A Score Test for Discreteness in GARCH Models

Henrik Amilon*

Department of Economics

Lund University

Abstract

The standard continuous-state GARCH model is misspecified if applied to returns calculated from discrete price series. We propose a modification of the above model for handling such cases, by modeling the dependent variable as an unobservable stochastic variable with certain observed outcomes. We further construct a score test that can be used to check if the proposed model differ significantly from the one we would have if all variables were observed, i.e. an underlying latent GARCH model. Using price data from some Australian stocks with high tick size to price ratios, we find the important result that in no case does the proposed model differ significantly from an unobservable continuous-state GARCH model.

Key words: GARCH, latent variables, generalized residuals, score test.

JEL Codes: C22, C24, C51, C52

---

*Allhelgona Kyrkog. 6a, SE-223 62 Lund, Sweden. Phone: +46 46 137876. Fax: +46 46 222 4118. E-mail: henrik.amilon@nek.lu.se. I am most grateful to Tobias Rydén for indispensable help on many issues. Financial support is received from the Quantitative Finance Research Group at the School of Finance and Economics, University of Technology, Sydney.

1 Introduction

The standard models of financial asset returns assume the dependent variable to be continuous. This is reasonable in modeling stock index returns, for example, but may be inappropriate in modeling returns from an individual asset, like a stock, since asset prices are traded in discrete units or ticks. For some assets, tick sizes might be small enough to justify a continuous-state model, but for other securities, such as low-priced stocks, the effects of discreteness are not necessarily negligible. It should also be clear that the problems with discreteness becomes more relevant for high frequency data and, ultimately, transaction data, since the price changes between two consecutive trading points then rarely are greater than a few ticks.

The breakdown of the continuous-state modeling in analyzing transaction data has lead some researchers to use statistical tools developed for cases where the dependent variable is discrete, as the price move in the ordered probit model of Hausman *et al.* (1992). Another line of work is found in Rydberg and Shephard (2003), who decompose the conditional distribution in three parts: activity, direction, and size of the price moves. Engle and Russel (1998) follow an alternative route, constructing a time series of a multinomial distribution. Hasbrouck (1999) constructs a microstructure model with deterministic and stochastic parts, and estimates it with Bayesian methods and Kalman filtering. The above references analyze *price changes*, while Amilon (2002) investigates *returns* from stocks with a high tick size to price ratio. The underlying latent GARCH process is estimated by replacing the unobservable variables in the log-likelihood function with their expectations, conditional on the observed information. This estimation procedure is motivated by the EM algorithm of Dempster *et al.* (1977), and has been used by Morgan and Trevor (1999) in the problem of censored GARCH returns, but it is only an approximation. This problem is further addressed in Wei (2002) who develops a censored GARCH model using Bayesin methodology and Gibbs sampling, unfortunately at the expense of a highly time-consuming estimation procedure. Fiorentini *et al.* (2002) perform more computational efficient exact likelihood inference of latent GARCH models, by the use of Markov chain Monte Carlo methods, see Chib (2001), and a Bayesian approach. They also apply their results through the estimation of a latent factor GARCH model to UK sectorial stock returns. Similar to Amilon (2002), the main differences compared to approximative models are in the estimates of the conditional variance parameters.

In this paper, we extend the continuous-state GARCH model of Bollerslev (1986) to handle cases where the return series are calculated from discrete prices, by modeling the dependent variable as an unobservable stochastic variable with certain observed outcomes. Inspired by the work of Hausman *et al.* (1992), the stochastic process is driven by the observable returns, and its log-likelihood function is easily derived. The proposed model have a rather loose connection to a latent equilibrium model which is hidden by discrete prices. We can, however, in the spirit of Gourieroux *et al.* (1985, 1987), construct generalized residuals and design a score test that can be used to test the proposed model from the one we would have if all variables were observed, i.e. the latent GARCH model. On the basis of the test we can then decide whether or not to move on to the more sophisticated, but also more computationally demanding, Bayesian estimation techniques. This is the very purpose of the paper.

The focus is on the AR-GARCH framework, as it is often used in modeling financial asset returns, see e.g. the survey in Bollerslev *et al.* (1994), although the same ideas most likely can be used for many other stochastic processes and latent GARCH models as well. The problem with discreteness is not restricted to financial series. A lot of econometric data, e.g. macroeconomic CPI series, are measured with low precision and could be risky to estimate with continuous-state statistical models.

In Section 2, we describe the model. Section 3 contain the results of the different estimations, while Section 4 presents the score test. A summary and concluding remarks are found in Section 5.

