15	−1.40	0.049	0.069	2.3	0.04	−1.69	0.049	0.069	2.3	−0.20	−1.32	0.049	0.066	1.6
Volatility	−0.12	−1.50	0.050	0.056	0.2	−0.09	−1.60	0.048	0.078	4.9	−0.16	−1.36	0.049	0.072	3.0
Log variance	−0.27	−1.33	0.050	0.050	0.0	−0.24	−1.44	0.048	0.093	10.9**	−0.10	−1.42	0.048	0.091	9.9**
Canadian dollar	−0.43	—	0.047	0.103	15.8**										
Variance	−0.06	−1.45	0.048	0.081	5.9	0.01	−1.70	0.048	0.072	3.2	−0.11	−1.21	0.048	0.091	9.7**
Volatility	−0.07	−1.44	0.049	0.078	4.8	−0.07	−1.47	0.048	0.084	7.1*	−0.09	−1.28	0.048	0.087	8.3*
Log variance	−0.10	−1.36	0.049	0.075	3.8	−0.13	−1.27	0.047	0.103	16.0**	−0.02	−1.56	0.046	0.099	14.3**
Deutsche mark	−1.03	—	0.048	0.084	7.1*										
Variance	−0.11	−1.47	0.049	0.060	0.6	0.25	−2.00	0.049	0.059	0.6	−0.16	−1.38	0.049	0.066	1.6
Volatility	−0.16	−1.44	0.050	0.056	0.3	−0.03	−1.67	0.049	0.066	1.6	−0.10	−1.45	0.049	0.066	1.6
Log variance	−0.22	−1.39	0.049	0.060	0.6	−0.10	−1.62	0.049	0.072	3.0	0.07	−1.68	0.049	0.075	3.8
Japanese yen	−1.02	—	0.048	0.097	12.6**										
Variance	−0.13	−1.42	0.050	0.050	0.0	0.29	−2.01	0.050	0.053	0.1	−0.09	−1.46	0.049	0.059	0.6
Volatility	−0.10	−1.49	0.050	0.047	0.1	−0.02	−1.68	0.049	0.059	0.6	−0.16	−1.31	0.050	0.053	0.1
Log variance	−0.37	−1.14	0.050	0.050	0.0	−0.21	−1.51	0.049	0.081	5.8	−0.09	−1.39	0.048	0.091	9.7**
Swiss franc	−1.17	—	0.048	0.094	11.1**										
Variance	−0.14	−1.40	0.049	0.063	1.0	0.18	−1.82	0.050	0.050	0.0	−0.18	−1.34	0.050	0.053	0.1
Volatility	−0.12	−1.46	0.049	0.063	1.0	−0.03	−1.62	0.050	0.053	0.1	−0.05	−1.50	0.049	0.059	0.6
Log variance	−0.23	−1.35	0.050	0.050	0.0	−0.16	−1.52	0.048	0.075	4.0	0.00	−1.56	0.049	0.075	3.9

continued

Table 8 (continued) Value-at-risk quantile regression results.

Asset/model	SARV models					LS-GARCH models					ML-GARCH models				
	$q_0 \times 100$	q_1	$\hat{\pi}_{1 0}$	$\hat{\pi}_{1 1}$	LR	$q_0 \times 100$	q_1	$\hat{\pi}_{1 0}$	$\hat{\pi}_{1 1}$	LR	$q_0 \times 100$	q_1	$\hat{\pi}_{1 0}$	$\hat{\pi}_{1 1}$	LR
Panel B: Equities															
NASDAQ	-1.57	—	0.045	0.152	54.7**										
Variance	0.04	-1.84	0.048	0.078	5.2	0.13	-1.87	0.050	0.060	0.7	-0.08	-1.68	0.048	0.078	5.2
Volatility	-0.10	-1.73	0.048	0.083	7.3*	-0.16	-1.67	0.049	0.072	3.3	-0.07	-1.70	0.048	0.086	8.6*
Log variance	-0.04	-1.87	0.048	0.080	6.0	-0.21	-1.65	0.048	0.092	11.1**	-0.17	-1.53	0.048	0.092	11.3**
NYSE	-1.39	—	0.047	0.111	22.4**										
Variance	-0.11	-1.50	0.048	0.092	11.2**	0.24	-1.84	0.049	0.074	3.8	-0.23	-1.32	0.048	0.094	12.4**
Volatility	-0.19	-1.42	0.047	0.097	14.1**	-0.21	-1.43	0.048	0.080	6.0*	-0.15	-1.43	0.047	0.100	15.5**
Log variance	-0.16	-1.49	0.048	0.077	5.1	-0.17	-1.58	0.047	0.101	16.0**	-0.17	-1.40	0.048	0.094	12.4**
S&P 500	-1.51	—	0.046	0.120	28.4**										
Variance	-0.04	-1.57	0.048	0.080	6.1*	0.31	-1.88	0.049	0.063	1.3	-0.23	-1.35	0.048	0.083	7.2*
Volatility	-0.14	-1.47	0.048	0.080	6.1*	-0.21	-1.44	0.049	0.069	2.5	-0.12	-1.46	0.048	0.086	8.6*
Log variance	-0.12	-1.55	0.048	0.078	5.2	-0.14	-1.60	0.048	0.083	7.4*	-0.10	-1.46	0.048	0.089	9.9**
FTSE	-1.53	—	0.046	0.132	37.3**										
Variance	0.18	-1.86	0.049	0.072	3.2	-0.12	-1.50	0.051	0.029	4.1	-0.07	-1.54	0.049	0.077	5.1
Volatility	0.05	-1.73	0.049	0.074	4.1	-0.18	-1.45	0.049	0.063	1.2	-0.05	-1.58	0.049	0.069	2.5
Log variance	-0.05	-1.66	0.049	0.077	5.1	0.15	-1.89	0.048	0.092	11.1**	0.09	-1.71	0.049	0.077	4.9
TOPIX	-1.51	—	0.043	0.176	77.3**										
Variance	0.03	-1.66	0.048	0.097	13.3**	0.45	-2.01	0.049	0.076	4.5	-0.12	-1.46	0.048	0.094	11.8**
Volatility	-0.06	-1.59	0.048	0.079	5.5	0.04	-1.77	0.048	0.079	5.5	-0.05	-1.54	0.048	0.079	5.5
Log variance	-0.10	-1.63	0.048	0.085	8.0*	-0.05	-1.72	0.047	0.103	16.5**	-0.06	-1.49	0.047	0.100	14.9**

The table reports the results of fitting quantile regressions to the volatility forecasts from our SARV, LS-GARCH, and ML-GARCH models to estimate the daily $\pi=5\%$ value-at-risk for long positions in various currencies (panel A) and equity indexes (panel B). We report the constant (q_0) and slope coefficients (q_1) for the quantile regressions and then, based on these estimates, the conditional probabilities of a value-at-risk exceedence given an exceedence ($\hat{\pi}_{1|1}$) or a nonexceedence ($\hat{\pi}_{1|0}$) on the previous day. LR is the likelihood ratio statistic for the joint hypothesis that the unconditional probability of an exceedence (π_1) equals π and that $\pi_{1|1} = \pi_{1|0} = \pi_1$. The statistic is asymptotically χ^2_2 under the null, and we use * and ** to denote rejections at the 5% and 1% levels. We also report the results for a constant variance model (i.e., $q_1 = 0$). The sample period for the currencies is April 1973 through September 2000 and the sample period for the equities is February 1971 through September 2000. We withhold the first 500 observations for each asset to initialize the sequence of volatility forecasts.

model. A single (double) asterisk indicates that the null can be rejected at the 5% (1%) significance level.

As might be expected, we can reject the constant variance model at the 1% level for every asset. For the other models, the rejections are much less frequent. None of the SARV models is rejected for any of the currencies (panel A). However, both the SARV variance and volatility models are rejected for three of the equities (panel B). The SARV log variance model delivers the best performance of the SARV models, with only one rejection (TOPIX). Turning to the GARCH results, we find that the maximum-likelihood models have a higher rejection rate than the least-squares models. The log variance model performs the worst: the least-squares (maximum-likelihood) version is rejected for two (three) of the currencies and all (four) of the equities, while the maximum-likelihood variance and volatility models are both rejected for one currency and three of the equities. The least-squares variance model performs the best, with no rejections for any of the assets, and the least-squares volatility model also performs well, with only two rejections at the 5% level.

Tables 9 and 10 provide additional evidence on the performance of the VaR estimates. We obtain these results by fitting overidentified versions of the quantile regressions, where the instrument vector for each regression includes a constant and the one-step-ahead volatility forecasts from the nine competing specifications. The first column in each table reports the results of our second specification test, which focuses on the orthogonality restrictions implied by conditional efficiency. The remaining columns report the asymptotic *t*-ratios that we use for the pairwise model comparisons. As in Table 8, a single (double) asterisk indicates statistical significance at the 5% (1%) level.

We begin with the SARV results in Table 9. As before, we can reject the constant variance model at the 1% level for every asset. However, we now find evidence that the SARV models are misspecified for the currencies (panel A). In fact, we can reject all of the SARV models for every currency except the CD. Apparently, for the currencies, the orthogonality restrictions pose a bigger challenge to the SARV models than the independence restriction considered earlier. But this does not seem to be true for the equities. Although we can reject all three models for the NASDAQ at the 5% level, the only other rejection in panel B is the volatility model for the FTSE.

Despite the many rejections in panel A, there is evidence that some of the SARV models do better than others for specific currencies. In particular, the SARV variance and volatility models outperform the SARV log variance model for the BP and JY. There are also some significant differences in performance between the SARV and GARCH models. All of the SARV models, for example, outperform both the LS- and ML-GARCH log variance models for the CD. However, it is not clear that the SARV models as a group do better than the GARCH models as a group. For instance, the LS- and ML-GARCH volatility models outperform the SARV log variance model for both the JY and SF.

The results for the equity indexes, on the other hand, provide more support for the view that SARV models are preferred to GARCH models. None of the differences in performance between the SARV models is statistically significant,

Table 9 SARV model value-at-risk orthogonality tests.

Asset	Model	κ -stat.	Δ -statistics									
			vs. CNS	vs. SARV models			vs. LS-GARCH models			vs. ML-GARCH models		
				VAR	VOL	LNV	VAR	VOL	LNV	VAR	VOL	LNV
Panel A: Currencies												
BP	Constant variance	149.5**										
	SARV variance	18.7*	5.00**	—	−0.08	2.09*	−0.03	0.00	0.45	−0.77	−0.43	0.17
	SARV volatility	18.4*	4.86**	0.08	—	2.13*	0.01	0.02	0.46	−0.65	−0.30	0.18
	SARV log variance	34.0**	4.41**	−2.09*	−2.13*	—	−1.52	−1.06	−0.58	−2.05*	−1.78	−0.82
CD	Constant variance	151.0**										
	SARV variance	13.1	5.33**	—	−0.80	−0.57	−0.42	0.65	3.00**	0.80	0.67	3.06**
	SARV volatility	11.2	5.33**	0.80	—	−0.14	−0.11	0.90	2.91**	1.03	0.92	3.00**
	SARV log variance	10.5	5.27**	0.57	0.14	—	−0.03	0.91	2.73**	0.90	0.84	2.85**
DM	Constant variance	207.4**										
	SARV variance	27.1**	4.86**	—	0.42	1.60	2.09*	−0.91	−1.20	0.20	−1.01	−1.19
	SARV volatility	30.4**	4.51**	−0.42	—	1.13	1.20	−1.00	−1.30	−0.17	−1.05	−1.47
	SARV log variance	38.7**	4.33**	−1.60	−1.13	—	0.53	−1.75	−1.87	−0.93	−1.77	−1.87
JY	Constant variance	237.9**										
	SARV variance	21.0**	5.61**	—	1.12	3.48**	1.30	−0.80	1.68	−1.59	0.59	2.04*
	SARV volatility	29.1**	5.39**	−1.12	—	3.10**	0.34	−1.51	1.13	−1.96*	−0.38	1.49
	SARV log variance	69.1**	4.41**	−3.48**	−3.10**	—	−2.75**	−3.52**	−1.32	−3.74**	−3.31**	−1.08