2 The Model

Let $\tilde{P}_t$ be the unobserved equilibrium price of a stock at time t, D_t its dividend, and $\tilde{r}_t$ the unobserved return of the stock, that is $\tilde{r}_t = \ln((\tilde{P}_t + D_t)/\tilde{P}_{t-1})$ with logarithmic returns, and $\tilde{r}_t = (\tilde{P}_t + D_t)/\tilde{P}_{t-1} - 1$ with percentage returns. The continuous-state AR(o)-GARCH(p, q) model, for notational reasons reduced to $o = 2$, and $p = q = 1$, takes the form

$$\begin{aligned} \tilde{r}_t &= \tilde{m}_t + \tilde{\varepsilon}_t, \qquad \tilde{\varepsilon}_t \mid I_{t-1} \backsim N(0, \sigma_t^2) \\ \sigma_t^2 &= \gamma_0 + \gamma_1 \tilde{\varepsilon}_{t-1}^2 + \gamma_2 \sigma_{t-1}^2, \end{aligned} \tag{1}$$

where the conditional mean process $\tilde{m}_t = \beta_0 + \beta_1 \tilde{r}_{t-1} + \beta_2 \tilde{r}_{t-2}$, and I_t is the unobserved information set at time t, that is $I_t = \left\{\tilde{P}_t, D_t, \tilde{P}_{t-1}, ...\right\}$. This model can, of course, not be estimated since $\tilde{P}_t$, and hence $\tilde{r}_t$, is not revealed to us. One simple approach would be to ignore that prices are discrete and the standard way estimate

$$\begin{aligned} r_t &= m_t + \varepsilon_t, \qquad \varepsilon_t \mid \Psi_{t-1} \backsim N(0, \sigma_t^2) \\ \sigma_t^2 &= \gamma_0 + \gamma_1 \varepsilon_{t-1}^2 + \gamma_2 \sigma_{t-1}^2, \end{aligned} \tag{2}$$

where r_t is computed from the observed prices, $\{P_s\}_1^T$, $m_t = \beta_0 + \beta_1 r_{t-1} + \beta_2 r_{t-2}$, and the observed information set $\Psi_t = \{P_t, D_t, P_{t-1}, ...\}$. This model is obviously misspecified, since ε_t cannot be a continuous random variable if r_t is discretely distributed around the conditional mean. A candidate for a model, not lacking from this misspecification, would be the following:

$$\begin{aligned} r_t^* &= m_t + \varepsilon_t^*, \qquad\qquad \varepsilon_t^* \mid \Psi_{t-1} \backsim N(0,\sigma_t^2) \\ \sigma_t^2 &= \gamma_0 + \gamma_1 \varepsilon_{t-1}^2 + \gamma_2 \sigma_{t-1}^2, \end{aligned} \tag{3}$$

where r_t^* is an unobservable, or latent, continuous random variable. Note that the expression for the conditional variance still is as in (2), that is $\varepsilon_t = r_t - m_t$. Hence, the conditional mean and variance processes are driven by the observable returns.

The basic idea, which is illustrated in Figure 1, is that we do not know r_t^*, but given the prices and the tick size, we can partition the state space of r_t^* in such a way that we know the boundaries, α_{1t} and α_{2t}, comprising r_t^*. Using a *rounding* procedure is one way of determining the boundaries. Say, for instance, that $P_{t-1} = 100, P_t = 101, D_t = 0$, and the tick size is $h = 1$. The observed percentage return is then 1%, but we can only conclude that r_t^* is between $\alpha_{1t} = (101 - 0.5 - 100)/100 = 0.5\%$, and $\alpha_{2t} = (101 + 0.5 - 100)/100 = 1.5\%$. More generally, modeling percentage returns, the boundaries are given by

$$\alpha_{1t} = \frac{P_t + D_t - h/2 - P_{t-1}}{P_{t-1}} \qquad \text{and} \qquad \alpha_{2t} = \frac{P_t + D_t + h/2 - P_{t-1}}{P_{t-1}}, \tag{4}$$

and for the more frequently used logarithmic returns by

$$\alpha_{1t} = \ln\left(\frac{P_t + D_t - h/2}{P_{t-1}}\right) \qquad \text{and} \qquad \alpha_{2t} = \ln\left(\frac{P_t + D_t + h/2}{P_{t-1}}\right). \tag{5}$$

The conditional distribution of r_t can now be expressed by assuming a suitable distribution of ε_t^*. With Gaussian error terms it is given by

$$\Pr(r_t \mid \Psi_{t-1}) = \Pr(\alpha_{1t} \le m_t + \varepsilon_t^* < \alpha_{2t} \mid \Psi_{t-1}) = \Phi\left(\frac{\alpha_{2t} - m_t}{\sigma_t}\right) - \Phi\left(\frac{\alpha_{1t} - m_t}{\sigma_t}\right), \tag{6}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. Once $\Pr(r_t \mid \Psi_{t-1})$ is specified, the log-likelihood function for a sample of T observations will be

$$\ln L(r_1, ..., r_T \mid \Psi_0) = \sum_{t=1}^{T} \ln l(r_t|\Psi_{t-1}) = \sum_{t=1}^{T} \ln \left(\Phi\left(\frac{\alpha_{2t} - m_t}{\sigma_t}\right) - \Phi\left(\frac{\alpha_{1t} - m_t}{\sigma_t}\right) \right). \qquad (7)$$

We can actually introduce an additional degree of freedom. What we seek is the mapping between the conditional distribution of r_t^* and that of r_t, where the aim is to partition the conditional state space of r_t^* in such a way that when the price process is between two boundaries, r_t is observed. In other words, the boundaries are not required to be symmetric as in the rounding procedure, but they must be non-overlapping. To be more specific, we can introduce yet another parameter, H, such that the expressions in (5) are replaced by

$$\alpha_{1t} = \ln\left(\frac{P_t + D_t - h/2 + H)}{P_{t-1}}\right) \qquad \text{and} \qquad \alpha_{2t} = \ln\left(\frac{P_t + D_t + h/2 + H}{P_{t-1}}\right), \qquad (8)$$

and similarly for the boundaries in (4). If $H = 0$, (5) and (8) are thus identical.