SF	Constant variance	198.3**										
	SARV variance	23.7**	4.84**	—	0.40	1.45	2.90**	−1.38	0.23	−0.89	−1.82	0.23
	SARV volatility	25.6**	4.64**	−0.40	—	1.12	2.57*	−1.42	0.07	−1.09	−1.87	0.07
	SARV log variance	36.2**	4.63**	−1.45	−1.12	—	1.89	−2.13*	−0.62	−1.81	−2.29*	−0.61
Panel B: Equities												
NASDAQ	Constant variance	296.8**										
	SARV variance	19.0*	7.27**	—	−0.55	0.07	0.66	0.86	2.75**	−0.83	0.40	2.92**
	SARV volatility	16.6*	7.54**	0.55	—	0.40	0.87	1.14	3.01**	−0.51	0.62	3.20**
	SARV log variance	19.4*	7.38**	−0.07	−0.40	—	0.62	0.84	2.47*	−1.22	0.29	2.72**
NYSE	Constant variance	150.8**										
	SARV variance	9.9	5.54**	—	−0.37	−0.42	2.70**	0.71	1.50	−1.12	−0.44	1.40
	SARV volatility	8.4	5.49**	0.37	—	−0.14	2.81**	0.82	1.60	−0.72	−0.19	1.49
	SARV log variance	7.5	5.27**	0.42	0.14	—	2.98**	1.12	1.42	−0.60	−0.06	1.38
S&P 500	Constant variance	163.3**										
	SARV variance	14.1	5.29**	—	−0.60	−0.62	2.94**	0.61	1.45	−1.25	−1.10	1.57
	SARV volatility	10.9	5.48**	0.60	—	−0.06	3.24**	1.01	1.75	−0.76	−0.64	1.87
	SARV log variance	10.5	5.58**	0.62	0.06	—	3.69**	1.41	1.69	−0.84	−0.63	1.82
FTSE	Constant variance	272.1**										
	SARV variance	15.3	7.70**	—	0.70	−1.26	3.31**	0.65	2.93**	−0.80	0.19	2.96**
	SARV volatility	18.8*	7.88**	−0.70	—	−1.60	3.12**	0.26	2.96**	−1.11	−0.21	2.88**
	SARV log variance	7.6	7.94**	1.26	1.60	—	3.92**	1.81	3.24**	0.20	1.50	3.44**

continued

Table 9 (continued) SARV model value-at-risk orthogonality tests.

			Δ -statistics									
Asset	Model	κ -stat.	vs.	vs. SARV models			vs. LS-GARCH models			vs. ML-GARCH models		
			CNS	VAR	VOL	LNV	VAR	VOL	LNV	VAR	VOL	LNV
TOPIX	Constant variance	412.4**										
	SARV variance	10.3	8.28**	—	−1.06	0.26	7.24**	1.00	1.84	1.22	0.46	1.52
	SARV volatility	5.4	8.36**	1.06	—	0.86	7.38**	1.48	2.19*	1.52	0.91	1.96
	SARV log variance	12.8	7.97**	−0.26	−0.86	—	6.85**	1.42	1.35	1.33	0.36	1.11

We estimate an overidentified version of the quantile regressions in Table 8 where the instrument vector contains ten competing volatility forecasts: those from the constant variance model (CNS) and those from the SARV, LS-GARCH, and ML-GARCH models in variance (VAR), volatility (VOL), and log variance (LNV). The table reports the Kitamura and Stutzer (1997) κ -statistic for misspecification for the constant variance model and for the three SARV models, as well as the Christoffersen, Hahn, and Inoue (2001) distance measure (Δ -statistic) for pairwise comparisons of the performance of each SARV model to that of the nine alternative models. The κ -statistic is asymptotically distributed χ^2_8 under the null (χ^2_0 for the constant variance model) and the Δ -statistic is asymptotically standard normal. Positive Δ -statistics indicate that the given SARV model outperforms the alternative. * and ** denote rejections at the 5% and 1% levels.

Table 10 GARCH model value-at-risk orthogonality tests.

		Δ -statistics						
		vs.	vs. LS-GARCH models			vs. ML-GARCH models		
Asset/model	κ -stat.	CNS	VAR	VOL	LNV	VAR	VOL	LNV
Panel A: Currencies								
British pound								
LS-GARCH variance	18.5*	5.09**	—	0.01	0.47	−1.39	−0.48	0.18
LS-GARCH volatility	18.7*	4.30**	−0.01	—	0.68	−0.51	−0.26	0.24
LS-GARCH log variance	24.5**	4.14**	−0.47	−0.68	—	−0.86	−0.78	−1.29
ML-GARCH variance	13.7	5.13**	1.39	0.51	0.86	—	0.57	0.58
ML-GARCH volatility	16.4*	5.11**	0.48	0.26	0.78	−0.57	—	0.43
ML-GARCH log variance	20.8**	4.36**	−0.18	−0.24	1.29	−0.58	−0.43	—
Canadian dollar								
LS-GARCH variance	10.4	5.20**	—	0.78	2.52*	0.87	0.76	2.56*
LS-GARCH volatility	16.1*	5.04**	−0.78	—	3.17**	0.21	−0.06	3.29**
LS-GARCH log variance	43.7**	3.90**	−2.52*	−3.17**	—	−2.33*	−3.24**	−0.06
ML-GARCH variance	17.7*	5.48**	−0.87	−0.21	2.33*	—	−0.25	2.25*
ML-GARCH volatility	15.9*	5.08**	−0.76	0.06	3.24**	0.25	—	3.28**
ML-GARCH log variance	43.6**	3.91**	−2.56*	−3.29**	0.06	−2.25*	−3.28**	—
Deutsche mark								
LS-GARCH variance	44.4**	4.55**	—	−2.73**	−2.24*	−2.35*	−2.68**	−2.05*
LS-GARCH volatility	19.8*	5.23**	2.73**	—	−0.84	1.51	−0.55	−0.88
LS-GARCH log variance	12.9	5.20**	2.24*	0.84	—	1.40	0.72	−0.47
ML-GARCH variance	28.6**	4.89**	2.35*	−1.51	−1.40	—	−1.61	−1.36
ML-GARCH volatility	18.7*	5.28**	2.68**	0.55	−0.72	1.61	—	−0.78
ML-GARCH log variance	10.1	4.99**	2.05*	0.88	0.47	1.36	0.78	—
Japanese yen								
LS-GARCH variance	33.0**	5.48**	—	−2.00*	0.75	−2.61**	−0.83	1.05
LS-GARCH volatility	14.2	5.84**	2.00*	—	2.45*	−1.01	2.18*	2.77**
LS-GARCH log variance	45.8**	5.31**	−0.75	−2.45*	—	−2.77**	−1.54	1.11
ML-GARCH variance	6.2	6.21**	2.61**	1.01	2.77**	—	1.91	3.02**
ML-GARCH volatility	25.7**	5.50**	0.83	−2.18*	1.54	−1.91	—	1.96
ML-GARCH log variance	50.7**	5.18**	−1.05	−2.77**	−1.11	−3.02**	−1.96	—
Swiss franc								
LS-GARCH variance	59.1**	3.80**	—	−3.87**	−1.85	−3.32**	−3.73**	−1.84
LS-GARCH volatility	12.6	5.08**	3.87**	—	1.42	0.46	−0.74	1.41
LS-GARCH log variance	26.6**	4.67**	1.85	−1.42	—	−0.83	−1.72	0.09
ML-GARCH variance	15.9*	4.83**	3.32**	−0.46	0.83	—	−1.01	0.83
ML-GARCH volatility	9.7	5.05**	3.73**	0.74	1.72	1.01	—	1.72
ML-GARCH log variance	26.6**	4.66**	1.84	−1.41	−0.09	−0.83	−1.72	—
Panel B: Equities								
NASDAQ								
LS-GARCH variance	27.6**	7.42**	—	−0.11	1.82	−1.23	−0.47	1.92
LS-GARCH volatility	26.6**	7.67**	0.11	—	2.43*	−1.79	−0.72	2.87**

continued

Table 10 (continued) GARCH model value-at-risk orthogonality tests.

Asset/model	κ -stat.	Δ -statistics						
		CNS	vs. LS-GARCH models			vs. ML-GARCH models		
			VAR	VOL	LNV	VAR	VOL	LNV
LS-GARCH log variance	58.1**	7.03**	-1.82	-2.43*	—	-3.02**	-2.97**	-0.13
ML-GARCH variance	12.9	7.69**	1.23	1.79	3.02**	—	1.22	3.34**
ML-GARCH volatility	21.7**	7.39**	0.47	0.72	2.97**	-1.22	—	3.32**
ML-GARCH log variance	57.3**	7.02**	-1.92	-2.87**	0.13	-3.34**	-3.32**	—
NYSE								
LS-GARCH variance	42.4**	4.74**	—	-2.63**	-0.96	-3.09**	-2.65**	-1.07
LS-GARCH volatility	15.5	5.20**	2.63**	—	0.95	-1.62	-1.25	0.88
LS-GARCH log variance	27.1**	4.96**	0.96	-0.95	—	-2.07*	-2.36*	-0.34
ML-GARCH variance	3.7	5.64**	3.09**	1.62	2.07*	—	0.66	1.95
ML-GARCH volatility	7.0	5.58**	2.65**	1.25	2.36*	-0.66	—	2.30*
ML-GARCH log variance	25.3**	4.75**	1.07	-0.88	0.34	-1.95	-2.30*	—
S&P 500								
LS-GARCH variance	56.9**	4.57**	—	-3.44**	-1.50	-4.07**	-3.56**	-1.52
LS-GARCH volatility	19.1*	5.67**	3.44**	—	1.34	-1.96*	-2.18*	1.54
LS-GARCH log variance	32.6**	5.15**	1.50	-1.34	—	-2.41*	-2.79**	0.32
ML-GARCH variance	5.7	6.03**	4.07**	1.96*	2.41*	—	0.16	2.56*
ML-GARCH volatility	6.6	5.82**	3.56**	2.18*	2.79**	-0.16	—	2.96**
ML-GARCH log variance	33.3**	5.19**	1.52	-1.54	-0.32	-2.56*	-2.96**	—
FTSE								
LS-GARCH variance	71.0**	6.66**	—	-3.59**	-0.72	-4.26**	-3.40**	-0.91
LS-GARCH volatility	21.5**	7.36**	3.59**	—	2.11*	-1.97*	-0.94	2.36*
LS-GARCH log variance	56.3**	7.18**	0.72	-2.11*	—	-2.96**	-2.60**	-0.28
ML-GARCH variance	8.6	8.09**	4.26**	1.97*	2.96**	—	1.45	3.12**
ML-GARCH volatility	16.8*	7.59**	3.40**	0.94	2.60**	-1.45	—	2.75**
ML-GARCH log variance	54.2**	7.05**	0.91	-2.36*	0.28	-3.12**	-2.75**	—
TOPIX								
LS-GARCH variance	242.0**	4.69**	—	-6.64**	-6.68**	-6.27**	-6.76**	-6.75**
LS-GARCH volatility	21.4**	7.87**	6.64**	—	0.84	0.45	-1.10	0.56
LS-GARCH log variance	34.9**	8.10**	6.68**	-0.84	—	-0.51	-1.32	-0.91
ML-GARCH variance	25.4**	7.63**	6.27**	-0.45	0.51	—	-1.57	0.24
ML-GARCH volatility	15.2	8.00**	6.76**	1.10	1.32	1.57	—	1.04
ML-GARCH log variance	29.6**	8.11**	6.75**	-0.56	0.91	-0.24	-1.04	—

We estimate an overidentified version of the quantile regressions in Table 8 where the instrument vector contains ten competing volatility forecasts: those from the constant variance model (CNS) and those from the SARV, LS-GARCH, and ML-GARCH models in variance (VAR), volatility (VOL), and log variance (LNV). The table reports the Kitamura and Stutzer (1997) κ -statistic for misspecification for the six GARCH models and the Christoffersen, Hahn, and Inoue (2001) distance measure (Δ -statistic) for pairwise comparisons of the performance of each GARCH model to that of the constant variance model and the other GARCH models. The κ -statistic is asymptotically distributed χ^2_8 under the null and the Δ -statistic is asymptotically standard normal. Positive Δ -statistics indicate that the given GARCH model outperforms the alternative. * and ** denote rejections at the 5% and 1% levels.

but there are many instances in which the SARV models do better than one or more of the GARCH models. For example, all of the SARV models outperform the LS-GARCH variance model for four of the indexes, and both the LS- and ML-GARCH log variance models for two of the indexes. At the same time, however, we find that both the LS- and ML-GARCH volatility models perform as well as any of the SARV models, with no significant differences reported. This finding is consistent with the relatively good performance of the GARCH volatility models in the simulations.

Table 10 examines the performance of the GARCH VaR estimates in greater detail. Like the SARV models, the GARCH models show clear evidence of misspecification for the currencies: every model is rejected in at least three of the specification tests. Although there does not appear to be much difference in the performance of the least-squares and maximum-likelihood models in this regard, the pairwise model comparisons prove more informative. Once again, we find that the least-squares volatility model delivers good performance, with no rejections in favor of any of the other models in panel A. The least-squares variance and log variance models, on the other hand, underperform by a significant margin in a number of cases. Among the maximum-likelihood models, the variance specification performs best, with no rejections in favor of the other models, and the volatility specification also performs well, with only one rejection in favor of the least-squares volatility model (for JY).

The results for the equity indexes (panel B) are very similar. On the basis of the specification test statistics, the maximum-likelihood variance and volatility models perform best, with only one and two rejections, respectively. These models also perform the best in the pairwise comparisons, with no rejections in favor of the other models. However, the least-squares volatility model also performs well, with only two rejections (S&P 500 and FTSE) in favor of the maximum-likelihood variance and volatility models. The least-squares variance and log variance models and the maximum-likelihood log variance model, on the other hand, are often outperformed in these tests. One significant difference from the currency results is that the least-squares variance model performs particularly poorly for some of the indexes. For example, in the case of the TOPIX, its Δ -statistics are in the -6 to -7 range for all pairwise comparisons with the other models. This suggests that, in the absence of other information, the volatility model is probably the most robust choice among the LS-GARCH models for conditional VaR estimation.