One may argue that if one believes that there exists a hidden underlying equilibrium price model, then the connection between (1) and (3) is rather loose, since the driving force in (3) is the observable data. We can, however, as shown in Section 4, test model (3) to model (1), with r_{t-j}^* and $\varepsilon_{t-k}^{*2}, j, k \geq 1$ included in the conditional mean and variance processes, respectively. If (3) is augmented with the same lagged endogenous variables, then the specification in (3) nests the one in (1). Furthermore, if the augmentations are deemed insignificant, the models are statistically equivalent.

3 The Estimates

In order to investigate how serious it is to neglect the effects of discreteness, we use daily prices of eight stocks traded at the Australian Stock Exchange (ASX) from January 31 2001 to January 31 2002. The market codes for the stocks are: DIO, ECP, IBA, MOS, MXO, SKR, SMX, and YTL. These low-priced stocks are chosen because they are highly liquid and have a high tick size to price ratio. On the ASX, stock prices lower than 10 cents have $h = 0.1$ cents, between 10 and 50 cents $h = 0.5$, while prices above 50 cents have a tick size of 1. The highest tick size to price ratios therefore appears when share prices are around and above 10 cents, and consequently are as large as 5%[1].

The series are fitted to the misspecified model (2) and to the modified model (3), henceforth model$_C$ and model$_D$. The choice of an AR(2) process in the conditional mean is purely ad hoc, but to some extent justified by the empirical evidence of low order autocorrelations in financial market returns. Besides, the purpose here is to examine differences in model estimates, and not to concern ourselves with model validation procedures. In the specification testing in Section 4, we further investigate the validity of this choice.

The parameter estimates and the asymptotic standard errors from the different time series and model specifications are presented in Table 1[2]. The estimates of β_1 and β_2, as well as their standard errors,

[1] When the tick size changes with the price level, the equations have to be slightly modified. Instead of a fixed h, it now differs with the price. For example, in the borderline case when $P_t = 10$, we have $h = 0.5$ in α_{2t} in (8), while $h = 0.1$ in α_{1t}.

[2] The optimization algorithm used is either the BHHH or the Newton-Raphson method in GAUSS. For many series, the Hessians are not positive definite. To facilitate the comparisons, all covariance matrices are computed as the inverse of the cross-product of the first derivatives. Unfortunately, the standard errors tend to be rather large with this method. These

are quite similar across the different models. One notices the quite large negative autocorrelations for MOS. The differences in the parameter estimates of the conditional variance process are more notable, although the results often are obscured by large standard errors. Only in some cases (typed in bold in Table 1) do the estimates from the proposed models differ, based on their asymptotic distributions, from what model$_C$ yields, at the 5 % significant level. For all series, $\hat{H}$ is negative and large compared to the tick levels (0.1, 0.5 and 1), but only significantly for IBA.

Comparing the estimates of the models to determine if they are statistically different is not the only subject of interest. One must remember that the ML estimates standard errors of the misspecified model$_C$ in principle are meaningless. Say, for example, that we want to test if the conditional variance for IBA is constant, that is if $\hat{\gamma}_1$ and $\hat{\gamma}_2$ equals zero. Statistical inference based on model$_C$ rather than the other model, would give rise to quite different and erroneous conclusions. In Table 1 we indicate (with †) those cases where model$_D$ gives opposite results than model$_C$, regarding the statistical significance of the parameter estimates (compared to 0), at the 5% level.

For one parameter, β_0, the estimates differ much more. This is not unexpected, bearing Figure 1 in mind. A downward shift of the boundaries (8), that is if $H < 0$, can to some extent be offset by a downward shift of the curve, that is a decrease in β_0. The effect is only partial since β_0 is common for all t, while the boundaries change with t, as the price evolves. A positive correlation between $\widehat{H}$ and $\hat{\beta}_0$ is confirmed in Table 1; if $\widehat{H} < 0$, $\hat{\beta}_0$ is lower compared to the other model.

It seems, from the data used here, as if the effects of discreteness in stock returns influence the parameter estimates and the standard errors. The findings may therefore have important real effects in

problems may be reduced by calculating the Hessians from first derivatives only. The estimates of the initial values in the conditional variance processes are of less importance and therefore not presented.

financial areas where the specification of the second moment is of significance, for example in volatility forecasting and portfolio optimization.