5. CONCLUSION

Our study of the performance of GARCH filters in a discrete-time SARV setting helps bridge the gap between the two main classes of volatility models. Although theory favors SARV models, we show that the MMS linear filters for three common SARV models imply the same forecasting rules as standard GARCH models. This relation suggests that GARCH models might perform well in forecasting stochastic volatility. To investigate this issue we fit both the SARV and GARCH models to artificial data to evaluate how they perform under controlled

conditions, and then we fit the models to daily currency and equity index returns to evaluate how they perform in a risk management application.

The simulation results indicate that the GARCH forecasts are less precise than the SARV forecasts, but it is not clear that the performance differences are large enough to be economically meaningful. Similarly there is no clear advantage to using either class of models to construct conditional VaR estimates for the currencies or the equity indexes. We do find, however, that among the GARCH models, the least-squares volatility specification performs well for both forecasting and conditional VaR estimation under a wide range of circumstances. Of course, we should note that these results are only preliminary in the sense that they are based on a limited set of SARV models that incorporate specific distributional assumptions. Nonetheless, the results should help narrow the scope of the ongoing debate over the relative merits of GARCH and SARV specifications.

In terms of methodology, our approach for constructing MMS linear filters provides a simple and surprisingly effective alternative to the computationally intensive techniques normally required for the econometric analysis of SARV models. Using the steady-state version of these filters, we can easily formulate a consistent least-squares estimator of the GARCH parameters associated with a given SARV process. Moreover, if we are primarily interested in estimating the underlying SARV parameters, we can simply fit the exact version of the filter using the standard recursion for the filter gain. This entails some loss of efficiency relative to MCMC estimation, but our simulation results suggest that the efficiency loss is not a major concern for samples of the size typically used in practice. Thus, by reducing the computational burden of estimating SARV models, our methods should facilitate wider acceptance of these models in applied work.

Subsequent research could extend the analysis of this article in several interesting directions. One possibility would be to examine multifactor SARV models and their associated GARCH filters. Another would be to look at multivariate specifications. Although SARV models do not appear to enjoy a systematic performance advantage in our VaR tests, this might change in a setting that requires the models to describe not only the dynamics of the individual volatilities, but also their comovements. In addition, the availability of simple econometric methods would be even more valuable in a multivariate SARV framework because of the added complexity of the models.

APPENDIX A: OPTIMAL FILTERING AND FORECASTING FOR SARV MODELS

This appendix shows how to (i) implement the optimal nonlinear filter for the SARV models and (ii) evaluate the likelihood function for the SARV models. Our approach is built around recently developed Monte Carlo procedures for performing the update and prediction steps implied by Bayes' theorem. Once we have the

output from the filter, it is straightforward to estimate $s_{t+k|t}$ using the approach discussed in Section 1.5.¹⁸

A.1 Optimal Nonlinear Filtering for SARV Models

Recall that $\mathcal{F}_t = (r_0, r_1, \dots, r_t)$ and let $\hat{\theta}$ denote an estimate of θ obtained using the methods discussed in Section 3.3 and Appendix B. We set $\theta = \hat{\theta}$ and use the following Monte Carlo procedure to perform the update and prediction steps of the optimal nonlinear filter. According to Bayes' theorem,

$$p(s_t|\mathcal{F}_t, \theta) \propto p(r_t|s_t, r_{t-1}, \theta)p(s_t|\mathcal{F}_{t-1}, \theta), \quad (52)$$

where

$$p(s_t|\mathcal{F}_{t-1}, \theta) = \int p(s_t|s_{t-1}, \theta)p(s_{t-1}|\mathcal{F}_{t-1}, \theta) ds_{t-1}. \quad (53)$$

Equations (2)–(4) imply that $p(s_t|s_{t-1}, \theta)$ is known once we specify a distribution for u_t . Thus, given $s_{t-1}^{(1)}, s_{t-1}^{(2)}, \dots, s_{t-1}^{(m)}$, a sample of m observations from $p(s_{t-1}|\mathcal{F}_{t-1}, \theta)$, we can approximate the distribution in Equation (53) as

$$p(s_t|\mathcal{F}_{t-1}, \theta) \simeq \frac{1}{m} \sum_{j=1}^m p(s_t|s_{t-1}^{(j)}, \theta). \quad (54)$$

Equation (52) then becomes

$$p(s_t|\mathcal{F}_t, \theta) \propto p(r_t|s_t, r_{t-1}, \theta) \frac{1}{m} \sum_{j=1}^m p(s_t|s_{t-1}^{(j)}, \theta). \quad (55)$$

The question is how to sample from $p(s_{t-1}|\mathcal{F}_{t-1}, \theta)$. We use a Monte Carlo approach known as an auxiliary particle filter. Our implementation of it closely follows Pitt and Shephard (1999) and Chib, Nardari, and Shephard (2002).

The filter has two stages. First we identify regions of high probability mass by creating a set of proposed values $s_t^{*(1)}, s_t^{*(2)}, \dots, s_t^{*(n)}$, where n is typically 5–10 times larger than m . Then we simulate from the density in Equation (54) with an emphasis on the high probability regions and resample the resulting values to produce a sample of size m from the distribution in Equation (55). The details are as follows.

Auxiliary particle filter

1. Given $s_{t-1}^{(1)}, s_{t-1}^{(2)}, \dots, s_{t-1}^{(m)}$, calculate

$$\tilde{s}_t^{*(j)} = \kappa + \phi s_{t-1}^{(j)} \quad (56)$$

$$\tilde{\pi}^{*(j)} = p(r_t|\tilde{s}_t^{*(j)}, r_{t-1}, \theta), \quad (57)$$

for $j = 1, 2, \dots, m$.

¹⁸ For discussion purposes we treat forecasting, optimal filtering, and likelihood evaluation as separate procedures. But they can easily be implemented simultaneously if desired.

2. Sample the integers $1, 2, \dots, m$ a total of n times with replacement using probabilities proportional to $\tilde{\pi}^{*(j)}$ for $j = 1, 2, \dots, m$. Let the resulting set of indexes be denoted by k_1, k_2, \dots, k_n .
3. For each k_j from step 2, simulate

$$s_t^{*(j)} = \kappa + \phi s_{t-1}^{(k_j)} + \gamma s_{t-1}^{(k_j)\rho} u_{t-1}^{(j)}, \quad (58)$$

where $\rho = 1/2$ for Model 1, $\rho = 0$ for Models 2 and 3, and $u_{t-1}^{(j)}$ has the appropriate distribution ($N(0,1)$ in our specification of the models).

4. Resample the values $s_t^{*(1)}, s_t^{*(2)}, \dots, s_t^{*(n)}$ a total of m times with replacement using probabilities proportional to

$$\pi^{*(j)} = \frac{p(r_t | s_t^{*(j)}, r_{t-1}, \boldsymbol{\theta})}{p(r_t | \tilde{s}_t^{*(k_j)}, r_{t-1}, \boldsymbol{\theta})}, \quad (59)$$

for $j = 1, 2, \dots, n$ to obtain $s_t^{(1)}, s_t^{(2)}, \dots, s_t^{(m)}$, a sample of m observations from $p(s_t | \mathcal{F}_t, \boldsymbol{\theta})$.

We initialize the algorithm by generating a sample $s_0^{(1)}, s_0^{(2)}, \dots, s_0^{(m)}$ from the marginal distribution of the s_t process.

From this algorithm it is apparent that the auxiliary particle filter is closely related to importance sampling in the context of Monte Carlo integration. As such, it is straightforward to implement. Pitt and Shephard (1999) discuss methods for computationally efficient sampling of indexes with replacement from populations with unequal probabilities. In the empirical analysis, we use $m = 1000$ and $n = 5000$.

A.2 The Likelihood Function

We approximate the likelihood of the model as $L(\hat{\boldsymbol{\theta}}) = \prod_{t=1}^T p(r_t | \mathcal{F}_{t-1}, \hat{\boldsymbol{\theta}})$. The value of $L(\hat{\boldsymbol{\theta}})$ is obtained using the following algorithm with $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$.

Likelihood evaluation

1. Set $t = 1$ and generate a sample of m observations $s_{t-1}^{(1)}, s_{t-1}^{(2)}, \dots, s_{t-1}^{(m)}$ from the marginal distribution of the s_t process.
2. Generate a sample of m observations $s_t^{(1)}, s_t^{(2)}, \dots, s_t^{(m)}$ from $p(s_t | \mathcal{F}_{t-1}, \boldsymbol{\theta})$. Specifically, for each value of $j \leq m$, simulate

$$s_t^{(j)} = \kappa + \phi s_{t-1}^{(j)} + \gamma s_{t-1}^{(j)\rho} u_{t-1}^{(j)}, \quad (60)$$

where $\rho = 1/2$ for Model 1, $\rho = 0$ for Models 2 and 3, and $u_{t-1}^{(j)}$ has the appropriate distribution ($N(0,1)$ in our specification of the models).

3. Estimate the one-step-ahead density of r_t as

$$p(r_t | \mathcal{F}_{t-1}, \boldsymbol{\theta}) \simeq \frac{1}{m} \sum_{j=1}^m p(r_t | s_t^{(j)}, r_{t-1}, \boldsymbol{\theta}). \quad (61)$$

4. Use the particle filter to generate a sample of m observations $s_t^{(1)}, s_t^{(2)}, \dots, s_t^{(m)}$ from $p(s_t | \mathcal{F}_t, \theta)$.
5. If $t < T$, then add one to the current value of t and go to step 2.
6. Compute $L(\theta) = \prod_{t=1}^T p(r_t | \mathcal{F}_{t-1}, \theta)$.

APPENDIX B: MCMC SAMPLING SCHEME

This appendix describes our sampling scheme. We assume some familiarity with MCMC methods. Good introductions to these methods are given by Gelfand and Smith (1990), Casella and George (1992), and Chib and Greenberg (1995).

B.1 Sampling Scheme for the Variances, Volatilities, and Log Variances

Suppose the data consist of the return sequence $\{r_t\}_{t=0}^T$. Let $\mathbf{r} = (r_1, r_2, \dots, r_T)'$ and $\mathbf{s} = (s_1, s_2, \dots, s_T)'$. We want to generate from the full conditional distribution of \mathbf{s} . This distribution takes the form

$$p(\mathbf{s} | r_0, s_0, \mathbf{r}, \theta) \propto p(\mathbf{r} | r_0, \mathbf{s}, \theta) p(\mathbf{s} | s_0, \theta), \quad (62)$$

where $s_0 = \kappa / (1 - \phi)$. A sampling scheme typically mixes better if we generate in blocks using a randomly chosen block size for each sampling operation. The method for randomizing the block size does not appear to be critical. After some experimentation, we settled on an approach in which the block size is set equal to one plus a random draw from a Poisson distribution with parameter three. This corresponds to generating an average of four elements of \mathbf{s} at a time.

To illustrate, let $\mathbf{s}_{(t,\tau)} = (s_t, \dots, s_\tau)$ denote a block that starts on date t and ends on date τ . Our target density is

$$\begin{aligned} f(\mathbf{s}_{(t,\tau)}) &= p(\mathbf{s}_{(t,\tau)} | r_0, s_0, \mathbf{r}, \mathbf{s}_{-(t,\tau)}, \theta) \\ &\propto p(\mathbf{r}, \mathbf{s}, \theta | r_0, s_0) \\ &\propto \prod_{j=t}^{\tau+1} p(r_j | r_{j-1}, s_j, \theta) p(s_j | s_{j-1}, \theta) I(\sigma_j \in \mathbb{R}^+) \\ &\propto \prod_{j=t}^{\tau+1} \sigma_j^{-1} s_{j-1}^{-\rho} \exp\left(-\frac{1}{2} \frac{e_j^2}{\sigma_j^2}\right) \exp\left(-\frac{1}{2} \frac{v_{j-1}^2}{\gamma^2 s_{j-1}^{2\rho}}\right) I(\sigma_j \in \mathbb{R}^+), \end{aligned} \quad (63)$$

where $\rho = 1/2$ for Model 1, $\rho = 0$ for Models 2 and 3, $I(\cdot)$ is the indicator function, and $\mathbf{s}_{-(t,\tau)}$ denotes all the elements of \mathbf{s} except those in the block. Note that the presence of the indicator function reflects the nonnegativity constraint on the volatilities.

Since it is not practical to sample directly from $f(\mathbf{s}_{(t,\tau)})$, we generate the block from a proposal density $g(\mathbf{s}_{(t,\tau)})$ and use the Metropolis-Hastings algorithm to perform the required correction. Suppose, for example, that $\mathbf{s}_{(t,\tau)}^{(i)}$ denotes the

value of the block after the i th sampling operation. We obtain $\mathbf{s}_{(t,\tau)}^{(i+1)}$ by generating a candidate block $\mathbf{s}_{(t,\tau)}^{(c)}$ from $g(\mathbf{s}_{(t,\tau)})$ and accepting it with probability a_c where

$$a_c = \min \left\{ 1, \frac{f(\mathbf{s}_{(t,\tau)}^{(c)})g(\mathbf{s}_{(t,\tau)}^{(i)})}{f(\mathbf{s}_{(t,\tau)}^{(i)})g(\mathbf{s}_{(t,\tau)}^{(c)})} \right\}. \quad (64)$$

If the candidate is accepted, then $\mathbf{s}_{(t,\tau)}^{(i+1)} = \mathbf{s}_{(t,\tau)}^{(c)}$; otherwise, $\mathbf{s}_{(t,\tau)}^{(i+1)} = \mathbf{s}_{(t,\tau)}^{(i)}$.