4 The Score Test

A most useful tool in regression diagnostics is the analysis of model residuals. Here, model$_C$ is misspecified and gives no true residuals, while the residuals from model (1) and model$_D$ are not computable, since we do not observe $\tilde{r}_t$ or the latent variable r_t^*. We can, however, as suggested by Gourieroux *et al.* (1985, 1987) define a quantity called the generalized residual $\hat{\varepsilon}_t$, which is the best prediction of the model error given the information set, the maximum likelihood estimates, and the observation at time t:

$$\hat{\varepsilon}_t = E[\varepsilon_t^* \mid r_t, \Psi_{t-1}; \hat{\boldsymbol{\theta}}], \tag{9}$$

where $\hat{\boldsymbol{\theta}} = \{\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}, \widehat{H}\}$ are the maximum likelihood estimates. With Gaussian error terms, $\hat{\varepsilon}_t$ is the expectation of a doubly truncated normally distributed variable, yielding:

$$\hat{\varepsilon}_t = \hat{\sigma}_t \frac{\varphi(\hat{c}_{1t}) - \varphi(\hat{c}_{2t})}{\Phi(\hat{c}_{2t}) - \Phi(\hat{c}_{1t})}, \tag{10}$$

with

$$\hat{c}_{1t} = \frac{1}{\hat{\sigma}_t}\left(\widehat{\alpha}_{1t} - \hat{m}_t\right) \qquad \text{and} \qquad \hat{c}_{2t} = \frac{1}{\hat{\sigma}_t}\left(\widehat{\alpha}_{2t} - \hat{m}_t\right), \tag{11}$$

where $\widehat{\alpha}_{1t}$ and $\widehat{\alpha}_{2t}$ are estimates of (8), $\hat{m}_t = \hat{\beta}_0 + \hat{\beta}_1 r_{t-1} + \hat{\beta}_2 r_{t-2}$, $\hat{\sigma}_t = \sqrt{\hat{\gamma}_0 + \hat{\gamma}_1 e_{t-1}^2 + \hat{\gamma}_2 \hat{\sigma}_{t-1}^2}$, and $e_t = r_t - \hat{m}_t$ are the observed prediction errors.

Unfortunately, the generalized residuals cannot normally be used as ordinary residuals in different diagnostic tests, see Gourieroux *et al.* (1985). Valid specification tests, often used in binary choice models, are the Likelihood ratio test, the Wald test, and the Lagrange multiplier or score test. The last test have the advantage of being carried out by using the estimates from the restricted model only, which may sometimes save the computational effort, and is crucial in testing for lagged unobservable variables. Gourieroux *et al.* (1987) show how the score test can be expressed in terms of the generalized residuals, and present numerous examples of its use.

It is obviously of great interest to test if model$_D$ is statistically equivalent to an underlying equilibrium price model like (1). Suppose that model$_D$ is assumed to be given by

$$r_t^* = \sum_{j=1}^{J} \delta_j r_{t-j}^* + m_t + \varepsilon_t^*, \qquad \varepsilon_t^* \mid \Psi_{t-1} \backsim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \sum_{k=1}^{K} \alpha_k \varepsilon_{t-k}^{*2} + \gamma_0 + \gamma_1 \varepsilon_{t-1}^2 + \gamma_2 \sigma_{t-1}^2 = \sum_{k=1}^{K} \alpha_k \left(r_{t-k}^* - \sum_{j=1}^{J} \delta_j r_{t-j-k}^* - m_{t-k} \right)^2 + \gamma_0 + \gamma_1 \varepsilon_{t-1}^2 + \gamma_2 \sigma_{t-1}^2, \tag{12}$$

with $j, k = 1, 2, ..., \mid \boldsymbol{\delta} \mid < 1$. The logarithm of the conditional density of the latent variable is

$$\ln l^* \left(r_t^* | \Psi_{t-1}; \boldsymbol{\theta} \right) \equiv \ln l_t^* = -\frac{1}{2} \ln \left(\sigma_t^2 \right) - \frac{1}{2} \frac{\varepsilon_t^{*2}}{\sigma_t^2} = -\frac{1}{2} \ln \left(\sigma_t^2 \right) - \frac{1}{2\sigma_t^2} \left(r_t^* - \sum_{j=1}^{J} \delta_j r_{t-j}^* - m_t \right)^2. \tag{13}$$