In practice, it is important to make sure that the sampling scheme converges quickly and produces iterates that display good movement around the support of the target density. This can usually be accomplished by selecting a proposal density that is a good approximation to $f(\cdot)$. A strategy that often works well is to use a multivariate normal proposal obtained by performing a second-order Taylor series expansion of $\log f(\cdot)$ about the mode of the distribution. But this is too computationally intensive for our purposes. Instead, we use a simpler approach that still produces good movement and reasonable acceptance rates.¹⁹ The idea, which exploits the fact that volatility is highly persistent, is to obtain a multivariate normal approximation to $p(\mathbf{s}_{(t,\tau)} | \mathbf{s}_{-(t,\tau)}, \boldsymbol{\theta})$ by approximating the conditional mean and variance of $s_{j+1} - s_j$ as locally constant on the interval $t-1 \leq j \leq \tau$.

Let $\mathbf{1}$ and \mathbf{L} denote a $k \times 1$ vector of ones and a lower-triangular $k \times k$ matrix of ones, respectively, where $k = \tau - t + 1$ is the number of elements in the block. To implement our approach we approximate the evolution of the s_t process as

$$s_{j+1} = s_j + \kappa + (\phi - 1)\bar{s} + \gamma\bar{s}^\rho u_j, \quad t-1 \leq j \leq \tau, \quad (65)$$

where $\bar{s} = (s_{t-1} + s_{\tau+1})/2$. Under this approximation,

$$\mathbf{s}_{(t,\tau)} | \mathbf{s}_{-(t,\tau)}, \boldsymbol{\theta} \sim N(\mathbf{m}, \mathbf{B}), \quad (66)$$

where

$$\mathbf{m} = \frac{s_{t-1}(k+1)\mathbf{1} + (s_{\tau+1} - s_{t-1})\mathbf{L}\mathbf{1}}{k+1} \quad (67)$$

$$\mathbf{B} = \left(\frac{(k+1)\mathbf{L}\mathbf{L}' - \mathbf{L}\mathbf{1}\mathbf{1}'\mathbf{L}'}{k+1} \right) (\gamma\bar{s}^\rho)^2. \quad (68)$$

This is our proposal density. A significant advantage of this proposal is that \mathbf{B}^{-1} has a very simple form. It is given by $\mathbf{D}(\gamma\bar{s}^\rho)^{-2}$, where \mathbf{D} is a tridiagonal matrix with $D_{ii} = 2$ and $D_{ij} = -1$ for $1 \leq i \leq k$ and $j = i \pm 1$. As a result, we can generate the variances and compute the acceptance probability without performing any numerical matrix inversions. This makes for a fast sampling scheme.

B.2 Sampling Scheme for the Parameters

Note that μ and δ are set equal to zero for the artificial data and estimated in a first-stage regression for the real data. Therefore we need only consider κ , ϕ , and γ . For

¹⁹ The acceptance rate varies somewhat depending on the model, but it is generally in the neighborhood of 90% under our strategy of generating an average of four elements of \mathbf{s} at a time.

all three models we require $-1 \leq \phi \leq 1$ and $\gamma \geq 0$. In addition, we impose the constraint $\kappa \geq 0$ for Models 1 and 2. We incorporate these constraints by specifying a joint prior of the form

$$p(\boldsymbol{\theta}) \propto (1/\gamma^2)I(\boldsymbol{\theta} \in C), \quad (69)$$

where C is the set of parameter values for which the constraints are satisfied. This corresponds to a jointly uniform improper prior on κ and ϕ combined with a standard noninformative prior on γ^2 in the region of the parameter space where the inequalities hold.

B.2.1 Generating κ and ϕ We generate κ and ϕ together as a block. Let $\mathbf{b}' = (\kappa, \phi)$. The full conditional distribution of \mathbf{b} is

$$\begin{aligned} f(\mathbf{b}) &= p(\mathbf{b}|r_0, s_0, \mathbf{r}, \mathbf{s}, \boldsymbol{\theta}_{-b}) \\ &\propto p(\mathbf{r}, \mathbf{s}, \boldsymbol{\theta}|r_0, s_0) \\ &\propto \prod_{j=1}^T \exp\left(-\frac{1}{2} \frac{v_{j-1}^2}{\gamma^2 s_{j-1}^{2\rho}}\right) I(\boldsymbol{\theta} \in C). \end{aligned} \quad (70)$$

The right side of Equation (70) is proportional to a truncated bivariate normal distribution. To see this, let $\mathbf{x}' = (x'_1, \dots, x'_T)$ and $\mathbf{y}' = (y_1, \dots, y_T)$, where

$$x_j = \left(\frac{1}{\gamma s_{j-1}^\rho}, \frac{s_{j-1}}{\gamma s_{j-1}^\rho} \right), \quad (71)$$

$$y_j = \frac{s_j}{\gamma s_{j-1}^\rho}. \quad (72)$$

Using this notation, we can express our target density as

$$\begin{aligned} f(\mathbf{b}) &\propto \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{x}\mathbf{b})'(\mathbf{y} - \mathbf{x}\mathbf{b})\right) I(\boldsymbol{\theta} \in C) \\ &\propto \exp\left(-\frac{1}{2}(\mathbf{b} - \bar{\mathbf{b}})'(\mathbf{x}'\mathbf{x})(\mathbf{b} - \bar{\mathbf{b}})\right) I(\boldsymbol{\theta} \in C), \end{aligned} \quad (73)$$

where $\bar{\mathbf{b}} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y}$. Equation (73) implies that \mathbf{b} is bivariate normal with mean vector $\bar{\mathbf{b}}$ and covariance matrix $(\mathbf{x}'\mathbf{x})^{-1}$ in the region of the parameter space where the inequality constraints hold.

B.2.2 Generating γ Let $\boldsymbol{\epsilon}' = (\epsilon_1, \dots, \epsilon_T)$, where

$$\epsilon_j = (s_j - \kappa - \phi s_{j-1})/s_{j-1}^\rho. \quad (74)$$

The full conditional distribution of γ^2 is

$$\begin{aligned} f(\gamma^2) &= p(\gamma^2|r_0, s_0, \mathbf{r}, \mathbf{s}, \boldsymbol{\theta}_{-\gamma}) \\ &\propto p(\mathbf{r}, \mathbf{s}, \boldsymbol{\theta}|r_0, s_0) \\ &\propto \frac{1}{\gamma^{2(T/2+1)}} \exp\left(-\frac{1}{2} \frac{\boldsymbol{\epsilon}'\boldsymbol{\epsilon}}{\gamma^2}\right) I(\boldsymbol{\theta} \in C). \end{aligned} \quad (75)$$

The right side of Equation (75) is proportional to an inverse gamma distribution with parameters $T/2$ and $\epsilon'\epsilon/2$ in the region of the parameter space where the inequality constraints hold.