Partial differentiation of element j in $\boldsymbol{\delta}$ gives

$$\frac{\partial \ln l_t^*}{\partial \delta_j} = -\frac{1}{2\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \delta_j} + \frac{r_{t-j}^* \varepsilon_t^*}{\sigma_t^2} + \frac{\varepsilon_t^{*2}}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta_j}, \tag{14}$$

and since

$$\frac{\partial \sigma_t^2}{\partial \delta_j} = 2 \sum_k \alpha_k \frac{\partial \varepsilon_{t-k}^*}{\partial \delta_j} \varepsilon_{t-k}^* + \gamma_2 \frac{\partial \sigma_{t-1}^2}{\partial \delta_j} = -2 \sum_k \alpha_k r_{t-k-j}^* \varepsilon_{t-k}^* + \gamma_2 \frac{\partial \sigma_{t-1}^2}{\partial \delta_j} \tag{15}$$

we have that under the null hypothesis that $\boldsymbol{\delta} = \boldsymbol{\alpha} = \mathbf{0}$, $\partial \sigma_t^2 / \partial \delta_j = 0$, and the latent score increment of δ_j is

$$\left. \frac{\partial \ln l_t^*}{\partial \delta_j} \right|_{\boldsymbol{\delta} = \boldsymbol{\alpha} = \mathbf{0}, \boldsymbol{\theta} = \hat{\boldsymbol{\theta}}} \equiv \frac{\partial \ln \hat{l}_t^*}{\partial \delta_j} = \frac{r_{t-j}^* \varepsilon_t^*}{\hat{\sigma}_t^2}. \tag{16}$$

Similarly, partial differentiation of element k in $\boldsymbol{\alpha}$ yields

$$\frac{\partial \ln l_t^*}{\partial \alpha_k} = -\frac{1}{2\sigma_t^2}\frac{\partial \sigma_t^2}{\partial \alpha_k} + \frac{\varepsilon_t^{*2}}{2\sigma_t^4}\frac{\partial \sigma_t^2}{\partial \alpha_k}. \tag{17}$$

Since

$$\frac{\partial \sigma_t^2}{\partial \alpha_k} = \varepsilon_{t-k}^{*2} + \gamma_2 \frac{\partial \sigma_{t-1}^2}{\partial \alpha_k} = \sum_{i=0}^{t-k-1} \gamma_2^i \varepsilon_{t-k-i}^{*2}, \tag{18}$$

the latent score increment of α_k under $H_0 : \boldsymbol{\delta} = \boldsymbol{\alpha} = \mathbf{0}$, is

$$\frac{\partial \ln \hat{l}_t^*}{\partial \alpha_k} = \frac{1}{2\hat{\sigma}_t^2}\left(\frac{\varepsilon_t^{*2}}{\hat{\sigma}_t^2} - 1\right)\sum_{i=0}^{t-k-1} \hat{\gamma}_2^i \varepsilon_{t-k-i}^{*2}. \tag{19}$$

Louis (1982) and Gourieroux *et al.* (1985) show that the observable score is the expectation of the latent score, given the observable variables. In Appendix A, we show how this also implies that the observable score increment is the expectation of the latent score increment, conditioned on the observables. For the jth element in $\boldsymbol{\delta}$,we have

$$\frac{\partial \ln \hat{l}_t}{\partial \delta_j} = E\left[\frac{\partial \ln \hat{l}_t^*}{\partial \delta_j}\middle| \{r_u\}_{u=1}^T\right] = E\left[\frac{\partial \ln \hat{l}_t^*}{\partial \delta_j}\middle| \{r_u\}_{u=1}^t\right] = \frac{\hat{r}_{t-j}\hat{\varepsilon}_t}{\hat{\sigma}_t^2}, \tag{20}$$

where $\hat{r}_t = E[r_t^* \mid r_t, \Psi_{t-1}; \hat{\boldsymbol{\theta}}] = \hat{m}_t + \hat{\varepsilon}_t$, because r_{t-j}^* and ε_t^* are independent, given the observable data.

As for the elements in $\boldsymbol{\alpha}$ we know that given the observables, ε_t^{*2} and ε_{t-k-i}^{*2} are independent, so what remains is an expression for the conditional expectation of ε_t^{*2}. The variance of a doubly truncated normally distributed variable is given in Johnson and Kotz (1970) as

$$Var\left[\varepsilon_t^* | \Psi_t\right] = \sigma_t^2 \left(1 + \frac{c_{1t}\varphi\left(c_{1t}\right) - c_{2t}\varphi\left(c_{2t}\right)}{\Phi(c_{2t}) - \Phi(c_{1t})} - \left(\frac{\varphi(c_{1t}) - \varphi(c_{2t})}{\Phi(c_{2t}) - \Phi(c_{1t})}\right)^2\right), \tag{21}$$

yielding

$$\widehat{\varepsilon_t^2} = E\left[\varepsilon_t^{*2} | r_t, \Psi_{t-1}; \hat{\boldsymbol{\theta}}\right] = \hat{\sigma}_t^2 \left(1 + \frac{\hat{c}_{1t}\varphi\left(\hat{c}_{1t}\right) - \hat{c}_{2t}\varphi\left(\hat{c}_{2t}\right)}{\Phi(\hat{c}_{2t}) - \Phi(\hat{c}_{1t})}\right). \tag{22}$$

For the kth element in $\boldsymbol{\alpha}$, the observable score increment thus is

$$\frac{\partial \ln \hat{l}_t}{\partial \alpha_k} = \frac{1}{2\hat{\sigma}_t^2}\left(\frac{\widehat{\varepsilon_t^2}}{\hat{\sigma}_t^2} - 1\right)\sum_{i=0}^{t-k-1} \hat{\gamma}_2^i \widehat{\varepsilon_{t-k-i}^2}, \tag{23}$$

and the score test statistic for R variables in $\boldsymbol{\lambda} = [\boldsymbol{\delta}\ \boldsymbol{\alpha}]$ becomes

$$\hat{\xi}_R = \sum_{t=J+K+1}^{T} \frac{\partial \ln \hat{l}_t}{\partial \boldsymbol{\lambda}} \left[\sum_{t=J+K+1}^{T} \frac{\partial \ln \hat{l}_t}{\partial \boldsymbol{\lambda}'}\frac{\partial \ln \hat{l}_t}{\partial \boldsymbol{\lambda}}\right]^{-1} \sum_{t=J+K+1}^{T} \frac{\partial \ln \hat{l}_t}{\partial \boldsymbol{\lambda}'}, \tag{24}$$

which is asymptotically $\chi^2(R)$ under the null hypothesis.

In Table 2, we report the score statistics, $\hat{\xi}_R$, together with the corresponding p-values, in testing if the proposed model$_D$ is significantly different from some unobservable continuous-state AR(o)-GARCH(p, q) models. We see that none of the score statistics are significant at any conventional level. Since we have included two lags in model$_D$, the two left-most columns are expected to be smaller, which also is confirmed in Table 2.

The score statistics increase with the number of lagged endogenous variables, but so does the degree-of-freedom in the corresponding χ^2- distributions. Testing our model$_D$ against an AR(5)-GARCH(1, 1) model give larger statistics for ECP and IBA but we still cannot reject the null hypothesis. With this knowledge at hand we might want to test model$_D$ against a GARCH(1,1) model with *only* lag 5 in the conditional mean process. In doing so we would get a p-value of 0.03 for ECP, but this procedure would definitely be data-mining. Besides, including this lag in the mean process m_t of model$_D$ would most likely also improve the specification for ECP and, hence, reducing the score statistic. We are safe to say that in no case can we reject H_0 : model$_D$ equals a latent AR(o)-GARCH(p, q) model.

5 Summary and Conclusions

We extend the continuous-state AR-GARCH framework in order to handle situations where the dependent variable is conditionally discrete, such as would be the case in modeling returns from individual stock prices. By the use of price data from some Australian stocks with high tick size to price ratios, we discover that the parameter estimates and the asymptotic standard errors differ when comparing the modified model to the standard, but misspecified, AR-GARCH model. The differences are mainly found to be in the parameter estimates of conditional variance process, although the results often are obscured by large standard errors.

The proposed model has a well-defined likelihood function and is very easy to estimate. It is therefore of great interest to test if this model is significantly different from an underlying latent AR-GARCH model which is hidden by discrete prices. Using the concept of generalized residuals suggested by Gourieroux *et al.* (1985, 1987), we construct a score test which does exactly that. We find the strong result that in no case does the proposed model differ significantly from an unobservable continuous-state AR-GARCH model. This finding could be of great practical importance, since it shows that a switch to more exact, but also more computationally expensive, models may not be needed.

Throughout the paper we have been assuming Gaussian error terms, but see no reason why any suitable distribution could not be used. The moments of the truncated distribution may then not be explicitly given, and must be computed numerically.

References

Amilon, H. (2002), "GARCH Estimation and Discrete Stock Prices: An Applocation to Low-priced Australian Stocks", Working Paper, Quantitative Finance Research Group, University of Technology, Sydney.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroscedasticity", *Journal of Econometrics*, 31, 307-327.

Bollerslev, T., Engle, R.F., and Nelson, D. (1994), "Arch Models", in Engle, R.F. and McFadden, D. (eds.), *Handbook of Econometrics*, Vol. 4, Elsevier, Amsterdam, 2959-3038.

Chib, S. (2001), "Markov Chain Monte Carlo Methods: Computation and Inference", in Heckman, J. and Leamer, E. (eds.), *Handbook of Econometrics*, Vol. 5, Elsevier, Amsterdam, 3569-3649.

Dempster, A., Laird, N., and Rubin, D. (1977), "Maximum Likelihood From Incomplete Data via the EM Algorithm", *Journal of the Royal Statistical Society,* Ser. B, 39, 1-38.

Fiorentini, G., Sentana, E., and Shephard, N. (2002), "Likelihood-based Estimation of Latent Generalised ARCH structures", Working Paper, Nuffield College, Oxford.

Gourieroux, C., Monfort, A., Renault, E., and Trognon, A. (1987), "Generalised Residuals", *Journal of Econometrics*, 34, 5-32.

Gourieroux, C., Monfort, A., and Trognon, A. (1985), "A General Approach to Serial Correlation", *Econometric Theory*, 1, 315-340.

Hasbrouck, J. (1999), "The Dynamics of Discrete Bid and Ask Quotes", *Journal of Finance*, 54, 2109-2142.

Hausman, J., Lo, A., and MacKinlay, A. (1992), "An Ordered Probit Analysis of Transaction Stock Prices", *Journal of Financial Economics*, 31, 319-379.