Received June 7, 2002; revised July 30, 2003; accepted September 3, 2003

REFERENCES

- Abramowitz, Milton, and Irene A. Stegun. (1970). *Handbook of Mathematical Functions*. New York: Dover Publications.
- Andersen, Torben G. (1994). "Stochastic Autoregressive Volatility: A Framework for Volatility Modeling." *Mathematical Finance* 4, 75–102.
- Andersen, Torben G. (1996). "Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility." *Journal of Finance* 51, 169–204.
- Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys. (2003). "Modeling and Forecasting Realized Volatility." *Econometrica* 71, 579–625.
- Barndorff-Nielsen, Ole E., and Neil Shephard. (2001). "Non-Gaussian Ornstein-Uhlenbeck-Based Models and Some of Their Uses in Financial Economics." *Journal of the Royal Statistical Society, Series B* 63, 167–207.
- Barndorff-Nielsen, Ole E., and Neil Shephard. (2002). "Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models." *Journal of the Royal Statistical Society, Series B* 64, 253–280.
- Bollerslev, Tim. (1986). "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31, 307–327.
- Bollerslev, Tim, and Peter E. Rossi. (1995). "Dan Nelson Remembered." *Journal of Business and Economic Statistics* 13, 361–364.
- Casella, George, and Edward I. George. (1992). "Explaining the Gibbs Sampler." *American Statistician* 46, 167–174.
- Chib, Siddhartha, and Edward Greenberg. (1995). "Understanding the Metropolis-Hastings Algorithm." *American Statistician* 49, 327–335.
- Chib, Siddhartha, Federico Nardari, and Neil Shephard. (2002). "Markov Chain Monte Carlo Methods for Stochastic Volatility Models." *Journal of Econometrics* 108, 281–316.
- Christoffersen, Peter F. (1998). "Evaluating Interval Forecasts." *International Economic Review* 39, 841–862.
- Christoffersen, Peter, Jinyong Hahn, and Atsushi Inoue. (2001). "Testing and Comparing Value-at-Risk Measures." *Journal of Empirical Finance* 8, 325–342.
- Daniélsson, Jon. (1998). "Multivariate Stochastic Volatility Models: Estimation and a Comparison with VGARCH Models." *Journal of Empirical Finance* 5, 155–173.
- Davidian, M., and R. J. Carroll. (1987). "Variance Function Estimation." *Journal of the American Statistical Association* 82, 1079–1091.
- Fleming, Jeff, Chris Kirby, and Barbara Ostdiek. (1998). "Information and Volatility Linkages in the Stock, Bond, and Money Markets." *Journal of Financial Economics* 49, 111–137.
- Gelfand, Alan E., and Adrian F. M. Smith. (1990). "Sampling-based Approaches to Calculating Marginal Densities." *Journal of the American Statistical Association* 85, 398–409.
- Geweke, John. (1986). "Modeling the Persistence of Conditional Variances: A Comment." *Econometric Reviews* 5, 57–61.

- Gordon, N. J., D. J. Salmond, and A. F. M. Smith. (1993). "Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation." *IEEE Proceedings - Part F, Radar and Signal Processing* 140, 107-113.
- Hamilton, James D. (1994). *Time Series Analysis*. Princeton, NJ: Princeton University Press.
- Harvey, Andrew C., Esther Ruiz, and Neil Shephard. (1994). "Multivariate Stochastic Variance Models." *Review of Economic Studies* 61, 247-264.
- Harvey, Andrew C., and Neil Shephard. (1996). "Estimation of an Asymmetric Model of Asset Prices." *Journal of Business and Economic Statistics* 14, 429-434.
- Heston, Steven L. (1993). "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options." *Review of Financial Studies* 6, 327-343.
- Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi. (1994). "Bayesian Analysis of Stochastic Volatility Models." *Journal of Business and Economic Statistics* 12, 371-389.
- Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi. (2001). "Bayesian Analysis of Stochastic Volatility Models with Fat-Tails and Correlated Errors." Working Paper, University of Chicago; forthcoming in *Journal of Econometrics*.
- Kim, Sangjoon, Neil Shephard, and Siddhartha Chib. (1998). "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models." *Review of Economic Studies* 65, 361-394.
- Kitagawa, Genshiro. (1996). "Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models." *Journal of Computational and Graphical Statistics* 5, 1-25.
- Kitamura, Yuichi, and Michael Stutzer. (1997). "An Information-Theoretic Alternative to Generalized Method of Moments Estimation." *Econometrica* 65, 861-874.
- Meddahi, Nour, and Éric Renault. (2002). "Temporal Aggregation of Volatility Models." Working Paper 2000s-22, CIRANO; forthcoming in *Journal of Econometrics*.
- Melino, Angelo, and Stuart M. Turnbull. (1990). "Pricing Foreign Currency Options with Stochastic Volatility." *Journal of Econometrics* 45, 239-265.
- Nelson, Daniel B. (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica* 59, 347-370.
- Nelson, Daniel B., and Dean P. Foster. (1994). "Asymptotic Filtering Theory for Univariate ARCH models." *Econometrica* 62, 1-41.
- Newey, Whitney K., and Kenneth D. West. (1987). "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55, 703-708.
- Pitt, Michael K., and Neil Shephard. (1999). "Filtering via Simulation: Auxiliary Particle Filters." *Journal of the American Statistical Association* 94, 590-599.
- Rich, Robert W., Jennie Raymond, and J. S. Butler. (1991). "Generalized Instrumental Variables Estimation of Autoregressive Conditional Heteroskedastic Models." *Economics Letters* 35, 179-185.
- Schwarz, Gideon. (1978). "Estimating the Dimension of a Model." *Annals of Statistics* 6, 461-464.
- Schwert, G. William. (1990). "Stock Volatility and the Crash of '87." *Review of Financial Studies* 3, 77-102.
- Stein, Elias M., and Jeremy C. Stein. (1991). "Stock Price Distributions with Stochastic Volatility: An Analytic Approach." *Review of Financial Studies* 4, 727-752.
- Tauchen, George E., and Mark Pitts. (1983). "The Price Variability-Volume Relationship on Speculative Markets." *Econometrica* 51, 485-505.
- Taylor, Stephen J. (1986). *Modelling Financial Time Series*. Chichester: John Wiley & Sons.