Johnson, N., and Kotz, S. (1970), *Distributions in statistics: continuous univariate distributions-1*, Houghton Mifflin, Boston.

Louis, T. (1982), "Finding the Observed Information Matrix When Using the EM Algorithm", *Journal of the Royal Statistical Society,* Ser. B, 44, 226-233.

Morgan, I., and Trevor, R. (1999), "Limit Moves as Censored Observations of Equilibrium Futures Price in GARCH Processes", *Journal of Business and Economic Statistics*, 17, 397-408.

Rydberg, T., and Shephard, N. (2003), "Dynamics of Trade-by-Trade Price Movements: Decomposition and Models", Forthcoming in *Journal of Financial Econometrics.*

Wei, S. (2002), "A Censored GARCH Model of Asset Returns with Price Limits", *Journal of Empirical Finance*, 9, 197-223.

A Appendix

For convenience, let $D_{\boldsymbol{\lambda}}$ be the notation for the partial derivative with respect to a parameter vector $\boldsymbol{\lambda}$, evaluated at its ML estimate, $\hat{S}_t$ and $\hat{S}_t^*$ the observable and latent score increments at time t, respectively. The observable score increment can be written as the difference of the following observable scores:

$$\hat{S}_t = D_{\boldsymbol{\lambda}} \ln \hat{l}(r_t|r_1, ..., r_{t-1}; \hat{\boldsymbol{\lambda}}) = D_{\boldsymbol{\lambda}} \ln \hat{l}(r_1, ..., r_t; \hat{\boldsymbol{\lambda}}) - D_{\boldsymbol{\lambda}} \ln \hat{l}(r_1, ..., r_{t-1}; \hat{\boldsymbol{\lambda}}).$$

From Louis (1982) and Gourieroux *et al.* (1985) we know that the observable score is the expectation of the latent score, given the observables. Hence

$$\hat{S}_t = E\left[D_{\boldsymbol{\lambda}} \ln \hat{l}^*(r_1^*, ..., r_t^*; \hat{\boldsymbol{\lambda}})|r_1, ..., r_t\right] - E\left[D_{\boldsymbol{\lambda}} \ln \hat{l}^*(r_1^*, ..., r_{t-1}^*; \hat{\boldsymbol{\lambda}})|r_1, ..., r_{t-1}\right] = E\left[\sum_{k=1}^{t} \hat{S}_k^*|r_1, ..., r_t\right] - E\left[\sum_{k=1}^{t-1} \hat{S}_k^*|r_1, ..., r_{t-1}\right].$$

Because of the Markov property of $\hat{S}_k^*$, that is $E\left[\hat{S}_k^*|r_1,...,r_k\right] = E\left[\hat{S}_k^*|r_1,...,r_T\right]$ with $T \geq k$, we have

$$\hat{S}_t = \sum_{k=1}^{t} E\left[\hat{S}_k^*|r_1,...,r_k\right] - \sum_{k=1}^{t-1} E\left[\hat{S}_k^*|r_1,...,r_k\right] = E\left[\hat{S}_t^*|r_1,...,r_t\right].$$

Hence, the observable score increment is equal to the expectation of the latent score increment, given the observable data.

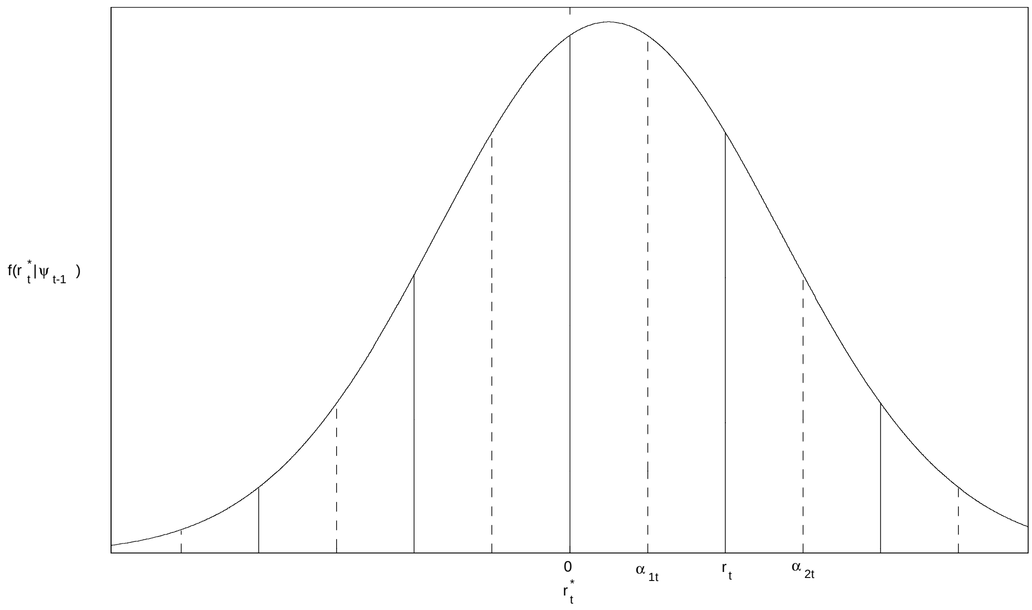

Figure 1: Shown is the conditional distribution of r_t^*. The observable states are marked with solid lines. The dashed lines partition the conditional state space of r_t^*. The boundaries comprising the realized return, r_t, are denoted α_{1t} and α_{2t}.

Table 1: ML estimates and log-likelihood values of the different models. Small numbers are asymptotic standard errors. Bold numbers indicate that the estimates from model$_D$ differ, based on their asymptotic distributions, from the point estimates of model$_C$, at the 5% significant level. Numbers marked with † indicate where model$_D$ gives opposite results than model$_C$, regarding the statistical significance, at the 5% level.

Stocks	Models	$\gamma_0 \times 10^3$	γ_1	γ_2	$\beta_0 \times 10^2$	β_1	β_2	H
DIO	model$_C$	0.457	0.227	0.359	0.037	-0.143	-0.141	
		0.189	0.098	0.200	0.228	0.081	0.080	
	model$_D$	0.100†	**0.100**	0.767†	-2.088	-0.116	-0.159	-0.135
		0.210	0.049	0.268	1.594	0.076	0.091	0.098
ECP	model$_C$	0.703	0.365	0.426	-0.620	-0.050	-0.050	
		0.261	0.113	0.127	0.354	0.104	0.094	
	model$_D$	0.677	0.344	0.447	-0.884	-0.047	-0.048	-0.162
		0.296	0.106	0.136	0.826	0.102	0.094	0.354
IBA	model$_C$	1.413	0.087	0.488	-0.026	-0.077	-0.034	
		0.504	0.033	0.033	0.049	0.022	0.034	
	model$_D$	2.332	0.128	0.163†	**-3.879**†	-0.046†	-0.034	**-0.792**
		0.967	0.062	0.313	1.795	0.082	0.061	0.370
MOS	model$_C$	0.931	0.566	0.065	-0.166	-0.469	-0.245	
		0.660	0.042	0.045	0.513	0.034	0.182	
	model$_D$	0.533	**0.356**	**0.351**†	**-1.782**†	**-0.406**	-0.224	-0.350
		3.607	0.042	0.045	0.509	0.035	0.182	0.337
MXO	model$_C$	1.434	0.254	0.000	-0.190	-0.172	-0.128	
		0.682	0.042	0.045	0.598	0.030	0.188	
	model$_D$	1.310	0.330	0.000	-1.114	-0.166†	-0.135	-0.209
		0.337	0.109	0.157	1.516	0.092	0.086	0.313
SKR	model$_C$	0.280	0.129	0.790	-0.170	-0.071	-0.112	
		0.402	0.049	0.174	0.335	0.076	0.076	
	model$_D$	0.068	**0.066**	0.921	-0.972	-0.070	-0.100	-0.076
		0.385	0.033	0.167	1.037	0.068	0.073	0.079
SMX	model$_C$	0.356	0.338	0.543	0.263	0.048	-0.053	
		0.254	0.122	0.156	0.267	0.085	0.080	
	model$_D$	0.241	0.295	0.604	-1.202	0.073	-0.057	-0.681
		0.218	0.112	0.155	1.640	0.079	0.082	0.725
YTL	model$_C$	0.153	0.103	0.847	-0.322	-0.015	-0.124	
		0.294	0.034	0.168	0.312	0.072	0.083	
	model$_D$	0.167	0.106	0.836	-0.710	-0.014	-0.123	-0.071
		0.299	0.035	0.172	0.627	0.072	0.083	0.087

Table 2: Score test statistics $\hat{\xi}_R$, as defined in (24), with $R = o+p$, in test of H_0 : model$_D$ equals a latent AR(o)-GARCH(p, q) model.

Stocks	AR(1)-GARCH(1,1)	AR(2)-GARCH(1,1)	AR(3)-GARCH(1,1)	AR(5)-GARCH(1,1)	AR(1)-GARCH(2,1)
DIO	0.271 0.873	0.284 0.963	1.389 0.962	3.051 0.802	1.566 0.667
ECP	1.007 0.604	1.827 0.609	2.452 0.653	8.053 0.234	1.445 0.695
IBA	0.077 0.962	0.069 0.995	0.292 0.990	5.632 0.466	3.734 0.292
MOS	0.710 0.701	1.569 0.666	2.233 0.693	2.366 0.883	2.507 0.474
MXO	0.114 0.945	0.507 0.917	0.512 0.972	1.841 0.934	0.769 0.857
SKR	0.164 0.921	0.156 0.984	0.912 0.923	1.389 0.967	1.957 0.581
SMX	0.647 0.724	0.708 0.871	0.921 0.922	6.317 0.389	1.588 0.662
YTL	0.171 0.918	0.366 0.947	0.789 0.940	1.223 0.976	1.155 0.764